A Preliminary Rubric Design to Evaluate Mixed Methods Research

Timothy J. Burrows

Dissertation submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
In
Educational Research and Evaluation

Elizabeth G. Creamer, Committee Chair
Penny L. Burge
Steven M. Culver
Gary E. Skaggs

March 26, 2013
Blacksburg, VA

Keywords: Mixed Methods, Evaluation, Rubric, Quality

A Preliminary Rubric Design to Evaluate Mixed Methods Research

Timothy J. Burrows

ABSTRACT

With the increase in frequency of the use of mixed methods, both in research publications and in externally funded grants there are increasing calls for a set of standards to assess the quality of mixed methods research. The purpose of this mixed methods study was to conduct a multi-phase analysis to create a preliminary rubric to evaluate mixed methods research articles. This study included four research questions:

1. What are the common evaluation criteria found in the contemporary methodological literature pertaining to the design of mixed methods research?
2. What evaluation criteria do experts in the field of mixed methods research perceive as the most important when distinguishing top-quality research in mixed methods?
3. What differences are there in the outcome of the rubric for evaluating mixed methods research identified from the literature compared to those advocated most uniformly by a panel of mixed methods research experts?
4. What are disciplinary differences between the use of mixed methods and views about evaluating it, including the role of paradigms in mixed methods research?

In the first phase of this multi-phase mixed methods study I used an inductive qualitative process to identify the quality criteria endorsed by 12 methodologists with a long-term involvement in mixed methods research. In the second phase of this study I conducted a quantitative analysis to pilot test a set of criteria identified in the qualitative phases. The sample for both phases of this study was comprised of the same eight males

and four females from multiple nationalities. Respondents to the on-line survey rated all 14 items as being important, with 11 of the 14 items being rated as very important or higher.

When considered together, findings from the two phases of this study provide a interesting view of attitudes about the use and application of quality standards to the mixed methods literature. While there was agreement about what elements were important to evaluate, there was not an agreement about the idea that one set of standards could be applied to all mixed methods studies.

Acknowledgements

I would like to thank my advisor, Dr. Elizabeth Creamer for her untiring work on assisting with this project and providing guidance throughout my doctoral studies. I would also like to thank my committee, Dr. Burge, Dr. Culver, and Dr. Skaggs for taking the time to ensure a study of the highest quality. I would also like to thank my friend Chris Altman for providing much support throughout this journey. Finally, I would like to thank everyone who assisted in one way or another in this study including Janice Austin, Molly Hall, Tom Long, Chloe Ruff, and Maria Stack.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1: Introduction

Three research methodologies dominate the fields of modern inquiry. These research methodologies include quantitative designs, qualitative designs, and mixed methods designs (Creswell, 2009). Each of these methodologies has a unique purpose that provides a specific perspective of the phenomenon being studied. "Qualitative research is a means for exploring and understanding the meaning of individuals or groups that ascribe to a social or human problem" (Creswell, 2009; p. 4). Quantitative methodologies allow the researcher to measure the orderliness and predictability of activities and decisions that occur in almost any situation (Pedhazur & Schmelkin, 1991). Mixed methods research enables the researcher to use components from both quantitative and qualitative methodologies to provide a greater understanding than just quantitative or qualitative inquiry can provide (Creswell, 2009).

Of the three research methodologies, mixed methods has been present for the least amount of time (the mid-1980s), thus having less of an opportunity for consensus of quality to be achieved as either the quantitative or qualitative methodologies. In an attempt to address this emerging method, this dissertation will examined the evaluation of mixed methods research articles as they are found in contemporary literature in hopes of leading to a more refined methodology.

Importance of Mixed Methods Research

Mixed methods research is said to provide a number of advantages. The use of mixed methods research provides both the researcher and the audience with an understanding that cannot be gleaned by utilizing only a purely quantitative or qualitative design (Creswell, 2009). By combining both quantitative and qualitative designs, validity

is enhanced because the data can be triangulated from multiple sources and methods (Greene, Caracelli, & Graham, 1989; Tashakkori & Teddlie, 2010). The use of multiple methods helps to offset the weakness of any single method (Tashakkori & Teddlie, 2010). Mixed methods research also possesses the characteristic of methodological eclecticism, or the "selecting and then synergistically integrating the most appropriate techniques from a myriad of qualitative, quantitative, and mixed methods to more thoroughly investigate a phenomenon of interest" (Tashakkori & Teddlie, 2010, p.8).

Although mixed methods research has many uses, there are situations where it is inappropriate. According to Creswell and Plano Clark, mixed methods research may not be appropriate under certain conditions, which include a lack of (1) sufficient time to analyze two types of data, (2) sufficient amounts of qualitative and quantitative data, and (3) available skills to undertake the study (Creswell & Plano Clark, 2011). Tashakkori and Teddlie add to Creswell and Plano Clark's conditions by also addressing that sometimes a purely quantitative or qualitative approach is the best method for the research question being presented (Tashakkori & Teddlie, 2010). The use of mixed methods should be driven by the need of the project and available resources.

In the social science disciplines, this adoption of a new methodology that allowed the merging of qualitative and qualitative practices gained acceptance rather quickly due to the social sciences' goals of better understanding humans and working towards improving real-world practices (Greene, 2007). Mixed methods allows for the corroboration of results from both quantitative and qualitative data, gather a greater understanding or clarification of one data type with another, and find paradoxes that may lead to a new understanding (Greene et al., 1989). The ability to accomplish such tasks

led to a greater acceptance among researchers conducting real-world evaluations (Greene, 2007). This greater acceptance of mixed methods research has led to government organizations such as the National Institutes of Health (NIH) and the National Science Foundation (NSF) to endorse the use of mixed methods (National Institutes of Health, 2012; National Science Foundation, 1997). Currently, major grants provided by organizations such as the NIH and the NSF have specific guidelines for mixed methods research proposals (National Institutes of Health, 2012; National Science Foundation, 1997).

Since the mid 1980s, mixed methods research articles have become more prevalent in research journals across a wide variety of disciplines. In 1989, Greene et al. conducted a content analysis of mixed methods research articles published between 1980 and 1988 (Greene et al., 1989). This comprehensive study included a sample of 57 empirical mixed methods evaluations (Greene et al., 1989). In 2006, Allen Bryman conducted another content analysis of mixed methods research and was able to use a sample of 262 articles (Bryman, 2006). One must assume that, in order to for Bryman to develop a larger sample size, the overall population must have also increased. Even though the overall number of mixed methods research publications is increasing at a steady rate, the quality and application of mixed methods research is largely undefined by the literature (Bryman, 2006, Greene et al., 1989; Gambrel & Butler, 2011).

Definition of Mixed Methods Research

Prior to examining the challenges found within the literature, we must first define exactly what mixed methods research is. Simply stated, mixed methods research is an "approach to inquiry that combines or associates both qualitative and quantitative forms"

(Creswell, 2009, p. 4). Unfortunately mixed methods research does not have a single definition, nor one that is simple. Mixed methods research has many formal definitions as defined by various researchers including, but not limited to Greene et al. (1989), Creswell and Plano Clark (2011), and Johnson, Onwuegbuzie, and Turner (2007). Definitions vary by important methodological issues such as the emphasis given to mixing, or the interrelating of a mixed methods study's qualitative and quantitative parts, whether it can be one or multiple linked studies, and the priority given to having been initially designed as a mixed methods study (Creswell & Plano Clark, 2011). Each of these individual definitions was examined so a note can be taken of similarities and differences found between the definitions offered by different researchers. Greene et al. (1989) defined mixed methods research as:

> We defined mixed-methods designs as those that include at least one quantitative method (designed to collect numbers) and one qualitative method (designed to collect words), where neither type of method is inherently linked to any particular inquiry paradigm (Greene et al., 1989, p. 256).

The focus of the definition by Greene et al. was primarily on the mixing of methods and the philosophies that traditionally bound them (Creswell & Plano Clark, 2011). Greene et al. may have focused on inquiry paradigms to set mixed methods research apart from previous methods following the paradigm wars of this time period.

Following Greene et al., Creswell and Plano Clark produced their own definition of mixed methods research, one that included both a methodological and philosophical orientation (Creswell & Plano Clark, 2011). They stated:

> Mixed methods research is a research design with philosophical assumptions as well as methods of inquiry. As a methodology, it involves philosophical assumptions that guide the direction of the collection of and analysis and the mixture of qualitative and quantitative approaches in many phases of the research process. As a method, it focuses on collecting, analyzing, and mixing both quantitative and qualitative data in a single study or series of studies. Its central premise is that the use of quantitative and qualitative approaches, in combination, provides a better understanding of research problems than either approach alone (Creswell & Plano Clark, 2007, p.5).

Creswell's and Plano Clark's definition contains components of the Greene et al. definition including mixing and philosophies, but adds more precise language about the purpose for utilizing mixed methods research. The idea of mixing occurring in more than one phase of the research points to the importance of data integration. In addition to the level of mixing, the purpose for providing a better understanding of the phenomena being studied also sets this definition apart from Greene et al.'s (1989).

In the most recent edition of the *Sage Handbook of Mixed Methods in Social Research & Behavioral Research* (2010), Tashakkori and Teddlie provide an additional definition of mixed methods research. Tashakkori and Teddlie use Johnson's definition of mixed methods research in the handbook. Johnson et al. defines mixed methods as, "the type of research in which a researcher or team of researchers combines elements of qualitative and quantitative research approaches for the broad purposes of breadth and depth of understanding and corroboration" (Johnson, Onwuegbuzie, & Turner, 2007;

Tashakkori and Teddlie, 2010, p.19). Johnson simplified the definition, but like Creswell and Plano Clark, addresses the purpose. Johnson also includes the combination of qualitative and quantitative methods in the study, but does not include an amount pertaining to the mixing of data.

When reading mixed methods literature, “multiple methods” is another term that is found. Multiple methods research is sometimes used in lieu of mixed methods, but has its own meaning. Tashakkori and Teddlie define multiple methods as "two or more studies using different methods, which address the same research question or different parts of the same research question or programmatic goals" (Tashakkori & Teddlie, 2010, p. 340). The absence of a reference to mixing distinguishes multi-method from the definitions provided by Creswell and Plano Clark and Johnson. A study is not a mixed methods study just because of the presence of two methods.

The previous three definitions are just a few of the many definitions that can be found in the mixed methods research literature. The purpose of reviewing these definitions is not to say one is better than the other, but to examine the subtle ways definitions vary. The definition of mixed methods research is evolving, just as the field itself is evolving. Greene et al.'s definition focused on mixing and philosophy, while Creswell and Plano Clark added a purpose for the research.

For this study, I will adopt the definition supplied by Creswell and Plano Clark (2010). I use this definition because it contains components of all the definitions listed above, including paradigms, mixing, and purpose. To follow the definition of Creswell and Plano Clark, certain assumptions must be made to carry out a mixed methods research study in the ways of philosophy and methods.

Assumptions of Mixed Methods Research

As with other research methodologies, assumptions are a critical part of mixed methods research (Howell, 2010; Patton, 2002). Three assumptions that impact a mixed methods research study that are (1) quantitative and qualitative data are compatible, (2) that all data have both quantitative and qualitative properties, and (3) that the pragmatic or constructive paradigms are primarily used. When these assumptions can be met, the mixed methods methodology becomes a viable option available to the researcher (Creswell & Plano Clark, 2011; Green et al., 1989; Maxwell & Mittapalli, 2010; Sandelowski, Voils, & Knafl, 2009).

Based on Creswell and Plano Clark's and Greene et al.'s definitions, perhaps the most important assumption is that quantitative and qualitative data can be used in combination (Creswell & Plano Clark, 2011; Greene et al., 1989). The combining of data is what ultimately leads to the ability for mixing, quantitizing, and qualitizing data[1]. This combining of data was not accepted during the paradigm wars of the 1970s and 1980s. During this paradigm war, qualitative researchers rejected the idea that numerical data was compatible with qualitative data (Maxwell, 2010). These researchers who take a hard-line approach to paradigms such as constructivism and post-structuralism may find it difficult to believe that quantitative and qualitative data are so similar (Maxwell, 2010)[2]. To make quantitative and qualitative data compatible and able to be mixed, the line between the types of data must be seen as commensurable (Sandelowski et al., 2009). This commensurability was seen as early as 1970, when Becker stated that qualitative researchers used terms such as *some*, *many*, and *most*, thus implying a quantity within the

[1] Mixing of data will be further explained in the literature review.
[2] The definition and use of applicable paradigms to this research will be discussed in more detail further on in the literature review portion of this paper.

qualitative data (Becker, 1970). Nearly 25 years later, Miles and Huberman (1994) added to Becker's earlier statement by claiming all data is either a "1" or a "0" (Sandelowski et al., 2009). By following the assumption that all data are commensurable and that it refers to the presence or absence of something being measured, a researcher can utilize both quantitative and qualitative data in a manner to combine the data, resulting (as previously stated) in triangulation of data, offsetting bias, and/or adding validity to the study (Greene et al., 1989).

The second assumption inherent in mixed methods research is linked to the first in the sense that it is required to accept the permeability of data. The assumption that all data have both quantitative and qualitative properties is based on Wolcott's (1994) premise of needing to communicate in order to collect and report all data (Sandelowski et al., 2009). Wolcott stated that "everything has the potential to be data, but nothing becomes data without the intervention of a researcher who takes note of some things to the exclusion of others" (Wolcott, 1994, p. 3-4). Wolcott's statement suggests that the data are always collected through qualitative means and then is transformed into a product that is designed to capture numeric data. By assuming that all data goes through this process of transformation, the researcher can then assume that data can be mixed together since they ultimately are one and the same (Sandelowski et al., 2009).

The final assumption that will be addressed in this study is that of paradigm use. Mixed methods research is primarily steeped in the pragmatist paradigm or the belief that the act of discovery outweighs the justification of knowledge (Creswell & Plano Clark, 2011; Rossman & Wilson, 1985; Small, 2011; Tashakkori & Teddlie, 2003). As with the definition of mixed methods research, the use of paradigms is also evolving, with the

recent discussion of the addition of realism as another appropriate possible paradigm (Maxwell & Mittapalli, 2010). Paradigms and their relation to mixed methods research will be discussed in further detail in the literature review section of this study.

Assumptions often pose problems to mixed methods research since they are often at odds with traditional methods. Since this methodology is still a burgeoning field, it possesses many hurdles that pose great challenges to the researcher and the reader (Creswell & Plano Clark, 2011; Gambrel & Butler, 2011; Greene et al, 1989; Maxwell, 2010).

Challenges Associated With Mixed Methods Research

The list of challenges that confront a mixed methods methodologist can be long and daunting. Since mixed methods research is a field that has only been around for the better part of a quarter century, there are plenty of growing pains that accompany its maturation (Small, 2011). Some of the challenges that a researcher can expect to see in mixed methods research are different terms, various reporting procedures, unique or unknown designs, and a lack of how quality is evaluated to name just a few (Bryman, 2007; Caracelli & Greene, 1993; Small, 2011). In this section, I will examine several known challenges and how they impact current practices in mixed methods research.

As an emerging field, mixed methods research finds itself in an ever changing terminological environment (Creswell & Plano Clark, 2011; Greene et al, 1989; Tashakkori & Teddlie, 2010). As a researcher reads mixed methods articles, they will notice very quickly that many terms are used with few of them being constants or even present at all (Gambrel & Butler, 2011). The conflict of terms occurs at all levels, from the term mixed methods itself all the way to the types of designs (Creswell & Plano

Clark, 2007). Some researchers refer to mixed methods as multi-method or multiple methods, although a clarification was given for the three by Creswell and Plano Clark (2007). The terms for research designs also vary greatly, depending on which author and edition of text the researcher prefers (Creswell & Plano Clark, 2007; Creswell & Plano Clark, 2011; Greene et al, 1989; Tashakkori & Teddlie, 2010).

Along with the lack of consensus on terminology, the methods for reporting mixed methods studies is also highly variable (Bryman, 2006; Engelbrecht, Harding, & DuPreez, 2007; Tully & Jacobs, 2010). Some researchers include both quantitative and qualitative findings in their publications, while others only report on one or the other methodology (Bryman, 2006; Engelbracht et al., 2007). While some researchers report on one method or another, others choose to include findings that are mixed (Kilgore, Jocuns, Yasuhara, & Atman, 2010). The lack of absolutes leaves the researcher free to interpret and use these terms freely, causing a wide array of reporting procedures in the literature.

When reading the mixed methods research literature, the different use of designs or the altogether lack of designs poses an additional challenge to the reader (Bryman, 2007). As stated previously, the terminology of designs relies on the various authors of mixed methods research designs (Creswell & Plano Clark, 2007; Creswell & Plano Clark, 2011; Greene et al, 1989; Tashakkori & Teddlie, 2010). Mixed methods research articles may use the design sets created by Greene et al. (1989) and mention designs such as triangulation or initiation or use the designs set forth by Creswell and Plano Clark (2011). Creswell's and Plano Clark's view's about the number of unique designs for mixed methods studies has evolved as well. The original group of four designs reported in the first volume of their textbook has been replaced with six designs in the most recent

edition of the textbook. Often authors omit references to a design in their publications, leaving the reader to wonder about the purpose and layout of the study that is being read (Sagebiel, 2006; Smith & Degniz, 2010).

With a lack of designs often comes a lack of mixing data, a component critical to the definition of mixed methods research used within this study (Raju, Sankar, and Xue, 2004; Sagebiel, 2006; Schultz & Christenses, 2004; Smith & Dengiz 2010). Mixing can occur in all or one of the following four places; (1) the research question, (2) data collection, (3) analysis of data, and (4) the interpretation of results (Creswell & Plano Clark, 2011).

When considering the amount of mixing and the quality of mixing, another challenge quickly arises. As stated previously, mixing is a component of multiple definitions and guidelines are given for when to mix and how often to mix, but no standard has been presented about how to judge the quality of mixing or in what stage of the research it should occur (Creswell & Plano Clark, 2011; Greene et al., 1989). Qualitative research includes methods to determine rigor through the use of saturation, triangulation, reflexivity, peer review, and negative case studies to name a few (Patton, 2002). Quantitative research has standards to ensure construct validity through rigorous procedures (Messick, 1995). Mixed methods research has the challenge of prescribing quality of mixing because it combines both quantitative and qualitative methods (Creswell & Plano Clark, 2011).

Additional challenges include two controversial types of mixing. These two types of mixing are data transformation, known as quantitizing and qualitizing (Tashakkori & Teddlie, 2010). Quantitizing is the transformation of qualitative data into quantitative

data (Sandelowski et al., 2009). Qualitizing is the transformation of qualitative data into quantitative data (Tashakkori & Teddlie, 2010). The ability to transform data in this fashion is related to the researcher's ability to accept the assumption that data is permeable and can be transposed upon one another (Sandelowski, 2001; Sandelowski et al., 2009). When transforming data, both quantitizing and qualitizing data contain specific challenges. When quantitizing data, certain assumptions may be violated when processing the statistical analyses due to sample sizes (Pedhazer & Schmelkin, 1991). Quantitizing also presents a challenge to the researcher by making the researcher choose how to assign values to qualitative data without absolutes, or a given description of what exactly the participant meant by their response. An subjective interpretation of data are required by the researcher. By placing a value on qualitative data, specific details may be lost. Similar challenges to those found when quantitizing also occur when qualitizing data. Qualitizing imposes a challenge of the researcher interpreting numerical values. Interpretations may vary or be misrepresented because of the instrument used to collect the quantitative data. Quantitizing and qualitizing are not exact procedures and when choosing to transform data, the researchers must determine whether they want to substitute the richness of qualitative data for the simplicity of quantitative data or vice versa (Sandelowski et al., 2009).

Another set of questions and challenges to mixed methods research is about sample size. Qualitative research often uses a small sample to collect a large amount of thick and rich descriptions about the topic being researched (Patton, 2002). Quantitative research, on the other hand, generally uses larger sample sizes based on the amount of variables and the total population size to often adhere to its own assumption of a

standardized distribution (Pedhazer & Schmelkin, 1991). The process of collecting quantitative data and conducting statistical analyses can be skewed if a small sample is used, which is often the case if the mixed methods research is based on a small qualitative sample (Howell, 2010). The opposite challenge can arise if the sample is based on a quantitative measure such as a survey with open ended questions. The researcher may end up with thousands of cases to code and interpret for qualitative analyses. This is another challenge that has not been addressed with a definitive answer, leaving the researcher to make his or her own choice in the matter. Another challenge associated with sample size is that it too is also used as a base for quality in the other methodologies (Patton, 2002; Pedhazer & Schmelkin, 1991). Mario Small adds some clarification to this challenge by explaining that as a minimum, "the application of techniques should not be contrary to the epistemological principles from which they are derived or to the technical problems for which they were intended" (Small, 2011, p.76).

This lack of consensus in the quality of mixing and sample size opens the debate on how to evaluate the overall quality of a mixed methods research study. As with mixing, the overall tendencies towards quality are left up to the researcher's own understanding of the mixed methods. Creswell and Plano Clark (2011) give guidelines for necessary components for a mixed methods study as do federal organizations such as the NIH and NSF (Creswell & Plano Clark, 2011; National Institutes of Health, 2012; National Science Foundation, 1997). Certain journals such as the *Journal of Mixed Methods* also give guidelines of what is considered to be a mixed methods study (Journal of Mixed Methods Research, 2011). The *Journal of Mixed Methods Research* explicitly requires that manuscripts integrate (mixing) and fit the definition of mixed methods

research by mixing the findings and drawing inferences using both qualitative and quantitative methods (Journal of Mixed Methods Research, 2011). The *Journal of Mixed Methods Research*, as with the other topics mentioned previously, provide only guidelines, not adopted standards. Guidelines are given to provide suggestions for the researcher, whereas standards provide the researcher specific objectives to accomplish. The lack of adopted standards leaves mixed methods methodologists in an uncertain position when trying not only to create a study, but also when trying to evaluate the quality of a mixed methods study.

A final challenge in mixed methods is bridging the gap between the theory and practice of using mixed methods research. A challenge that presents itself when comparing theory to practice is that the information found in textbooks is not always reflected in research publications. Several publications, including content analyses of mixed methods research have commented on the absence of references to standard mixed methods texts, as well as the lack of consistency in references to design names (Bryman, 2006; Gambrel & Butler, 2011; Greene et al., 1989; Plano Clark, Huddleston-Casas, Churchill, O'Neil Greene, & Garrett, 2008; Small, 2011). Examples of the gap between theory and practice can be found in the use of designs, purpose statements, and research questions to name a few (Creswell & Plano Clark, 2011; Kilgore et al., 2010; Lou, Liu, Chaung, & Tseng, 2011). The time lag between publications of the textbooks and the diffusion of the ideas into the field is one explanation for the gap between theory and practice. Time and resources such as compressing so much information within the confines of a journal article may also cause difficulties for the researcher to close the gap between theory and practice. The lack of time may prevent optimal research designs,

while resources such as article page limits may require the researcher to leave out select information.

To say the least, there are many challenges that confront researchers in mixed methods research. Researchers must contend with an ever changing vocabulary, a lack of agreement on how mixed methods data should be formally reported, which research design to use, what sample size is necessary, and how to ensure quality (Bryman, 2006; Caracelli & Greene, 1993; Creswell & Plano Clark, 2011; Gambrel & Butler, 2011). All of these challenges are connected to a fledgling methodology and may continue until a consensus is reached. How long until this happens, one can only guess. As noted by astrophysicist and science spokesperson, Neil DeGrasse Tyson, it only took the astronomy community 2,000 years to define what a planet is (Tyson, 2009). Universal criteria for defining quality in mixed methods research will assist in the furthering of the mixed methods methodology.

Need For One Universal Criteria for the Determining Quality of Mixed Methods Research

I will pose a triple-pronged argument to demonstrate the need for universal criteria for determining quality mixed methods research publications. The three points to this argument are (1) a summary of key results of previous mixed methods research content analyses, (2) what is missing from these content analyses, and (3) how an additional content analysis will aid in building consensus among mixed methods research.

Several key results emerged from a meta analysis of four content analyses of the mixed methods research literature (Bryman, 2006; Gambrel & Butler, 2011; Greene et

al., 1989; Plano Clark et al., 2008). These content analyses provide insight into the state of the literature in mixed methods. Greene et al. (1989) created a framework of five mixed methods designs to demonstrate a purpose for using mixed methods evaluations (Greene et al., 1989). Along with these purposes, Greene et al. (1989) also stated that many terms, such as triangulation were misused and many studies failed to integrate quantitative and qualitative methods. In 2007, Bryman added to Greene et al. (1989) by finding mixed methods research publications tended to be predominantly quantitative and that the purpose for these studies did not always correspond to the practices utilized (Bryman, 2007). In the 18 years following Greene et al., Bryman found that mixed methods research was still a practice without much standardization (Bryman, 2007). Following Bryman, Plano Clark et al. (2008) also found that quantitative methods predominated as the method for collecting data. The methods for collecting data generally consisted of surveys with open ended questions (Plano Clark et al., 2008). In addition to limited methods for collecting data, the use of mixed methods terminology was also very limited and not standardized (Plano Clark et al., 2008).

Most recently, Gambrel and Butler (2011) supported the previous content analyses and offered new insight in mixed methods research articles. Gambrel and Butler (2011) found that many of the articles that they reviewed in the field of marriage and family therapy did not give a concise purpose for using mixed methods research, nor did they define how the data were mixed (Gambrel & Butler, 2011). Like Bryman (2007) and Plano Clark et al. (2008), Gambrel and Butler (2011) also found that the majority of their articles were quantitatively focused. Butler and Gambrel (2011) added to the previous content analyses by including the lack of philosophical support for using mixed methods

research. The results of the four content analyses demonstrate that there continues to be many inconsistencies in mixed methods research. These content analyses show that among mixed methods research articles, there seems to be very little consensus when dealing with terminology, research designs, and support from the literature (Gill, Sharp, Mills, & Franzway, 2008; Larson & Ahonen, 2004; Lou, Liu, Chang, & Tseng, 2011).

While previous content analyses demonstrate a wide gap in practice found in the literature, several additional problems were not addressed. Of the four content analyses, two of these contained articles from single disciplines (e.g. Gambrel and Butler (2011) and Plano Clark et al. (2008). In addition to discipline, with the exception of Bryman (2007), a relatively small sample size of less than 60 was used by Bryman (2006), Gambrel & Butler (2011), Greene et al. (1989), and Plano Clark et al. (2008). Perhaps the largest problem with these content analyses is that most are dated and do not capture the burgeoning number of publications and changes in the vocabulary (Bryman, 2007; Greene et al, 1989; Plano Clark et al., 2008).

The analysis I have proposed will address the findings of the previous content analyses as well as topics that were not discussed previously. By bridging the gaps and connections of previous articles, a new content analysis can aid in building a consensus among mixed methods research, which in turn could advance standardization. The content analysis undertaken in this dissertation will attempt to advance standardization by aiming to be a point of convergence, something not previously seen in the literature. Available content analyses provide many links in a chain of events describing the evolution of mixed methods literature, but there has yet to be a link to solidify this chain.

It is clear from several content analyses that the variation of reported literature is as broad as the literature from which researchers learn their craft (Bryman, 2006; Gambrel & Butler, 2011; Greene et al., 1989; Plano Clark et al., 2008; Small, 2011). A standardization of terminology and procedure will assist to reduce the variations found within the reporting of mixed methods research (Bryman, 2006; Gambrel & Butler, 2011). A standard procedure should include a common vocabulary, design typology, research question and purpose statement, and appropriate level of mixing (Bryman, 2006; Bryman, 2007; Green et al., 1989; Plano Clark et al., 2008). Having a level of standardization is imperative for promoting a level of quality found in mixed methods research articles (Fitzpatrick, Sanders, & Worthen, 2011).

Research Purpose and Questions

The purpose of this mixed methods study is to conduct a cross-disciplinary, multi-phase content analysis to create a preliminary rubric to evaluate mixed methods research articles in multiple disciplinary fields. The four research questions that guide this study are:

1. What are the common evaluation criteria found in the contemporary methodological literature pertaining to the design of mixed methods research? (Qualitative & Quantitative)
2. What evaluation criteria do experts in the field of mixed methods research perceive as the most important when distinguishing top-quality research in mixed methods? (Qualitative)
3. What differences are there in the outcome of the rubric for evaluating mixed methods research identified from the literature compared to those advocated most

uniformly by a panel of mixed methods research experts? (Qualitative & Quantitative)

4. What are disciplinary differences between the use of mixed methods and views about evaluating it, including the role of paradigms in mixed methods research? (Qualitative)

The timing of this study is sequential with a quantitative prioritization. The research design is multi-phase (qual → QUAN). The formal description of the design for this study will be presented in the methods section.

Chapter 2: Literature Review

As mixed methods research becomes a more popular choice among researchers, a new method of evaluating quality is imperative for researchers to work collectively under an umbrella of universal understanding (Bryman, 2007; Gambrel & Butler, 2011). In this chapter, several topics will be addressed to provide a background for developing criteria to evaluate quality in mixed methods research. The principles of evaluation along with the purpose for rubrics guide this discussion through the various topics including current research designs, how data are mixed, the role of paradigms, and the possible findings from this research. The principles of evaluation are important to understand because they drive the methods used in this study. An in-depth look at evaluation is necessary due to the often misunderstood or misused meaning of the topic, especially when addressed in mixed methods research articles (Creswell & Plano Clark, 2011). Following the review of evaluation, examples of evaluation in current mixed methods publications are examined to provide evidence for the current state of evaluation of mixed methods research. To evaluate the quality of mixed methods research, several principle components of mixed methods research must also be identified. This research study will use both theory and practice to create a rubric that is practically useful and grounded in current mixed methods and evaluation theory.

The first section of this literature review will focus on the definition, purpose, and role in mixed methods research. The first component to be identified is the types of research designs currently found in mixed methods research studies. The different types of research designs pose separate considerations when considering how to evaluate their quality. The second component of mixed methods research to address is mixing, since it

is often described as the feature to define a mixed methods study (Creswell & Plano Clark, 2011; Greene et al., 1989). How a mixed methods research study is mixed leads to a required flexibility in the evaluation of quality. The final piece of the mixed methods process to be discussed in this literature review is the use of paradigms. Paradigms create a worldview for which the study is created within, which may affect the research purpose and design. The use of different research designs, types of mixing, and paradigms must all be taken into account when creating a method of evaluation, since one method is not determined superior to another. The possible findings presented in this section reflect the different ways in which the above components of mixed methods research may affect how the quality of mixed methods research is evaluated in the future.

Evaluation Criteria

Evaluation, like mixed methods research itself, has many definitions that have been evolving over the past 40 years (Mark, Henry, & Julnes, 2000; Schwandt, 2008; Scriven, 1967). These definitions tend to vary between practitioners and disciplines (Fitzpatrick, Sanders, & Worthen, 2011). The definition that is used for this study comes from Fitzpatrick et al. (2011). Fitzpatrick et al., define evaluation as "the identification, clarification, and the application of defensible criteria to determine an evaluation object's value (worth or merit) in relation to those criteria" (Fitzpatrick et al., 2011, p. 7). By properly evaluating mixed methods research, we hope to clarify the means that we are using to judge future studies and to overcome the failure to identify and clarify the components that are viewed as necessary (Fitzpatrick et al., 2011).

Evaluation uses three methods of inquiry and judgment to optimize the object that is being evaluated and to help stakeholders determine the object's value. The three

methods include (1) "determining the criteria and standards for judging quality and deciding whether those standards should be relative or absolute, (2) collecting relevant information, and (3) applying the standards to determine value, quality, utility, effectiveness, or significance" (Fitzpatrick et al., 2011, p. 7).

As with any evaluation, there are stakeholders who must be identified. The stakeholders for this evaluation are categorized by Greene and Henry in the *Encyclopedia of Evaluation* (2005) as people who have direct responsibility for that which is being evaluated (Greene & Henry, 2005). In this instance, the stakeholders are the researchers and practitioners who directly involve themselves with mixed methods research. It is these stakeholders who have a stake in the future direction of this methodology.

With a definition of evaluation in place and a population of who the stakeholders are, the next topic to be discussed is the purpose for conducting an evaluation. Following with the definition presented earlier, the primary purpose for conducting an evaluation is to judge on the value of what is being evaluated, which is mixed methods in this case (Fitzpatrick et al., 2011). This purpose also mirrors Scriven's (1996) purpose for conducting evaluations, which involves emphasizing the worth of the object in question. In addition to Fitzpatrick et al. (2011) and Scriven (1996), this evaluation also follows Talmage's (1982) inclusion that evaluations assist in the decision making process for those responsible for creating policies. To summarize the points made by Fitzpatrick et al. (2011), Scriven (1996), and Talmage (1982), the evaluation process of this study is not only designed to assign a value to the research that is presented in mixed methods, but also to guide those who lead the field in its advancement and growth. A rubric will be created to provide a system for judging quality of mixed methods research.

A rubric is a document that articulates the expectations for the object being evaluated by listing the critical criteria of what is deemed as necessary and assesses the levels of quality from poor to excellent (Reddy & Andrade, 2010). Reddy and Andrade (2010) state that rubrics tend to have three principle components: (1) the criteria that are being used to evaluate the object, (2) definitions of quality, and (3) a scoring strategy to measure the overall quality of the object. When creating a rubric, of the three components presented, perhaps the most important component is two: definition of quality. By providing quality definitions of what is expected, the evaluator will have an easier time making a clear and precise judgment (Reddy & Andrade, 2010). By creating a clear rubric to measure quality, mixed methods research studies take a significant step in having a broader impact in understanding its research as a whole (Sevian & Gonsalves, 2008).

Rubrics assist accountability by providing a measure to which works are compared. By creating an object that serves as an example of quality and using a scale from poor to excellent, all the stakeholders are able to make decisions based on a standard and not just selected literature (Fitzpatrick et al., 2011). Current assessments of quality of the mixed methods research publications are based on guidelines set forth in the literature, such as the NIH guidelines (2012), but as of this date, no current rubric takes into account the combination of literature and a scale of assigned value for the individual components. By applying the three components listed previously to a rubric and combining criteria found in previous literature, the rubric will allow researchers to create an atmosphere of accountability.

Accounting for the quality of mixed methods research is difficult due to the many defining characteristics of mixed methods research, such as mixing and design types. Currently, there are variety of viewpoints about what a mixed methods research study should include and what its purpose should be (Creswell & Plano Clark, 2011; Greene, 2007; Tashakkori & Teddlie, 2010). Although researchers have given several purposes for mixed methods research and guidelines for what makes a mixed methods study, they all lack a way to evaluate the quality of the study. Greene (2007), Creswell and Plano Clark (2011), and Tashakkori and Teddlie (2010) have each laid out their own interpretations of the dimensions that compose a mixed methods research study, but neglect to propose a way to evaluate them.

In her book, *Mixed Methods in Social Inquiry* (2007), Greene proposed seven characteristics of mixed methods research: (1) paradigms, (2) phenomena, (3) methods, (4) status, (5) implementation: independence, (6) implementation: timing, and (7) study. *Paradigms* refers to which paradigms are used and how they are implemented in the study (Greene, 2007). *Phenomena* refers to the degree in which qualitative and quantitative methods are used to assess the phenomena being studied. *Methods* is how the various selected methods are similar to and different from one another. *Status* represents the weighting of the qualitative and quantitative components of the study. *Implementation: independence* is the degree to which the selected methods are utilized and designed. *Implementation timing* refers to when the qualitative and quantitative methods occur in the study. Finally, the *study* represents the category of research in which the overall study fits. Greene's characteristics are described as applicable, but not

as absolutes, meaning that they may all apply to a mixed methods research study, but are not mandatory in its design.

Creswell and Plano Clark provide another set of characteristics of mixed methods research in their publication, *Designing and Conducting Mixed Methods Research* (Creswell & Plano Clark, 2011). Creswell and Plano Clark offer six characteristics that define a mixed methods research study. The six characteristics are (1) collecting both qualitative and quantitative data, (2) using rigorous methods for data collection and analysis, (3) mixing the data, (4) utilizing a mixed methods design, (5) conducting the study within a paradigm or worldview, and (6) reporting the findings using vocabulary consistent in the present literature (Creswell & Plano Clark, 2011).

Tashakkori and Teddlie (2010) follow Greene (2007) and Creswell and Plano Clark (2011) with the most comprehensive list of characteristics with nine in total. The first characteristic that Tashakkori and Teddlie set forth is methodological eclecticism, or the idea that the best method should be selected for the desired outcomes (Tashakkori & Teddlie, 2010). The second characteristic is paradigm pluralism, the belief that various paradigms may best serve the researcher in answering the research question(s). "The third characteristic of contemporary (mixed methods) research is an emphasis on diversity at all levels of the research", from the conceptual to the empirical (p. 9). The fourth characteristic is the emphasis on a continuum for analyzing the data rather than an either-or approach. The inclusion of both inductive and deductive logic in the same study is the fifth characteristic. The sixth characteristic is using a research question to guide the methods employed in the study. The inclusion of a specified research design is the seventh characteristic. The eighth characteristic is balance and compromise in the mixed

methods research community. The final characteristic includes the use of "visual representations and a common notation system" in the research.

When reviewing the characteristics of mixed methods research, several similarities and differences become apparent. Greene (2007), Creswell and Plano Clark (2011), and Tashakkori and Teddlie (2010) are similar in the regards to the presence of a paradigm and the collection of both qualitative and quantitative data. Although there are similarities between the selected researchers, the differences outnumber the similarities. Tashakkori and Teddlie (2010) do not explicitly refer to mixing, while Greene (2007) and Creswell and Plano Clark (2011) do not address the mixed methods community nor diversity of the research. Greene (2007) differs from Creswell and Plano Clark (2011) and Tashakkori and Teddlie (2010) by not including any discussion of design in her characteristics of mixed methods research. These similarities and differences of characteristics point to differences in the rubrics each might create to evaluate the quality of a mixed methods research article.

By examining the characteristics that define mixed methods research, a focus can then be placed on the criteria for defining quality in mixed methods research. An exact measure of quality in mixed methods research currently eludes researchers, but a semblance of quality may be derived from several criteria set forth to review mixed methods research. Five pieces of literature will aid in the discussion of the current state of evaluation in mixed methods research found in the literature.

Examples of Evaluation Criteria in Mixed Methods Research

Six examples of evaluation criteria for mixed methods research stand out in current literature pertaining to mixed methods research. These six sources include

articles, a textbook, and government grant reporting procedures. Similarities and differences in these sources create a broad picture of the present state of the evaluation of mixed methods research.

The first source of guidelines for evaluating mixed methods research is found in Creswell and Plano Clark's textbook, *Designing and Conducting Mixed Methods Research* (2011). Creswell and Plano Clark lay out four principles used in the evaluation process to judge the quality of mixed methods research. First, Creswell and Plano Clark examine the methods section, looking for both qualitative and quantitative data used in the response to research questions or hypotheses. The second step in the evaluation process according to Creswell and Plano Clark is a further look at the methods section to examine for rigor and persuasiveness, depending on the type of methods used. The third criterion for evaluating a mixed methods study is how the mixing of data was completed and where the mixing occurred. Finally, the use of mixed methods terminology was sought. Appropriate terms should be found in the title, research questions, and throughout the study in relations to design (Creswell & Plano Clark, 2011). Creswell and Plano Clark do not provide a rubric for assigning an actual value to the criteria.

The second source of guidelines for evaluating mixed methods research comes from the mixed methods research application review standards (2012) for grants recently offered by the NIH. The evaluation criteria set forth by the NIH is the most comprehensive set of criteria, pulling from several different sources, including the ones presented within this literature review such as Creswell and Plano Clark. NIH also presents four criteria for evaluating mixed methods research applications (National Institutes of Health, 2012). The first criterion is to determine if evidence is given for both

qualitative and quantitative procedures. The second criterion is to ensure a body of literature is presented to support the components that are found in the study. The third criterion is to consult existing mixed methods research quality criteria. The different quality criteria may include methods orientation, research process orientation, and the timing of phases found in the investigation orientation. The final criterion is a checklist for reviewing the application of mixed methods research procedures. This checklist may include the significance of the study, information about the researcher's skills and experience, the innovation of the study, the approach, and the environment (National Institutes of Health, 2012). Unlike Creswell and Plano Clark's book, the NIH guidelines provide a way to score the research. NIH scores the significance of the study, the investigators participating in the study, the innovation of the study, the approach taken to conduct the study, and environment of the study on a nine point scale. Studies with the higher score are seen as having a higher quality than those with a lower score. NIH heavily relies on the current body of literature, including Creswell and Plano Clark, but adheres to the definition of evaluation by offering an evaluation process that adds a value to the research presented.

The third source of mixed methods research evaluation guidelines is found in *Evaluating Mixed Methods Research Studies: A Mixed Methods Approach* (2009), by Leech, Dellinger, Brannagan, and Tanaka. Leech et al., take a different approach to evaluation by focusing on a validation framework. The five components of the validation framework include the foundational element, the elements of construct validation, inferential consistency, the utilization/historical element, and the consequential element (Leech et al., 2009). The first of these, the "foundational element provides evidence of

construct validation through evaluation of relevant research that may offer insight into the meaning of data, interpretation of data, and consequences of the use of data" (p. 19). The elements of construct validation use traditional qualitative and quantitative construct validation methods, but are not limited to traditional means. "Inferential consistency refers to how inferences from a study are logically consistent with other elements of the study" (p. 21). The utilization/historical element refers to how the use of a research method is supported by previous studies in the use of similar data. The consequential element examines the consequences of how the data are used and the inferences made from the data. Leech et al. have created a unique set of criteria to evaluate mixed methods research from the standpoint of content validity. Content validity is used to determine if the inferences made within the study are consistent as well as to determine if the data are relevant to the claims being made. These criteria for content validity leave much to the individual researcher to interpret. In addition to requiring a high level of interpretation, Leech et al. do not provide a scoring procedure for assigning definitive values to each of the five criteria presented.

The fourth source of guidelines for evaluating mixed methods research is found in *The Quality of Mixed Methods Studies in Health Services Research* (2008), by O'Cathain, Murphy, and Nicholl. O'Cathain et al. present a set of guidelines they refer to as the Good Reporting of a Mixed Methods Study (GRAMMS). GRAMMS is comprised of six guidelines. These guidelines include:

> (1) the justification for using mixed methods approach to the research question, (2) the design in terms of the purpose, priority, and sequence of the method, (3) the description of each method in terms of sampling, data

> collection, and analysis, (4) the description where integration has occurred, how it occurred, and who participated in it, (5) the description of any limitation of one method associated with the presence of the other method, and (6) any insights gained from mixing or integrating methods" (O'Cathain et al., 2008, p.97).

Like Creswell and Plano Clark and Leech et al., O'Cathain et al. set forth a set of guidelines for the evaluation of mixed methods, but fail to present a way to score the criteria presented in their literature. Besides not presenting a scoring guide, O'Cathain et al. also differs from the other sources by creating a comprehensive list of guidelines that point to a need for a deeper methodological discussion in each part of the study. O'Cathain et al. present guidelines for a more holistic description of a mixed methods research study.

Schifferdecker and Reed are the fifth source to present guidelines for evaluating mixed methods research (2009). Schifferdecker and Reed propose guidelines that focus on the creation of a mixed methods research study, but they do discuss how those guidelines are scored. Schifferdecker and Reed present a list of guidelines based on the goals of mixed methods research to establish necessary components for what should be included in a mixed methods research study. These guidelines do not meet the definition of evaluation, but do fall in a similar position as the previous sources with the exception of those presented by the NIH. Schifferdecker and Reed propose the exploration of tools for integrating the data and the method for reporting the findings in accordance to current mixed methods research literature. They also state that the study design should be identified as mixed methods and that the prominence of each data type should be

addressed. In addition to identification and prominence of data, Schifferdecker and Reed place an emphasis on adequate data collection procedures for each data type and how the data are analyzed in accordance with mixed methods research standards. Schifferdecker and Reed's guidelines can be followed to ensure some degree of quality to a mixed methods research study, but cannot be used as a definitive evaluation criterion. Schifferdecker and Reed's proposal differs from the others by focusing heavily on the data types and collection process. Much is left out in their schema for evaluating mixed methods research. Schifferdecker and Reed require the researcher to support their research by using additional sources for evaluating mixed methods research.

The final source of evaluation guidelines for mixed methods research is presented by Pluye, Gagnon, Griffiths, and Johnson-Lafleur (2009). Pluye et al. are the only ones to propose a scoring system, entitled the Mixed Methods Appraisal Tool (MMAT), for mixed methods research based on the idea that quality appraisal is important to a methodology. In addition to a scoring system that applies to qualitative, quantitative, and mixed methods research, which is done because of the belief that all three research methodologies are inherently integrated, Pluye et al. also provides guidelines for conducting a mixed methods research study. Pluye et al. reviewed 59 qualitative, quantitative, and mixed methods articles and coded them to create generic themes about what design components the articles contained. These themes led to the creation of the scoring system. Although the scoring system was created to measure the quality of the three research methods, only 3 of the 14 criteria are used to measure mixed methods research studies specifically. The three criteria used for mixed methods studies include (1) rationale of the mixed methods design, (2) use of both qualitative and quantitative

data collection and analysis procedures, and (3) integration of qualitative and quantitative data and/or results (Pluye et al., 2009, p. 540). Since the MMAT's creation in 2009, the MMAT was tested for reliability, revealing that the scoring system provided a high level of inter-rater reliability for 17 of the 19 criteria (Pace et al., 2012). Seventeen of the criteria received a Kappa score of .94, with the remaining two receiving a score between .83 and .89 (Pace et al., 2012). The MMAT provided a tool for appraising quality in mixed methods research and provides a strong foundation for future scoring rubrics (Pace et al., 2012).

Although the MMAT provided good inter-rater reliability, there are critical faults with the instrument and its reliability testing. The items of the MMAT are scored dichotomously by noting the presence or absence of each, allowing for a lack of variable quality rating. Another limitation is that only one mixed methods article was used in the reliability testing for the mixed methods criteria (Pace et al., 2012). Pluye et al. addressed several criteria that are previously discussed, but left out several others addressed by Creswell and Plano Clark (2011), Tashakkori and Teddlie (2010), and O'Cathain et al. (2008).

The articles presented in this section contain several similarities and differences for the criteria of what constitutes a quality mixed methods research study. Only the NIH standards (2012) and Pace et al. (2012) address the evaluation of quality in mixed methods research. Similarities between articles include looking for the presence of mixing, qualitative and quantitative designs, and the use of pertinent literature on mixed methods research to support the study (Creswell & Plano Clark, 2011; National Institutes of Health, 2012; O'Cathain et al., 2008; Pluye et al., 2012; & Schifferdecker & Reed,

2009). Differences between these sources range from practice (National Institutes of Health, 2012) to theoretical concepts (Leech et al., 2009). Some of these differences include the use of content validity to evaluate quality of a study and the in-depth examination of data. The similarities and differences found in these articles reveal the complex nature of creating a rubric that is capable of evaluating such a variety of studies with multiple components. Several of these components have been presented in these guidelines, such as research designs and mixing and paradigms, while other pieces described as being necessary to a mixed methods research study, such paradigms, have not been included.

Designs Found Within Mixed Methods Research

When looking towards an evaluation of mixed methods research studies, perhaps the most important piece of the evaluation puzzle is the research design typology (Creswell & Plano Clark, 2011; National Institutes of Health, 2012; O'Cathain et al., 2008; Pluye et al., 2012; & Schifferdecker & Reed, 2009). Research designs "are procedures for collecting, analyzing, interpreting, and reporting data in research studies" (Creswell & Plano Clark, 2011, p. 53). Design typologies include, but are not limited to, "number of methodological approaches used, number or strands or phrases, type of implementation process, stage of integration approaches, priority of methodological approach, function or the research study, and theoretical perspective" (Teddlie & Tashakkori, 2006, p. 13).According to Teddlie and Tashakkori (2006, p. 12), there are five reasons why research design typologies are important in mixed methods research. Those are:

1. Typologies guide the researcher during the design process of the mixed methods study. The design provides various paths that may be chosen in order to accomplish the goals of the study.
2. Typologies allow for a common language. An example of this common language is Morse's (1991, 2003) system of notations and abbreviations, which are still in use today.
3. Typologies assist in organizing structure within the field of mixed methods research. Various typologies give the researcher a choice of alternative organizational designs.
4. Typologies legitimize mixed methods research by providing distinct examples that are different from those found within quantitative or qualitative research designs.
5. Typologies are a pedagogical tool that helps in the teaching of alternative designs found within mixed methods research.

Research design typologies are useful tools when creating mixed methods research, but like the field of mixed methods research itself, they are in a state of change. Research designs made one of their first appearances in Greene et al.'s (1989) "Toward a Conceptual Framework for Mixed Method Evaluation Designs." Further steps in the creation of research designs were taken by Patton (1990), Sandelowski (2001), Teddlie and Tashakkori (2006), and Creswell and Plano Clark (2011) to name a few of the most popular typologies found in the current literature. The foundation for these designs ranged from the theory and practice found within mixed methods studies to the different strands used in studies (Greene et al., 1989; Teddlie & Tashakkori, 2006). In addition to

the creation of different designs, designs have also been removed or replaced by more current ones, which will be addressed in more detail later in this section (Creswell & Plano Clark, 2011). This constant change of design must be taken into account when evaluating the quality of mixed methods research. Since all design typologies serve several important purposes, they all must be considered equal in the overall process of conducting a quality mixed methods research study (Teddlie & Tashakkori, 2006). Each component of the research design typology adds to one another to create the final product.

Since design typologies play a significant role in the design of a mixed methods research study, special care must be taken when evaluating them. Designs may have to be weighted more heavily than other components due to their prominence in current guiding practices and evaluation criteria (Creswell & Plano Clark, 2011; National Institutes of Health, 2012; O'Cathain et al., 2008; Pluye et al., 2012; Schifferdecker & Reed, 2009). The quality of a mixed methods research study must include a rational for the design used, even if it is an unspecified design. Research designs are an important piece to a study because they provide the blueprint for how it is conducted. Different designs will be taken into account developing this rubric by creating a system of rating designs based on different theoretical perspectives (Creswell & Plano Clark, 2011; Greene et al., 1989; Teddlie & Tashakkori, 2006). A key evaluation point may come in the form of how well the researcher uses supporting literature to demonstrate the need and theory of the selected design.

Design typologies; derive from different theoretical backgrounds and purposes, which led the creation of the unique designs (Bryman, 2006). Three different design

viewpoints by Greene et al. (1989), Teddlie and Tashakkori (2006), and Creswell and Plano Clark (2011) are presented here. Greene et al. (1989) focus on the purpose of the study to define the design. The five designs proposed by Greene et al. (1989) each serve a different purpose including triangulation of data, enhancement of results, the informing of other methods, discovery of new paradoxes, and extending the range of inquiry (Greene et al., 1989). Teddlie and Tashakkori (2006) differ from Greene et al. (1989) by basing their design strategy on of the criteria which define mixed methods research, including the number of methods used and how the study is implemented (Teddlie & Tashakkori, 2006). Creswell and Plano Clark (2011) depart from Greene et al. (1989) and Teddlie and Tashakkori (2006) by taking the position that the researcher should use a design that supports the research problem and the reason for mixing (Creswell & Plano Clark, 2011). Each of these reasons for using different designs serves the purpose by assisting the researcher in preparing a quality research study. Research designs fall under the proverbial "there is more than one way to skin a cat". This is why the design typology must be acknowledged in the evaluation process of a quality mixed methods study.

There are similarities in the different major design typologies proposed in the literature. The three design systems addressed here focus on guiding the purpose of the researcher (Creswell & Plano Clark, 2011; Greene et al., 1989; Teddlie & Tashakkori, 2006). Based on a combination of mixed methods research and the designs proposed by these individuals, there is also an agreement on the acknowledgment of methodological approaches used in the study, number of phases, and a level of mixing or integration (Caracelli & Greene, 1993; Creswell & Plano Clark, 2011; Teddlie & Tashakkori, 2006).

This agreement points to a beginning point of how to evaluate research designs in mixed methods research.

Unfortunately, all are not in agreement in mixed methods designs. Disagreement is found not only *between* researchers, but also *within* researchers. Teddlie and Tashakkori (2006) differ from Creswell and Plano Clark (2011) on the value of priority. Priority refers to the importance of qualitative and quantitative methods used to answer the research questions (Creswell & Plano Clark, 2011). Teddlie and Tashakkori (2006) do not include priority in their design typology because they argue that the importance of qualitative or quantitative methods cannot be determined at the out start of the research project (Teddlie & Tashakkori, 2006). Teddlie and Tashakkori (2010) also differ from Greene (1989) by arguing against the use of theoretical perspectives because they only serve to shape the purpose, which supersedes the research design (Teddlie & Tashakkori, 2006). Creswell and Plano Clark (2011) go further by changing their typology of research designs. In the second edition of *Designing and Conducting Mixed Methods Research*, Creswell and Plano Clark remove the triangulation design and introduce the multi-phase and transformative designs (Creswell & Plano Clark, 2011). These changes were done "to bring multiple designs together" (Creswell & Plano Clark, 2011, p. 69). These few differences shed light on the fact that differences are present in viewpoints about design typologies and must be taken into account when evaluating and assortment of designs.

Designs are deeply embedded in the literature as being important to mixed methods research. Within the literature, there are both similarities and differences among researchers. As with many facets of mixed methods research, designs will go through

many changes as well as fuel future debates. Of the mixed methods debates, designs are not the only topic being discussed in great depth.

Mixing

Along with designs, mixing is another piece of the mixed methods research process that is often debated in the literature as a critical component to the mixed methods process (Bryman, 2007; Caracelli & Greene, 1993; Creswell & Plano Clark, 2011). Woolley defines mixing as "the extent that (qualitative and quantitative) components are explicitly related to each other within a single study and in such a way as to be mutually illuminating, thereby producing findings that are greater than the sum of the parts" (Woolley as cited in Bazeley, 2009, p. 204). Although there is no clear consensus on the level or amount of mixing required in a study, the literature does provide evidence that it is an important enough piece to be included in the evaluation of quality of mixed methods research (Bryman, 2007; Caracelli & Greene, 1993; Creswell & Plano Clark, 2011; Tashakkori & Teddlie, 2010). Mixing can also be referred to as integration or triangulation. It is mixing that allows the researcher to create a study that connects the qualitative and quantitative data. Without mixing, the study becomes a multi-method study with qualitative and quantitative data analyzed to tell two distinct stories. By not including mixing, the primary concepts of a combined methodology and one that builds upon its parts to enhance its strength, is lost (Bryman, 2007; Caracelli & Greene, 1993; Creswell & Plano Clark, 2011).

The mixing of quantitative and qualitative data can occur in several distinct ways within a mixed methods study. According to Creswell and Plano Clark (2011), mixing may appear in five forms: including triangulation, merging, linking, embedding, and

transformation (Creswell & Plano Clark, 2011). Mixing is not limited to just the analysis phase and may occur in multiple phases, including the research design, data collection, analysis, and findings. The type of mixing used and where the data is mixed depends on the study and the data being used. The first type of mixing, triangulation, is the use of multiple types of data to create a level of support for the research studies findings. Next, merging is the cross tabulation of quantitative and qualitative data. The third and most basic type of mixing, linking, occurs when one type of data is directly related to another type of data within one study. The fourth category of mixing, the embedding of data occurs in design phase by placing one type of data within a larger data set of a different type. Data transformation, the fifth type of mixing, is when quantitative data are transformed into qualitative data or vice versa to allow for quantitative or qualitative analysis (Caracelli & Greene, 1993). "The primary purposes of [data] transformation, including consolidation, are to enable further higher-order analyses" (Greene, 2007, p.145).

Mixing is what makes mixed methods distinctive from other research methodologies, thus making it an important part of the criteria for evaluating mixed methods research. Mixing serves as the bridge between qualitative and quantitative methodologies. It is for these reasons that I consider it likely that it will be a major feature of a rubric. Currently there are no rubrics for evaluating the quality of mixing, so this rubric will focus on the presence of mixing and to what it adds to the overall study based on expert recommendations through interviews during the rubric construction process. Further research on the quality of mixing may be undertaken once the initial framework for quality is constructed for mixed methods research.

Paradigms

Paradigms refer to a set of epistemological and ontological assumptions, or beliefs, about the knowability and beliefs of the social world as we understand it (Harrits, 2011). Paradigms are the assumptions made about our social practices and the ways we can construct an understanding of this reality (Blakie, 2010). Paradigms can also "refer to a common research practice, existing within a research community, and carrying with it a shared identity as a set of problems" that are related to the advancement of knowledge (Harrits, 2011, p. 152). The use of paradigms directs research efforts in both the practical and methodological level (Harrits, 2011).

This research study does not make a case for which paradigm is most effective or most important in mixed methods research, but instead points out the purpose for paradigms in mixed methods studies and why they are important as a whole. Paradigms are linked to quality by providing assumptions and beliefs in which the research is grounded, thus making the study more transparent. Paradigms create a theoretical foundation, providing support for the purpose and design of the study. The use of paradigms is demonstrated by the researcher's reflexivity on his or her worldview and how the study fits that worldview. The more evidence provided by the researcher supporting his or her research framework, the higher the level of overall quality. In addition to supporting the research, the disclosure of paradigms also strengthens the trustworthiness criteria of the qualitative and quantitative component of the mixed methods research study. By showing that paradigms serve an important role in the research process, I argue that on this premise, it is likely that interview participants will include paradigms in the evaluation process of mixed methods research.

Several paradigms are found in current mixed methods literature. This section reviews the most commonly found paradigms in mixed methods research. These paradigms include pragmatism, realism, transformative, and the dialectic stance (Maxwell & Mittapalli, 2010; Mertens, 2007).

Pragmatism is most often the framework as the paradigm underlying mixed methods research. Pragmatism is a way to free the researcher of "mental and practical constraints" created by choosing between positivism and constructivism (Feilzer, 2009, p. 8). Pragmatism frees the researcher by "focus[ing] on the consequences of research, on the primary importance of the question asked rather than the methods, and the use of multiple methods of data collection to inform the problems under study" (Creswell & Plano Clark, 2011, p. 415). Pragmatism is often thought to be the philosophical partner to mixed methods research by presenting a set of assumptions about the methods of inquiry (Denscombe, 2008). Pragmatists question the duality of positivism and constructivism by converging methods to create knowledge (Feilzer, 2009; Rorty, 1999). The ideas behind pragmatism date back to 1925, when Dewey remarked that the truth must be found, regardless if it is an objective truth or a subjective inquiry (Dewey, 1925). The idea of seeking "the truth" lends itself well to the field of evaluation and research, making it a popular choice for many mixed methodologists (Denscombe, 2008; Fitzpatrick et al., 2011). Admittedly, pragmatism is not the only choice of paradigm for mixed methods research, nor am I advocating that it is the best choice, but pragmatism does present a worldview that has a direct impact on a research design.

In addition to pragmatism, realism is another paradigm that has made its way into the mixed methods research literature. Realism is based on the belief that "there is a real

world with which we interact, and to which our concepts and theories refer" (Maxwell & Mittapalli, 2010). Realism provides a philosophical stance that is well-suited with both quantitative and qualitative research (Maxwell, 2004). Realism includes several subtypes, including critical realism, experimental realism, and scientific realism (Maxwell & Mittapalli, 2010). The distinguishing factor of these three forms of realism "is that they deny that we have any objective or certain knowledge of the world, and accept the possibility of alternative valid accounts of any phenomenon" (Maxwell & Mittapalli, 2010). Like pragmatism, realism allows the researcher to maintain a philosophical stance that is compatible with both quantitative and qualitative research methods, thus allowing for dialogue between the two fields (Greene, 2000). Realism is useful in mixed methods research by viewing a research design as a real entity, not just a model for research (Maxwell & Mittapalli, 2010). Realism is also useful by providing the reader of a mixed methods research article the understanding of the *real* design as "logic-in-use" (Maxwell & Mittapalli, 2010, p. 153). Realism appeals to many researchers by embracing the knowledge of the real world, its construction, and its perspective (Greene, 2007). Since there are multiple perspectives of the world, multiple scientific observations can be used in the creation of conceptual schemes and categories in a study (Lakoff, 1987). Realism provides both quantitative and qualitative ways of understanding the studied phenomenon (Maxwell, 2004).

In addition to pragmatism and realism, the transformative paradigm "provides an overarching framework for addressing issues of social justice and consequent methodological decisions" in mixed methods research (Mertens, 2007, p. 212). The transformative paradigm focuses on the researcher instead of a total worldview. The

transformative paradigm assumes that knowledge is influenced by the researcher, reflecting their power and relationship with society (Mertens, 1999). A researcher following the transformative paradigm recognizes inequalities and injustices found within society and desires to challenge these norms or the status quo (Mertens, 2007). The transformative paradigm's primary premise is that power differentials must be addressed at every stage of the research process (Mertens, 2007). Studies that follow the transformative paradigm must divulge their philosophical underpinnings in publications so that the reader clearly comprehends what is transpiring and why it is conducted in such a manner to be transparent to the audience. The transformative paradigm promotes and stresses a consciousness in inquiry (Greene, 2007). An example of the transformative paradigm used for advocacy can be read in Merten's (1997) work promoting the interests of the deaf and hearing impaired communities.

The final paradigm that I will discuss is the dialectic stance. The dialectic stance maintains the goal of creating a dialogue between different perspectives being studied in order to deepen the understandings gained (Maxwell & Mittapalli, 2010). The intertwining of different perspectives is motivated by good of the public (Greene & Henry, 2005). "Greene considers this to be the most valuable stance for mixed methods research because the juxtaposition of different lenses or 'mental models' that it requires is most likely to produce generative insights" while creating meaningful engagement between different groups (Maxwell & Mittapalli, 2010, p. 147). Questions pertaining to social inquiry begin with the purpose, which often derives from the paradigm (Greene, 2007). The dialectic stance allows for researchers using different paradigms to discuss matters, thus creating more than just a triangulation of data (Maxwell & Mittapalli,

2010). These matters may include purposes such as the evaluation of programs (Greene, 2000). The dialectic stance can improve research and evaluations by combining paradigms by juxtaposing differences to produce new understandings (Greene, 2007). The ideas surrounding the dialectic stance provide the researcher the opportunity to design studies that foster further discussion and to have a philosophical framework in which to work.

Since research paradigms are included in the way mixed methods research is justified and conducted, the concept of paradigms can be a useful tool in the comparison of different mixed methods research strategies (Harrits, 2011). Ontological, epistemological, and axiological assumptions are part of the values of researchers and evaluators and the research they conduct (Mark, Henry, & Julnes, 2000). Although researchers include these assumptions in their work, they often do not make their paradigm preferences explicit to their audience (Alise & Teddlie, 2010). Of 600 articles reviewed, Alise and Teddlie (2010) found that researchers reporting empirical studies rarely identified paradigmatic orientations and that post positivists and constructivists were less likely to describe their philosophical support that others due to a perceived understanding of field specific paradigm use. Alise and Teddlie recommend that authors should describe their paradigm preference and how it affects the study (Alise & Teddlie, 2010). Paradigms fill an important need within research, for both the researcher and the audience. It is for these reasons that paradigms should be included in the evaluation process for mixed methods research.

Summary

Several key components that detail the evaluation of mixed methods are acknowledged in the mixed methods research literature. These components include the current ways mixed methods studies are evaluated, research designs found in mixed methods research, the role of mixing data, and the use of paradigms. Discussion about tools to evaluate mixed methods research has only recently begun to emerge in the literature. The evaluation tools discussed here utilize guidelines found in the mixed methods literature, but do not offer an in-depth, comprehensive measure of the multiple components found in mixed methods research designs, including mixing and reflexivity about paradigm use. Mixing is regarded as a critical piece in the definition of a mixed methods study, but the amount of mixing and type of mixing is not taken into account in the evaluation tools reviewed in this section. Paradigms support the quality of a mixed methods study by providing assumptions and a framework for the study, but are also left out of the evaluation tools available in the literature. The addition of research designs, mixing of data, and paradigms in the rubric proposed in this study will create a more comprehensive tool for evaluating the quality of mixed methods research studies.

Possible Findings of This Research

Results of this study may include both an agreement and disagreement among experts in the field of mixed methods research. In addition to these similar and different stances on several areas and criteria within the mixed methods methodology, new criteria are expected to emerge from the interviews conducted with the experts chosen for this study. An updated look into the current thoughts on what makes a mixed methods study

as well as new found criteria for describing a mixed methods study will not only add to the current literature, but also update positions found in past publications.

My analysis of current mixed methods research literature leads me to believe that the experts will agree on the necessity of several critical components that are said to comprise a mixed methods study. Greene (1989), Creswell and Plano Clark (2011), and Tashakkori and Teddlie (2010) all agree on the use of research designs, paradigms, and the mixing of data when conducted mixed methods research. Since the three primary developers of mixed methods research designs all agree on these three components, I believe that it is likely that most of the experts interviewed will also agree. I also believe that most of the interviewed participants will agree on these three points because they are inherent to the fundamental practices of mixed methods research. A research design is necessary to create a blueprint of how the study will be carried out. In addition to a design, paradigms are inherently important because they provide a degree of transparency by providing the context in which the research is framed. Mixing will also probably be addressed because without it, by current definitions a study is simply multi-method. Although I believe there will be some agreement among the experts about different components of mixed methods research, I believe that there will be a greater level of disagreement found when analyzing the qualitative data. The individuality of the researcher will likely provide many opinions that are embedded in both experience and personal beliefs.

Although I believe that the experts will agree that mixing is critical to mixed methods research, I do not believe that they will agree on how much mixing will be required or in what phase(s) the mixing should occur in the study. Another source of

disagreement may be in the use of vocabulary. Since several researchers currently use different terms, it may be difficult to come to a consensus on which terms should be used and why. The level of disagreement may be much greater than these listed here, but I believe that these two components will have the greatest level of disagreement. Much of the current disagreement found in the mixed methods research literature may not be active disagreement, but rather a product of what is a more pressing topic at the time of the publication.

I expect to find several points of both agreement and disagreement by interviewing experts on mixed methods research. I addition to these, I also expect that several new criteria will emerge from the qualitative data. By asking the experts the same five questions and providing a time to speak candidly on the subject, it is likely that I will find new themes that may be based on current research that has yet to become popular in the literature. The mixed methods methodology has proven to be an evolving field and it is for this reason that I highly suspect new criteria to become evident. I cannot predict what the new criteria will be due to the multitude of current criteria, but for this very reason I expect that new criteria will arise. I do not believe that the current criteria found in the mixed methods research literature have reached a point of saturation because new findings are still being revealed in the literature.

Chapter 3: Methods

Research Questions

The purpose of this mixed methods study is to conduct a cross-disciplinary, multi-phase content analysis to create a preliminary rubric to evaluate mixed methods research articles in multiple disciplinary fields. The four research questions that guide this study are:

1. What are the common evaluation criteria found in the contemporary methodological literature pertaining to the design of mixed methods research? (Qualitative & Quantitative)
2. What evaluation criteria do experts in the field of mixed methods research perceive as the most important when distinguishing top-quality research in mixed methods? (Qualitative)
3. What differences are there in the outcome of the rubric for evaluating mixed methods research indentified from the literature compared to those advocated most uniformly by a panel of mixed methods research experts? (Qualitative & Quantitative)
4. What are disciplinary differences between the use of mixed methods and views about evaluating it, including the role of paradigms in mixed methods research? (Qualitative)

Introduction

This multi-phase mixed methods exploratory research design was comprised of two distinct phases to answer the four research questions proposed in this study. The

process used for this multi-phase mixed methods study is illustrated below. The diagram presented in Figure 3.1 depicts the sequential timing and quantitative priority.

Figure 3.1. The Design for this Study's Multi-Phase MMR Design

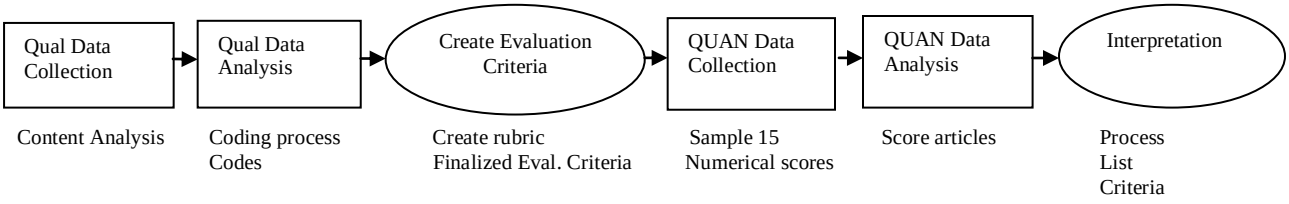

Figure 3.1. Multi-phase mixed methods research design.

The first phase was a qualitative piece where interviews were used to create a rubric for evaluating mixed methods research articles. The second phase was a quantitative piece to conduct a pilot test of the rubric. A rubric is a scoring guide for assessing the performance of a product (Wolf & Stevens, 2007). In this study, the product is a guide to assess the quality of mixed methods research.

This methods section is organized by phase of the study. This information includes a sampling plan, procedures for data collection, and a description of how the data were analyzed. Following a description of the unique properties of each phase, information is provided about how and where mixing occurred throughout this study as well as a description of the limitations of this study.

The methods used in this study adhered to strict ethical and professional guidelines and were approved by the Virginia Tech Institutional Review Board (IRB) to ensure the protection of human subjects. An initial IRB protocol was submitted that included the initial e-mail used to contact the participants, the participant consent form, and the interview protocol. A revision to the IRB protocol was made to include the second e-mail and website used for collecting the data used to rank the importance of the

themes generated from the interviews. A revision was used since the initial interview data was not collected at the time of the initial IRB protocol.

Phase One

Sampling

The sampling plan in phase one was a purposeful sample of mixed methods research experts. The sample was created by contacting a group of 52 experts from the mixed methods research field. "Experts" in this study is defined as any author who wrote or co-wrote a chapter in the *Sage Handbook of Mixed Methods in Social & Behavioral Research* (2010). The authors of these chapters are considered experts because of this book's recognition as a leading textbook in the field. The authors within this text have displayed a significant contribution to the development and progress of mixed methods research as a methodology.

An email (see Appendix C) was sent to each of the 52 authors and co-authors requesting their participation in this study. Two electronic reminders (resending of Appendix C) were sent out to non-respondents in seven day intervals following the initial mailing. The desired response rate for this phase was 38%, or 20 participants in order to ensure a diverse pool of researchers. If less than 12 individuals agreed to participate in this study, a snowball sampling method was used to generate at least 12 participants. Each participant was asked to provide the names of researchers they perceived to be experts in the field of mixed methods research. These individuals were then contacted using the same method as the original 52 authors until 12 participants agreed to participate.

Procedures

Qualitative interview data collected in the first phase was used in the second part of the initial phase to develop the rubric. The interviews were conducted by utilizing an interview protocol (see Appendix B) primarily containing open-ended questions.

The interview questions reflected the components of mixed methods research as described in the literature. The protocol questions focused on why quality is important and what components of mixed methods research are the most important and why. Questions were specifically asked about the component of mixing data because of its importance and often misuse or mislabeling. Specific questions also addressed the role of paradigms and their importance within mixed methods research. In addition to quality, important components of mixed methods research, mixing, and paradigms, a question was also included about the viewpoint of mixed methods research based on the participant's discipline. The six questions of the interview protocol provided information that was used to create the scoring rubric. In addition to providing information to create the rubric, the interview data were also used to answer research questions two and four.

The interviews were conducted via telephone or computer based communication software, such as Skype. The conversations were audio recorded using a digital recorder. Prior to conducting the interview, participants electronically signed a digital copy of the consent form, which included whether their names could or could not be used in the reporting of the data. Before the interviews took place, I practiced the protocol on a faculty member who is experienced in mixed methods research to ensure an accurate and professional interview process.

Once the interviews were complete, the digital recordings were outsourced to a local transcription service. The transcriber was directed to transcribe the interview word for word with each line of text being numbered. This was done to provide a precise and accurate account of the interview. Pseudonyms were used to maintain the confidentiality of the participants. The pseudonyms were non-gender or racial specific and were labeled as Participant 1, Participant 2, etc. The names and locations that were given as part of examples were changed to protect other individuals and institutions.

As the transcripts were received, an inductive coding process occurred for each interview. This inductive coding process included line-by-line coding, which led to the creation of several themes. In this study, a theme refers to a recurring idea that is created from an inductive analysis of qualitative data. The themes which were developed from the inductive coding were then transformed into variables and used to create items for the rubric. A model item is presented in Figure 3.2.

Figure 3.2. Example Item Found in the Quality Rubric

Standard	Strongly Disagree (0)	Disagree (1)	Somewhat Disagree (2)	Somewhat Agree (3)	Agree (4)	Strongly Agree (5)
The results of the study are clearly stated.						

Figure 3.2. An example of a rubric item.

Once the coding was complete and a list of themes was generated, an intensity matrix was developed to document the occurrence of themes. The matrix was created by using x and y axis chart giving the name of the theme and how many times it occurred throughout the interviews. In addition to the intensity matrix, correlations between the themes were analyzed to highlight the relationship between themes. The themes that occurred most often with the highest level of correlation to the amount of times addressed

by the participants were chosen as the most important to transform into the initial rubric items. The final rubric will contain eight items; so therefore, the number of themes was limited to 16 in order to provide twice as many items for the purpose of refining and replacing poorly performing items on the rubric.

From a selection of primary themes, a list of items was created. This list acted as a rough draft for the rubric (see Appendix E). The themes were transformed into items that could be polytomously scored on a six-point Likert scale. The six-point scale specifically represented each item, eliminating a generalized scale. A six-point Likert scale was adopted for its ability to be used as a continuous variable. The six-point scale also allowed for tests of unidimensionality, ensuring that one construct was being measured. For this study, the desired construct, or the theoretical conception of a process, was quality of mixed methods research articles.

Once a rough draft of this rubric was created of the items designed from the interview themes, a copy was be sent back to each participant. The participants were then asked to rank the items by importance. Each item of the rough draft rubric was then scored. Each item was scored using mean score for each item. Each item with a mean score of 4.0 or higher was included in the rubric. This score was chosen because it represents the minimum score for an item to be rated important. The items with a score of four or greater were selected to be on the final rubric (see Appendix E) that was pilot tested in phase two.

In cases where the participant agreed to provide an interview, but declined to rate the items, the participant was not contacted a second time. The frequency of themes from their interview was used to score item importance by assigning the most points to the

items that were most frequently reflected by their themes. The frequency of themes was determined by adding the number of times a theme was documented in the individual interview data.

Analysis

The qualitative analyses of phase one was reported in research article one, found in chapter four of this manuscript. Research article one includes thick and rich descriptions of the themes produced from the inductive coding of the interviews. In addition to the qualitative analysis, the intensity matrix provided evidence of how the primary themes were selected for use in the rough draft of the rubric.

Phase One Trustworthiness and Validity

Member checking, an audit trail, peer review, and reflexivity were used to ensure a high level of trustworthiness in the first phase of this study (Patton, 2002). The transcripts were sent to the participants to make certain the accuracy of the interview and to allow the participants the chance to add to any previous statements. In addition to the member checking, an audit trail, made up of procedural memos, was created to trace and outline the decision making processes of this study. A peer review of the coding schemes supported the member checking and audit trail by creating a high level of acceptance for the selected codes. Finally, reflexivity, or the process of being self-analytical, defined my own assumptions as a mixed methods researcher (Patton, 2002). These four steps developed the reliability of the study by adhering to the guidelines set forth for qualitative research by Patton (2002).

The peer review of the coding scheme was conducted by faculty knowledgeable in qualitative research. The faculty member chosen instructs courses in qualitative

research. Four interviews were coded by the researcher and a mixed methods faculty member to ensure inter-rater reliability of at least 80%. An inter-rater reliability of 80% is acceptable according to Patton (2002).

In addition to the trustworthiness of the qualitative component of phase one, Kane's (2011) criteria for validity was applied to the instrument development piece of this phase. Kane stated that validity is created through the interpretation of the score and the plausibility of the instrument's proposed uses (Kane, 2011). The claim of this instrument, the rubric, is that it can be used to score the quality of mixed methods research articles. The evidence that supports this claim comes from the data (participants being interviewed and their responses to the protocol). The claim is warranted through a series of assumptions which back the claim. As experts, it is assumed that the themes that arose during the interview process represented the critical parts of a mixed methods study. By identifying the themes that appeared with the greatest frequency, it was assumed that these were the most important themes amongst the participants in regards of quality. By turning these themes into items, it was then assumed that the items reflect the most important issues in judging the quality of mixed methods research. Finally, by scoring the items, the rubric was then made more concise by selecting those items which were scored the most vital by the experts. By following this logical sequence, the rubric's score of the quality of mixed methods research articles was determined to be valid and also potentially generalizable to a larger population, which was attempted in the pilot study in phase two.

Phase Two

Sampling

The sampling of phase two occurred in two ways. The first sampling plan outlined the procedures used to gather a collection of mixed methods articles, which were used to test the generalizability and validity of the rubric. The second sampling plan detailed how the reviewers were chosen to provide inter-rater reliability of the rubric.

A stratified sampling plan was created to guide the selection of articles suitable for validating the rubric. Ninety-three mixed methods articles were selected from the Ebscohost general database. The general database was narrowed down to only include articles found within the education database within Ebscohost. To further the selection process, keywords were used to identify the articles as mixed methods research. The mixed methods research articles used in this study were published between 2008 and 2012. A refined search of article abstracts that included the keywords "mixed methods", "multi-method", and "qualitative and quantitative" revealed a total of 153 articles. Each of the 153 articles was reviewed in detail to identify appropriate mixed methods research articles. The guidelines for this selection were based on the mixed methods quality guidelines created by Creswell and Plano Clark (2011). After applying Creswell and Plano Clark's guidelines, the selection was narrowed down to 93 articles. From these 93 articles, a random selection of 15 articles was chosen for use in this pilot study by employing a random number generator to choose 15 of the 93 articles. The 15 selected articles were then used by the selected reviewers to establish a level of instrument validity. Fifteen articles were selected due to the available resources, including time and

reviewers. The articles chosen for this study were limited to their availability in the Ebscohost general database.

A purposeful sampling plan was also used to collect five mixed methodologists to act as article raters. The five researchers were selected based on location to limit costs and time delays. In addition to location, the researchers were also chosen based on experience, meaning they have published at least one article using mixed methods research or have taken a mixed methods research course. The researchers were doctoral students from a large (more than 40,000 students) southeastern, research one institution. Once the sampling procedures were in place, the procedures for phase two were developed.

Procedures

The phase two procedures were conducted in several steps. The first step was to select a list of 15 articles to be reviewed. The second step was to assign the articles to the reviewers for evaluation. The third step was to compare the ratings of the reviewers to create a level of score validity and also produce a degree of unidimensionality for the construct being scored. An initial training session was given to the reviewers in order to instruct them on the proper use of the rubric and to give them the opportunity to ask any questions about the items. The training session was conducted by giving each rater the same article and reviewing the article as a group. The raters first used the rubric to rate the article and then compared their results. The raters then discussed how they arrived at their results to ensure they were interpreting the rubric uniformly. Once phase one was complete, the articles were distributed to the selected reviewers to begin the second step. Each reviewer was be assigned a number that he or she placed on the review form.

Once the reviewers received the articles, the second step of procedures began. The reviewers received the articles numerically labeled from 1 to 15. The reviewers began with number one and proceeded to score as many articles as they could over a one week period using the final rubric. This allowed for the reviewers to overlap the articles they were reviewing as well as create a benchmark for timeliness for how long it took them to rate a number of articles. The reviewers did not assign their names to the scoring rubric, so that the review could be done in a blind review approach. (A secondary goal of this rubric was to create a scoring system that could be completed in a short period of time by keying in on specific elements of an article.) I scored each article to create a baseline score for comparison to the other reviewers. In addition to the scoring, the reviewers also documented the time spent on reviewing each article on the rubric. The scoring of the articles concluded the second step of the phase two procedures.

The final step of the phase two procedures supported the validity of the rubric's score and tested the unidimensionality of the rubric. The data of the scoring rubrics was uploaded into jMetrik v.2.10 Rasch modeling software to determine the levels of discrepancy and fit for each item. The data was also uploaded into SPSS v.20 statistical software in order to run an exploratory factor analysis (EFA).

Analysis

The analyses of phase two is not presented in the conclusion of this manuscript due to the lack of statistical significance. The analysis of phase two included the results of the scoring rubric pilot study and recommendations for future revisions. The results of the individual reviewers were given as well as the output from jMetrik, revealing the individual item statistics.. In addition to the jMetrik output, the EFA provided by SPSS

demonstrates the level of unidimensionality of the "quality" construct. The EFA produced Eigen values for the data, showing the possible number of factors as well as a diagram of how the items loaded on the factor(s). Overall, the analysis showed how the rubric will be validated in the future with a larger sample to create statistically significant data for determining the overall quality of the rubric.

Mixing of Phase One and Two

As discussed previously in the literature review, mixing is a critical component to any mixed methods research study. The mixing of data in this study occurred in four distinct ways. The first type of mixing found was the intertwining of inductive and deductive research methodologies. The second type of mixing found was quantitizing. The third type of mixing found was the creation of polytomous variables. The fourth and final type of mixing found within this study was linking.

The first type of mixing occurred at the introduction of this research study in the research design phase. During the research design phase, research questions combined the use of qualitative and quantitative methods. The combination of inductive and deductive research methods allowed for the study to provide a holistic view of the judgment of quality found in mixed methods research articles (Bryman, 2006). Although this type of mixing is straightforward, it provided a mixed method's framework for the entire argument of this research.

The second type of mixing found in this study was the use of quantitization to create a ranking system for the themes created from the qualitative phase. By using the transcripts of the participant interviews, numerical data were created through the use of frequency tables and correlations of those frequencies. The qualitative data from the

interviews was transformed into quantitative data by tabulating the frequencies of the codes and themes found within the interviews. This process of transforming qualitative data into quantitative data provided a way to numerically represent the number of qualitative data, thus providing a means to conduct statistical analyses of qualitative data. Additional information was gleaned by conducting statistical analyses of qualitative data that otherwise was not possible through qualitative means alone.

The third type of mixing found in this study was the creation of polytomous variables in the qualitative phase which were used to create the initial rubric. The variables were directly created from the themes generated by the participant interviews. The variables were created by determining which themes were most recurring and rated as most important.

The final type of mixing, linking, occurs in the discussion phase of this research. The quantitative findings of the items fit statistics and EFA will be linked to the qualitative findings from the qualitative phase. By linking qualitative and quantitative methods, an answer to the final research question will be provided.

The use of multiple mixing types during multiple phases ensured a thorough integration of data. This study was designed to mix the data at several key phases to provide a high level of triangulation as well as provide a deeper understanding of the phenomena being studied. According to Creswell and Plano Clark (2011), mixing should occur in at least two separate phases (Creswell & Plano Clark, 2011). This research study doubles the recommended amount of mixing to strengthen its definition as a mixed methods research study.

Limitations

As with any study, this study contained several limitations that may have affected the results to some degree. This study's limitations included time, sampling procedures, resources, and sample size. The process of designing a new instrument can take several years to produce a rigorous product. From a consensus of experts in the field which was then applied to a sample of mixed methods articles, this study was only designed to create an initial instrument which is tested at the pilot test phase. This instrument is not designed to be a finalized product. Much additional research will be required to move this instrument from the beta phase to the in-use phase. In addition to the limited time given during the dissertation process, resources also limited this study. A larger sample size would have been desirable, but the time to transcribe and the monetary resources to hire professionals to transcribe additional transcripts was not feasible. A large sample size might have provided a greater diversity of viewpoints on this topic. As a future research topic, additional funding and time can be allocated towards this project to improve the findings of this study. Given unlimited resources and time, articles would have been selected from multiple fields of study, including engineering, health care, and life science. An increase in articles would have meant an increase of personnel to review the articles. The benefits of a larger sample size include more accurate statistical analyses and a larger degree of generalizability. The final limitation came from my own personal attitudes towards quality of mixed methods research articles. Although this was addressed in reflexive practices during the qualitative phase of the study, individual biases and lenses cannot be fully removed from the narrative. These recognized limitations impacted

this study to an unspecified degree, but actions have been taken to limit each one so they have the least amount of negative impact possible.

Chapter 4: Article 1

What is quality? A look at what makes a quality mixed methods research study

Timothy Burrows

Virginia Tech

Abstract

Currently, in the field of mixed methods research, there are many different viewpoints on how to measure quality of mixed methods research. The purpose of this study is to conduct a cross-disciplinary, multi-phase study to create a preliminary guideline to evaluate mixed methods research articles in multiple disciplinary fields. This study utilized a content analysis of the current mixed methods literature to determine the current practices for evaluating the quality of mixed methods research. In addition to the content analysis, 12 mixed methods experts (eight males and four females) were interviewed to determine what elements of a mixed methods study are most important when evaluating quality. The findings of this study revealed that three major points of agreement include the idea that strict guidelines for quality may not work and in fact stunt the development of the methodology, the discussion of paradigms is important to provide a greater understanding of how and why a study was conducted, and that mixed methods studies should be transparent to allow for replicability and quality assurance. The findings of this study present the mixed methods research community with a better understanding of why certain elements of mixed methods research should be included when evaluating quality and why more standards and guidelines for evaluating mixed methods research may be most beneficial for the development and growth of mixed methods research.

Introduction

An important question to ask when reading a mixed methods research publication or report, as with any research, is one about quality. How can the reader judge if a mixed methods study was conducted well or poorly (O'Cathain, 2010). According to Creswell and Plano Clark (2007), it is important to be able to assess the quality of mixed methods research (MMR), but as of yet, there is no consensus on a standard to evaluate the quality of the research. Quality is important in mixed methods because it represents the credibility of the study (Creswell & Plano Clark, 2011). Although quality is an important factor in research, it is a difficult obstacle to overcome within the field of MMR since the ideas that define "defensible methods and warrented knowledge are fundamentally philosophical matters" (Greene, 2007; p. 165).

Quality assurance in MMR can be achieved in several ways. One of these ways is to use a set of guidelines that have been accepted by the mixed methods community to ensure the most important components of a mixed methods study are present. Examples of these guidelines include those presented by Creswell and Plano Clark (2011), Leech, Dellinger, Brannagan, and Tanaka (2009), and O'Cathain, Murphy, and Nicholl (2008). Those guidelines were designed to guide the researcher in their work, but are not used to assign a comparative value to the study. To assign a value, rubrics have been created to assess the level of quality for MMR. Two examples of rubrics are presented from the National Institutes of Health (2012) and Pluye, Gagnon, Griffiths, and Johnson-Lafleur (2009). These guidelines and rubrics present a snapshot of how quality is currently being evaluated in MMR.

This article aims to examine strategies for evaluating the quality of MMR and to create additional support for assisting in the evaluation of quality for MMR. I utilize both a review of the current literature pertaining to standards for evaluation of quality in MMR publications and report on my analysis of interviews conducted with experts in the field of MMR. Interviewing has provided a wealth of information about the current state of MMR and offers insight that is unobtainable through other means such as surveys and content analyses (Leech, 2010). Through the literature review and interviews, I will demonstrate the need for developing an additional framework for evaluating quality in MMR, and identify quality components that are important to providing and maintaining this quality.

This article presents the qualitative findings from the first phase of a larger multi-phase mixed methods study that used an exploratory research design with a sequential timing and quantitative priority. The purpose of the larger study is to conduct a cross-disciplinary, multi-phase study to create a preliminary guideline to evaluate MMR articles in multiple disciplinary fields. In the discussion section, findings from the qualitative phase of the study are considered in light of results from the larger study. The purpose of this study was to conduct a cross-disciplinary, multi-phase study to create a preliminary guideline to evaluate MMR articles in multiple disciplinary fields.

This study answered the following two research questions:

1. What are the common evaluation criteria found in the contemporary methodological literature pertaining to the design of mixed methods research?
2. What are mixed methods research experts views on the role of paradigms when evaluating the quality of mixed methods research?

Literature Review

As MMR becomes a more popular choice among researchers, a new method of evaluating quality is imperative for researchers to work collectively under an umbrella of universal understanding (Bryman, 2007; Gambrel & Butler, 2011). In order to work under some level of universal understanding, it is important to have a level of quality present in all MMR. Currently, several guidelines have been proposed to offer ways to evaluate the quality of MMR (Creswell & Plano Clark, 2011; Leech, Dellinger, Brannagan, & Tanaka, 2009; NIH, 2012; O'Cathain, Murphy, & Nicholl, 2008). Guidelines for evaluating the quality of MMR have been around for many years, but it has only been recently that the evaluation process has transitioned into the creation of rubrics for the purpose of scoring the quality of MMR (Pluye, Gagnon, Griffiths, and Johnson-Lafleur, 2009). Although this step is seen as controversial by many scholars, it demonstrates a methodology on the move, giving evidence of its evolution. This section discusses the importance of quality research and several examples of current guidelines for evaluating or assessing the quality of MMR. Although there were several examples of guidelines for what makes a good mixed methods study, the examples that were selected for this article were mostly in agreement with the selected definition of MMR. Along with guidelines, rubrics for evaluating MMR is also discussed.

Guidelines are a guide for researchers to reference in assisting them to create a study that meets accepted practices. Three examples of guidelines are presented here to demonstrate different approaches that can be taken to evaluate the quality of MMR. These guidelines offer a way for researchers to focus on various components of the research to ensure a level of acceptable quality. These three examples of current

guidelines illustrate how different scholars from various disciplines are attempting to address the issue of quality.

The first source of guidelines for evaluating MMR is found in Creswell and Plano Clark's textbook, *Designing and Conducting Mixed Methods Research* (2011). Creswell and Plano Clark lay out four principles used in the evaluation process to judge the quality of MMR publications. First, Creswell and Plano Clark examine the methods section, looking for both qualitative and quantitative data used in the response to research questions or hypotheses. The second step in the evaluation process according to Creswell and Plano Clark is a further look at the methods section to examine for rigor and persuasiveness, depending on the type of methods used. The third criterion for evaluating a mixed methods study is how the mixing of data was completed and where the mixing occurred. Finally, the use of mixed methods terminology was sought. Appropriate terms should be found in the title, research questions, and throughout the study in relations to design (Creswell & Plano Clark, 2011).

A second source of guidelines for evaluating MMR is found in *The Quality of Mixed Methods Studies in Health Services Research* (2008), by O'Cathain, Murphy, and Nicholl. O'Cathain et al. present a set of guidelines they refer to as the Good Reporting of a Mixed Methods Study (GRAMMS). The purpose of a framework for assessing quality is to provide understanding (O'Cathain, 2010).

O'Cathain et al. differs from the other sources by creating a comprehensive list of guidelines that point to a need for a deeper methodological discussion in each part of the study. In order to be considered comprehensive, the framework for addressing quality

should cover a wide range of topics being studied (O'Cathain, 2010). O'Cathain et al. present guidelines for a more holistic description of a MMR study.

The third source of MMR evaluation guidelines is found in *Evaluating Mixed Methods Research Studies: A Mixed Methods Approach* (2009), by Leech, Dellinger, Brannagan, and Tanaka. Leech et al., take a different approach to evaluation by focusing on a validation framework. This framework was developed due to the vast number of terms for validity and the numerous ways for defining validity (Dellinger & Leech, 2007). This validation framework addresses five components in an attempt to unify thoughts about validity in this area. The five components of the validation framework include the foundational element, the elements of construct validation, inferential consistency, the utilization/historical element, and the consequential element (Leech, Dellinger, Brannagan, & Tanaka, 2009). The first of these, the "foundational element provides evidence of construct validation through evaluation of relevant research that may offer insight into the meaning of data, interpretation of data, and consequences of the use of data" (Leech et al., 2009, p. 19). The elements of construct validation use traditional qualitative and quantitative construct validation methods, but are not limited to traditional means. "Inferential consistency refers to how inferences from a study are logically consistent with other elements of the study" (Leech et al., 2009, p. 21). The utilization/historical element refers to how the use of a research method is supported by previous studies in the use of similar data. The consequential element examines the consequences of how the data are used and the inferences made from the data. Leech et al. created a set of criteria to evaluate MMR from the standpoint of content validity. Content validity is used to determine if the inferences made within the study are

consistent as well as to determine if the data are relevant to the claims being made. These criteria for content validity leave much to the individual researcher to interpret.

Guidelines are not designed to measure criteria in a quantitative manor. In order to evaluate quality quantitatively, a scoring rubric must be used. In the case of article selection for a peer reviewed journal or acceptance for a conference proposal, it is often necessary to rank an article or report. This is where rubrics offer an additional layer to guidelines. The idea of assigning a score however, is not accepted by all researchers.

Here, two examples of rubrics for evaluating the quality of MMR are discussed. Unlike Creswell and Plano Clark's book, the NIH guidelines provide a way to score the research. NIH scores the significance of the study, the investigators participating in the study, the innovation of the study, the approach taken to conduct the study, and environment of the study on a nine point scale. The National Institutes of Health (NIH) offers a rubric for evaluating MMR in their application review standards (2012) for grants. The evaluation criteria set forth by the NIH is the most comprehensive set of criteria, pulling from several different sources, including the ones presented within this literature review such as Creswell and Plano Clark. NIH also presents four criteria for evaluating MMR applications. The first criterion is to determine if evidence is given for both qualitative and quantitative procedures. The second criterion is to ensure a body of literature is presented to support the components that are found in the study. The third criterion is to consult existing MMR quality criteria. The different quality criteria may include methods orientation, research process orientation, and the timing of phases found in the investigation orientation. The final criterion is a checklist for reviewing the application of MMR procedures. Studies with the higher score are seen as having a

higher quality than those with a lower score. This checklist may include the significance of the study, information about the researcher's skills and experience, the innovation of the study, the approach, and the environment.

In addition to NIH's rubric, Pluye, Gagnon, Griffiths, and Johnson-Lafleur (2009) propose a scoring system, entitled the Mixed Methods Appraisal Tool (MMAT), for MMR based on the idea that quality appraisal is important to a methodology. In addition to a scoring system that applies to qualitative, quantitative, and MMR, which is done because of the belief that all three research methodologies are inherently integrated, Pluye et al. also provides guidelines for conducting a MMR study (Pluye et al., 2009). Pluye et al. reviewed 59 qualitative, quantitative, and mixed methods articles and coded them to create generic categories about what design components the articles contained. These categories led to the creation of the scoring system. Although the scoring system was created to measure the quality of the three research methods, only 3 of the 14 criteria are used to measure MMR studies specifically. The three criteria used for mixed methods studies include (1) rationale of the mixed methods design, (2) use of both qualitative and quantitative data collection and analysis procedures, and (3) integration of qualitative and quantitative data and/or results (Pluye et al., 2009, p. 540). Since the MMAT's creation in 2009, the MMAT was tested for reliability, revealing that the scoring system provided a high level of inter-rater reliability for 17 of the 19 criteria (Pace, Pluye, Bartlett, Macaulay, Salsberg, Jagosh, & Seller, 2012). Seventeen of the criteria received a Kappa score of .94, with the remaining two receiving a score between .83 and .89 (Pace, et al., 2012). The MMAT provided a tool for appraising quality in MMR and provides a strong foundation for future scoring rubrics (Pace et al., 2012).

By reviewing examples of guidelines and rubrics, it is apparent that there are both similarities and differences between them that go beyond their defined purposes. Most of the examples selected here discussed the importance of using qualitative and quantitative data to add information to the study. In addition to the value added component, the way and purpose for the integration of the data was also a common feature for evaluating a quality mixed methods study. Lastly, the methods for conducting the study are given a high priority in both the guidelines and rubrics. Several common features reveal a collective understanding of quality, but several differences also expose a level of individuality.

Differences between the selected examples show how the understanding of MMR varies between researchers. Differences are seen in the emphasis placed on literature reviews and the rigor of a study. While some examples emphasized the literature and rigor, another recognized the importance of using appropriate vocabulary. The largest difference seen between these examples is their purpose. The guidelines cover a wide array of topics to assist the researcher, while the rubrics are designed to score and assign a value to the study.

Methods

This study has been conducted from the perspective of a pragmatic paradigm, using the knowledge gained in this study to create a product that can benefit the practice of MMR (Schwandt, 2008). Both of the research questions are answered through a literature review and by conducting interviews with experts in the field of MMR. This section details how the qualitative data were collected and analyzed, and how the trustworthiness of the study was ensured.

Sampling

The sampling plan used in this phase of the study was a purposeful sample of MMR experts. The sample was created by contacting a group of 52 experts from the MMR field. "Experts" in this study is defined as any author who wrote or co-wrote a chapter in the second edition of the *Sage Handbook of Mixed Methods in Social & Behavioral Research* (2010). The authors of these chapters are considered experts because of this book's recognition as a leading textbook in the field. The authors within this text have also displayed a significant contribution to the development and progress of MMR as a methodology.

An email was sent to each of the 52 authors and co-authors requesting their participation in this study. Two electronic reminders were sent out to non-respondents in seven day intervals following the initial mailing. A total of 20 individuals responded from the email solicitation, or a response rate of 38%. Of the 20 initial respondents, 12 agreed to participate in the study (23%). The relatively low response rate may be explained in part by the timing of the invitation to participate and the timeline set for completing the interviews. The invitations to participate were sent out in late May with a July deadline for conducting the interviews. Many participants may have been out of the office for summer or already committed to other projects. The short timeline was necessary to keep within the timeframe for the study's deadline.

The 12 respondents represented a multi-national and multi-disciplinary community. The twelve participants included eight males and four females. Two of the participants were faculty at universities located in the United Kingdom, one participant was a United Kingdom resident working in the United States, and nine were faculty

members from the United States. The participants also varied in years of experience, ranging from approximately five years to over thirty years. The participants also varied in methodological backgrounds, including researchers that operate primarily from either a qualitative and/or quantitative paradigms.

Data Collection

The interview protocol (Appendix) reflected the components of MMR as found in current literature regarding quality. The protocol questions focused on why quality is important and what components of MMR are the most important and why. Questions were specifically asked about the component of mixing data because of its importance and of its often misuse or mislabeling. In addition to the topic of mixing, specific questions also addressed the role of paradigms and their importance within MMR. Finally, in addition to quality, important components of MMR, mixing, and paradigms, a question was also included about the viewpoint of MMR based on the participant's discipline. Following the principle protocol questions, the participants were asked to participate in a follow up questionnaire, which will be covered in more detail in a future article.

The protocol was approved by the Virginia Tech Institutional Review Board (IRB) to ensure that any harm to the subjects was kept to an absolute minimum and that the protocol was carried out in an ethical manner. Consent for the interview was obtained by either written or verbal permission prior to the interview. This was addressed in the second email confirming the participation of the subjects. A copy of the consent form was sent with a copy of the protocol with the confirmation email. Participants then had

the choice to electronically sign or give a verbal agreement prior to the start of the interview.

The interviews were conducted via telephone or the computer based communication software Skype. Both telephone and Skype were given as an option to give the participants more flexibility in using a technology that was more comfortable to them. In addition to flexibility, Skype offered the ability to conduct international interviews at a minimal cost. Although Skype offered a free alternative to international calling, the quality of the transcripts were adversely impacted. This impact was minor, but some data were lost due to the quality of the call. All of the conversations were audio recorded using a digital recorder. Once the interviews were complete, the digital recordings were outsourced to a local transcription service. The transcription service provided secure transactions of data files through the use of a personal Dropbox folder and was returned with the names of the participants removed. The transcription service also agreed not to release any of the information in regards to the confidentiality statement of the IRB. This service was selected based on high recommendations from other researchers and previous experiences with past research projects. Pseudonyms were used to maintain the confidentiality between the researchers and the participants. The names and locations that were given as part of examples were also changed to protect other individuals and institutions. The final versions of the transcripts were then edited for clarity and/or reworded to remove additional references that might personally identify the participant.

As the completed transcripts were received, an inductive coding process occurred for each interview. This inductive coding process included line-by-line coding, which led

to the creation of several categories. In this study, a category refers to a recurring idea that is created from an inductive analysis of qualitative data. The initial coding process generated 125 codes. Some examples of these codes included engagement-active, paradigm-purpose, and mixing-challenges to name a few. These 125 codes were then collapsed into categories which were created by grouping similar codes together. The initial 125 codes were collapsed into 20 categories. Examples of categories included quality-value added, transparency, and mixed methods research-evolving. Several of the initial codes, such as quantitative-role, were removed due to lack of connectivity to the research questions or a lack of general connection to the topic of study.

Following the finalizing of the initial coding scheme which was conducted by hand, the transcripts were imported in NVIVO 10 qualitative software and then recoded. The continuity in coding between the inductive line-by-line coding and the deductive coding by categories was most likely over 90%. This percentage was calculated by comparing the number of initial codes to the number of categories, which included multiple codes. A generalized word count also supported evidence of this percentage. Due to the use of different software at each process, this number is only an approximation and an exact number cannot be obtained.

Once the coding was complete and a list of categories was generated, an intensity matrix was developed to document the occurrence of categories. The matrix was created by using x and y axis chart giving the name of the categories and how many times it occurred throughout the interviews (table 1 in the findings section). In addition to providing information on frequency, the table also identified how many of the

participants discussed certain topics. Both the coded interviews and the intensity matrix were used in the analysis of this study.

Analysis

The analysis of this study was done through the use of a literature review, interview data, and an intensity matrix. The literature review provided a detailed understanding of the criteria that are offered in the literature to evaluate the quality of MMR. The data from the interviews were then used to both support and refute common practices found within the literature pertaining to the evaluation of quality. Finally the intensity matrix was designed to connect the amount which topics were discussed in the interviews and how many individuals commented on them to the literature on this topic.

Trustworthiness of the analysis was ensured through the use of member checking, an audit trail, peer review, and reflexivity (Patton, 2002). The transcripts were sent to the participants to make certain the accuracy of the interview and to allow the participants the chance to add to or correct any previous statements. Participants responded mostly to correct the spelling of names and fill in some of the missing segments of the interviews that were lost due to the quality of the recoding (mostly in the cases using Skype). In addition to the member checking, an audit trail, made up of procedural memos stored in NVIVO, was created to trace and outline the decision making processes of this study. A peer review of the coding schemes among the co-authored supported the member checking and audit trail by creating a level of acceptance of over 80% for the selected codes. Finally, reflexivity, or the process of being self-analytical, defined my own assumptions as a mixed methods researcher and biases that may have been introduced to the study (Patton, 2002).

In addition to creating a trustworthy study, one of the goals of this study, as with any study, was to provide a level of transparency that could lead to its replication. In order to maintain these factors, the primary limitations of this study must be addressed. The largest limitation may be the sample size and the participant population. Individuals who disagreed with the idea of rating quality were quick to address their disapproval in initial emails. This may have led to a population of similar thinking individuals, preventing a wider scope of data. A large sample size might have provided a greater diversity of viewpoints on this topic. It would have been valuable to find a participant who opposed this study who would have agreed to be interviewed. Several individuals also commented on their unavailability due to busy summer schedules. In addition to the sampling, my own personal bias as a beginning researcher in the evaluation field may influence the findings of this study. Many attempts were made to limit this bias, but it cannot be removed completely due to the human element of research. The final limitations are the lack of monetary resources and time. Given a larger budget and timeframe, additional participants could have been interviewed, thus providing a larger sample for greater generalizability.

Findings

The findings detailed in this article do not represent all the facets of the interviews. These findings reveal the disciplinary differences and similarities that researchers have towards the evaluation of quality in MMR. In addition to the similarities and differences about quality, the role of paradigms in MMR is also examined. First, findings are presented that describe level of engagement to provide support for the participants being considered experts in the field of MMR. Secondly, information is

provided to detail the differences of viewpoints that were provided pertaining to the use of guidelines and rubrics to measure the quality of mixed methods studies. In addition to the viewpoints about quality, viewpoints on the use of paradigms and their added value to a mixed methods study were also given. Finally, the importance of a MMR study being transparent is addressed. First, the intensity matrix of categories is examined to provide a way to identify where a consensus among experts about evaluation and criteria where there is less agreement.

Intensity Matrix of Categories

Table 4.1 provides the frequencies for the categories that were developed from the interviews conducted with the 12 participating experts. Each of the categories is listed in this table below. The table provides the number of responses by each participant for each of the nine categories used in this findings section. A number of three or more is considered to be an in-depth discussion because this represents a larger portion of the interview spent on one category over others.

Table 4.1

Intensity Matrix of categories depicting evaluation of mixed methods research, participant engagement, mixing, paradigm use, standardization, and transparency of evaluating quality.

	Engagement - longevity	Engagement - advancement	Criteria - evaluation	Criteria - using multiple methods	Criteria - quality	Mixing - purpose
Interview 1	1	1	1	0	1	2
Interview 2	1	1	2	0	1	7
Interview 3	2	0	1	0	1	2
Interview 4	3	0	1	1	0	0
Interview 5	3	0	1	0	0	4
Interview 6	2	2	0	0	0	0
Interview 7	1	2	1	0	0	3
Interview 8	2	0	2	3	1	3
Interview 9	1	0	1	0	0	2
Interview 10	1	1	2	1	0	2
Interview 11	2	1	0	0	1	7
Interview 12	1	1	3	1	5	1
TOTAL	20	9	15	6	10	33
Response %	100%	58%	83%	33%	50%	100%

Table 4.1 cont

Intensity Matrix of categories depicting evaluation of mixed methods research, participant engagement, mixing, paradigm use, standardization, and transparency of evaluating quality.

	Mixing - study	Paradigm - differences	Paradigm - multiple	Paradigm - purpose	Standardization	Transparency	Total
Interview 1	0*	1	1	2	1	2	13
Interview 2	1	2	0	4	1	1	21
Interview 3	2	3	0	3	3	0	17
Interview 4	1	3	0	2	7	4	22
Interview 5	5	0	1	3	0	1	18
Interview 6	2	0	3	3	2	2	16
Interview 7	3	0	1	1	2	3	17
Interview 8	3	1	1	2	0	0	18
Interview 9	1	0	1	2	4	0	12
Interview 10	2	2	0	4	0	2	17
Interview 11	2	0	0	3	1	0	17
Interview 12	1	1	3	1	3	1	22
TOTAL	23	13	11	30	24	16	**210**
Response %	100%	58%	58%	100%	75%	67%	

Note: * denotes that this participant was not asked this question due to time constraints. These categories represent the topics presented in this article.

Engagement of the participants in MMR was detailed by two categories, "engagement-longevity" and "engagement-advancement". Engagement-longevity refers to examples of the length of time that the participants have been participating in the mixed methods community. Engagement-advancement is the engagement between the participants and others to promote and/or advance the field of MMR. All of the participants (100%) discussed their experience using mixed methods. Question one of the protocol (Appendix) asked about the participants involvement in the mixed methods community and all the participants included the amount of time they have been practicing

or participating in the MMR. Of the 12 participants, 7 (58%) gave examples of how they are promoting the advancement of MMR. Examples of this are described later in the paper. Two participants discussed engagement-longevity in depth by making at least three references to this category. These categories were mostly answered in one or two extended statements.

When asked about evaluating MMR, three categories were developed. These three categories include "criteria-evaluation", "criteria-using multiple methods", and "criteria-quality." Criteria-evaluation is comprised of examples of criteria that are used to evaluate MMR as a whole. Criteria-using multiple methods includes examples of evaluating the quality of individual methods used within a mixed methods study. Criteria-Quality refers to the relationship between the quality of a mixed methods study and the evidence of the study being robust, evidence based, and/or reflecting on what was learned from the study. Ten of the 12 participants spoke directly about the criteria they use for evaluating MMR (83%). In addition to speaking directly about evaluating a mixed methods study as an aggregate, four of the 12 participants (25%) focused on criteria for evaluating the individual methods that make up a mixed methods study. These four participants focused on evaluating the individual qualitative and quantitative strands according to their own current methodological practices. This is an example of a holistic evaluation of quality compared to a more internal evaluation of quality. For evaluation purposes, six of the participants associate a mixed methods study as a singular method, while four participants associate a mixed methods study as being a grouping of qualitative and quantitative methods. Of the 12 participants, only six (50%) made explicit references that MMR should be evidence based, robust, and reflect what is learned from the study. In

addition to the three categories of criteria, a fourth category of "standardization" was also used in this section to provide insight on the views towards universal guidelines for evaluating MMR. Of the 12 participants, nine (75%) discussed issues concerning the standardization of evaluation criteria for evaluating the quality of MMR. Each of the three categories of criteria had one participant who spoke in-depth (more than three times) about criteria. Four of the participants spoke in-depth about the standardization for evaluating the criteria for quality in MMR. Participant 4 spoke more about standardization (seven times) than any other participant on any other category. The categories of criteria had the least amount of in-depth discussion. This may be due to quality being a relatively unaddressed topic in MMR.

The purpose of mixing along with how, why, and where it should occur in a MMR study were acknowledged by the categories "mixing purpose" and "mixing study." All of the participants (100%) provided responses to both of these categories. Five of the participants (42%) made at least three references to "mixing-purpose", while three (25%) made at least three references to "mixing-study."

The role of paradigms in MMR was addressed by three categories. These categories were: "paradigm-differences", "paradigm-multiple", and "paradigm-purpose." Paradigm-differences is defined as examples of how paradigms differ between people based on personal bias, position, and definition. Paradigm-multiple applied to examples of how the use of multiple paradigms within one study. Paradigm-purpose is the role and purpose of paradigms in a study and why they are important to the study. Paradigm-purpose had the most references. Six of the participants (50%) made at least three separate references to the purposes that paradigms play in MMR. The participants

provided mixed responses when directly asked about the importance of paradigms in MMR. Seven of the researchers (58%) discussed the issue of differences related to the use of paradigms within mixed methods studies. These differences include the definition of paradigm to the role of the paradigms used in the study. The same number (58%), although different participants discussed the use of multiple paradigms within a single mixed methods study and the implications that may be connected to them. One hundred percent of the participants addressed the overall purpose of paradigms in MMR. Two participants spoke in-depth about paradigm-differences, while two spoke in-depth about the paradigm-multiple category.

The idea of transparency in MMR emerged as a distinct category. The category of "Transparency" refers to examples about the general role of transparency and what it adds to a mixed methods study. Eight of the participants (67%) discussed the importance of transparency to a mixed methods study. Although this topic did not surface in all the interviews, it was addressed as important in each instance it was mentioned. Two participants referred to transparency in-depth.

Each of the five groups outlined above are discussed in greater detail in the following sections. Quotes taken from the interviews provide an in-depth look at how both similar and different viewpoints occur in the roles of quality, paradigms, and transparency in the emerging culture of MMR.

Levels of Engagement

The participants for this study represent a group of scholars that have been connected to the field of MMR for many years and have made numerous contributions to

the advancement of the field. Participant 11 spoke of his long career in MMR even before the field had a name.

> Well my career is 40 years long, and that's how long I've probably been using mixed-methods. When I started, I'm not sure there even was a phrase or a field called mixed-methods research, actually, although since then it's emerged. My advisors and experience in graduate school was in the cross cultural study of child development and parenting, and I did field work in [name of a country], I've done many studies [within and outside the US]. And I was trained by [names deleted], who are anthropologists who did disciplinary and multi method research on child development.

While participant 11 spoke about his long career in mixed methods before it was considered a formal research methodology, another participant discussed her career as it pertained to conducting research during the initial development of MMR. Participant 3 has been working on mixed methods for over 30 years and has contributed to the literature in many ways. Participant 3 described her contributions below.

> I have been involved with mixed methods research and thinking about mixed methods since the early 1980's. I'm an [evaluator], and I think our work, being kind of the ultimate practical social science, calls out at times for outside the box in terms of methods. And it was through some practical work at that time that I, along with a couple of colleagues, said we need to think more seriously about solving all these disagreements and paradigm wars and getting on with the business of doing good social science. So that's kind of how it happened…how it started; with a couple of graduate

> students we did [a study]…and that launched my scholarship in the area. I guess prior to that there were a couple of articles that were reflecting on our evaluation practice, and why mixed methods seemed to be a good response to that. When a couple of graduate students did a review, which was published in the late 1980's, that formally launched my scholarly interest in the field. So I've been engaged for quite some time. I consider myself still engaged with the mixed methods community and the interest in it.

Of the 12 participants interviewed in this study, at least eight of the participants have careers that span over 20 years in the field of MMR, but participant 10 may be the most prolific writer of the group. Participant 10 detailed his career as follows.

> I participated in a variety of ways: authored a couple of textbooks, to writing numerous chapters about mixed research. I had the pleasure of co-authoring more than 120 works on mixed research, mixed methods. Being an editor of a sorts, for a special issue on mixed-methods research, being a guest editor for the [mixed methods journal], being on the review board of the [mixed methods journal]. And I've attended [multiple] conferences, of the International Mixed Methods conferences.

In addition to long careers and making substantial contributions to mixed methods as a methodology, participants also provided examples of the use of MMR outside the academic circle. Participant 2 provided information on how he uses his expertise to help others improve their mixed methods knowledge.

> I have taught a two-day class on using mixed methods in development research and program evaluation which gets 30 people or so every year…so there's a lot of interest within (our agency)…and we get not only staff members taking that class, but people from other agencies in the [metropolitan] area. So it's something we started out doing because we wanted to figure out how to combine qualitative and quantitative methods in our own joint research initially and our separate research since then…but it's just been one thing that's…either through our own advocacy or through just the range of different problems that people now find themselves confronting…the need to figure out how to judicially combine different methods in response to particular problems has become a more routine issue in a range of different disciplines and fields.

The participants of this study not only provided decades of experience, but they also have contributed a large number of widely cited publications pertaining to MMR. Through their research, presentations, and instruction, these individuals have promoted the growth and development of MMR, thus confirming them as experts. Although the participants have a commonality by their contributions to advance the field of mixed methods, they held a variety of different viewpoints when it came to evaluating the quality of mixed methods.

Different Viewpoints of Evaluating Quality

During the interviews with these 12 experts, it became apparent that one topic elicited a good deal of disagreement. That was the idea of standardizing the criteria for evaluating MMR. The 12 participants all disagreed with a standardized set of guidelines

to measure quality. Participant 4 proposed more of a hard-lined approach towards the use of a rubric for evaluating quality in a mixed methods study. Participant 4 made two points in his response.

> I'm not sympathetic to the view that we need to establish uniform standards of some kind that people can use to evaluate a particular study…that is--a list of criteria, because to me it is not that simple. I don't think there's a rubric you can use for that; I think it depends very much on the specific study.

At the same time, the need for certain commonalities in an evaluation rubric was addressed. Participant 3 perhaps best articulated this need for quality without applying one standard to all MMR while noting that the field is still emerging. Participant 3 said,

> we have to have widely accepted ways of establishing a warrant or inferences…saying that these are worth paying attention to. I also think it's a real test of mixed methods as its own methodology, which maybe it is and maybe it isn't, and I don't know that I care whether it is…it's not that I think oh it has to be…I think maybe, maybe not, I don't know. The jury is out on that. But I think it's a real test…can mixed methodologists and practitioners come up with ways of establishing warrants…I don't want to use validity or credibility…but ways of saying that these come from the data, they match the context, they state something of worth. I think it's a critical challenge. I think there's some good work out there. Again, I especially don't think that we need to come up with one set of criteria or one way of saying that this is quality work, but I do think we

need to come up with some ways (to measure quality work). Otherwise we generate inferences and we don't if they're any good.

Participant 7 offered an explanation for why it is not yet time for a single standard to evaluate MMR, while at the same time suggesting that some standards may be beneficial.

I think that we needed to be messy for awhile because the field really isn't ready to come to consensus about anything, and until we have this time of exploration in figuring out where we want to go as a field, I think it's too early. So right now I think that it's good that it's messy. I think though eventually for the sake of the field, and especially for new researchers conducting mixed methods studies, we do need to come to some kind of consensus where we have language that is similar. Like in quantitative research, if you talk about multiple regression, everyone knows what that is; there's no questioning. Well what is that really? What is your definition of that? On the other hand with mixed methods research, we don't have that kind of common language, and so that makes evaluating other people's mixed methods studies very difficult because they'll use language that you're not exactly sure what they're really meaning. I think one of the things that can help with that, while the field is still messy, it would be helpful if researchers would cite other people, so if they say they're doing a certain type of mixed methods design, cite who coined that term, so then the reader would know that this is what the author is talking about.

All 12 participants agreed that standards or a single set of guidelines to evaluate the quality of mixed methods is not practical. Although there was a consensus that a single set of guidelines is not practical, the participants did have several differences in their specific viewpoints. Participants 1, 4, and 6 felt that it was not at all appropriate to have standards, while participant 3, 7, and 9 did not agree with one standard, but did acknowledge that some standards were beneficial. Participant 12 added to the idea of flexibility to standards. One standard cannot fit every study, but a set of flexible standards can create a level of acceptable quality. The participants of this study have made it very clear that the idea of standardization is a contested topic; one that will continue to be debated as mixed methods grows out of its infancy.

Viewpoints on Mixing

With MMR based on the principle of integrating both qualitative and quantitative data, the participants were asked about their opinions towards what degree mixing is necessary in a quality mixed methods study and how much mixing should occur. Participant 12 spoke in depth on two purposes of integrating data within a study.

> One is synergistic legitimation, which is talking about how one can argue and justify the study based on what's gained from the two [methods] together. But another one is called integration legitimation. Because that's something… if you go to the Journal of Mixed Methods Research and search on the key word "integration," you're going to see that virtually every article talks about that, almost every article. Because in a loose sense what one needs to show is that these two things together are integrated least in thought and conclusions to show what meta inferences

> or what better value one has from doing this study, rather than say two separate studies.

Participant 11 presents that the purpose of mixing data is to provide more information than one single methodology can provide on its own. On the subject of how often this should occur in a study, Participant 1 provides an insight on how to choose the amount of mixing and why mixing should occur in a study.

> I would say that mixing is really, really important. In terms of the degree of mixing, I would say like a continuum—, one or two phases, all the way through, every single phase. But the more phases that you mix, including when you get into mixing of [methods] and paradigms and so forth, the more mixed the study becomes, the more complex the study becomes and the more imbedded it becomes. So for me, those kinds of studies take me to another level.

Participant 1 stated that mixing should occur to form a synergy between the methods used and that the more mixing that is done, the more integrated the data becomes, but is not necessarily in a specific amount. Mixing is important to the quality of a MMR study because it provides the purpose for combining qualitative and quantitative data sets with a synergy to provide additional information that can only be revealed through MMR.

Different Viewpoints of Paradigms

When reviewing the literature about what constitutes quality in a mixed methods study, paradigms are sometimes discussed and other times ignored. To get a better understanding of the place of paradigms in the process of evaluating quality, each participant was asked their opinion about paradigms and their role in quality. Like the

viewpoint on the use of rubrics and guidelines, participants expressed a variety of views about the role of paradigms in MMR. Eleven of the 12 participants (92%) agreed that paradigms serve an important role in providing a level of quality to a study.

Participant 4 touched upon the importance of paradigms and what they add to a study by saying,

> the philosophical belief that a researcher has about the phenomena they're studying and about the nature of research in general—yes, I think those are important. They definitely shape how the researcher proceeds. They are lenses that show you some things and obscure others, they create some validity threats, but you couldn't do research without them. It's like trying to see, the lens in my eye is distorting what I see.

Participant 6 also identified the need for understanding the researcher's motives behind their work. Without a framework to interpret the study as was meant by the researcher, meaning may be lost. Participant 6 expanded on the points made by participant 4 in the following passage.

> I think…conceptually…any time you do any kind of research where you're going to do any kind of meaning making, when you're going to interpret something, you've got to have a framework so that you can sort of justify the conclusions you have drawn and the interpretations you make. And so just like the methods you talked about earlier, a paradigm you use, you've got to hook your meaning on something. Kind of the theoretic framework for a paradigm is you're going to…let me see if I can make a sentence out of this…you can't just claim things out of the blue,

> you've got to have hooks to put your meaning on, or there's nothing to put it on. And so using a paradigm, whether it's received or tweaked, I think is a necessary framework to be able to make any kind of claim. You can't just say something is true because it's true, or something's this way because it's this way, you've got to have a contextualizing framework to be able to make any kind of claim.

Only participant 9 had the viewpoint that paradigms are not necessary when evaluating the quality of a study and that they are not as important as other factors. Participant 9 described his feelings towards the importance of including paradigms in a mixed methods study. He specifically addressed the difficulty in making arguments about paradigms. The differences in philosophies can become a point of contention. For this, at least in part, Participant 9 stated,

> We followed up with some articles in [a Journal] and one of the things that we actually in a sense avoided is the idea of paradigms, and one of the reasons for that it seems, it's a point of contention, where it's that whole idea of…let's see how can I put it…I've had some really interesting conversations with people about things like mixing methods, I mean literally there are some people in my university who believe it is impossible to mix methods…and you cannot mix philosophies…I'm like okay, I didn't know there were any pure philosophies, but sure! To me, it's probably not as important as other things, and it's probably because in my brain I've kind of downplayed it just because it really is seemingly a harder argument to have. I think it's something that can be left out.

Eleven of the 12 participants agreed that paradigms are important to the quality of a mixed methods study. Paradigms provide a framework for why and how a study is designed and carried out. Paradigms provide the motives, which is a clue to properly interpret a study. Without the appropriate support, the reader may find it difficult to arrive at the same understanding as the author or another reader. It is the ability to provide this structure that allows the proper interpretation of the study that makes it an important piece of the quality evaluation puzzle. The use of paradigms contributes to a higher level of transparency for a mixed methods study.

Transparency

Although different viewpoints emerged during the course of the interviews about quality and the use of paradigms, the one point where there was agreement was the importance of a study being transparent and how that impacts the overall quality of MMR. Mixed methods research offers a variety of definitions, designs, and ways of integrating data.

One example where transparency is useful in evaluating the quality of a mixed methods study is the concept of mixing data. Some researchers prefer the term mixing, while others prefer integration. To some they mean the same, while to others they are two different concepts. As for the design portion, mixing can occur in one or more phases. The amount of mixing also varies between designs. How the data are mixed is an additional factor in the study. Does the mixing occur as the linking of data or are the data transformed? It is simple to understand the level of complexity in MMR just by briefly examining one concept. Transparency ensures that the reader is of the same understanding that the researcher intended.

In responding to what to look for when evaluating the quality of MMR, Participant 10 discussed how transparency can help the reader understand a study better and placed particular emphasis on the importance of clarifying the design.

> As long as they state what they're doing, and as long as they're clear, I'm very much a visual person so I do like where possible to see some kind of visual representation—especially if it's a complex design, that always helps, that maps out the design. But there are lots of designs out there—that's one of our strongest areas. There are more designs than we know what to do with. There are probably over 100 different types of designs now out there. And so yeah, it's just a question of choosing one that's appropriate and making it clear. And not assuming that the readers will be aware of that design, so presenting it in a way that gives enough information that someone who's never ever seen the design can get a grasp, and also giving appropriate citations.

It is often unclear to what needs to be included in a mixed methods article. Participant 7 spoke of a situation where the lack of information makes it difficult to understand a study. Here is her account.

> One of the things that some of my colleagues that I've written with and I have struggled with is many mixed methods studies when they're written up don't include all of the aspects that need to be there. I review for [multiple journals], and they get mixed methods studies occasionally. What's frustrating is that people who write about mixed methods studies don't include enough information many times and I think it's because

there's this strict page limit, and with a mixed methods study you really need extra pages to include everything, and if you don't include everything, it's sometimes very difficult to understand what people have done in their study.

The participants of this study have elaborated on why being clear is important and what it adds to the overall quality of a study. Transparency is the piece of a mixed methods study that transcends the boundaries of multiple definitions, designs, and differences in how the data are integrated. Transparency allows the reader to understand the paradigm in which a study was conducted. A high level of transparency can also clarify the design and give enough details to allow for replicability. A transparent study allows for the reader to identify the robustness and trustworthiness of the study and to identify what the study adds to the body of knowledge. Finally, transparency can allow the reader to determine if mixing was done appropriately and if it added to the findings of the study. Transparency is particularly important to MMR because it is, as Participant 3 described, "messy" at this point in time.

Of all the categories discussed in this section, transparency may be the most important because it provides a lens or maybe better put, a cleanser, to evaluate quality on multiple facets. The information gleaned from these interviews can allow us to reconsider what makes a quality study and how this differs from previous discussions on this subject.

Discussion

The interviews conducted for this study revealed that experts in MMR hold diverse views about the evaluation of quality, the use of paradigms, and transparency.

The twelve experts interviewed in this study have an array of experience and come from a diverse background, thus providing an in-depth look into the views towards the overall quality of a mixed methods study, the purpose of mixing, the use and role of paradigms in MMR, and the importance of transparency within a mixed methods study.

The first viewpoint to be examined is the viewpoint of evaluating the quality of MMR. Through the interviews with the twelve experts, it became clear that quality is an important aspect to MMR, just as it is with other research methodologies. This viewpoint supports Creswell and Plano Clark's (2007), discussion on the necessity of quality in MMR. The difference in viewpoints came from how to measure the level of quality within a mixed methods study, a problem that has been addressed by both Creswell and Plano Clark (2007) and O'Cathain (2008). Standards may be necessary, but they cannot be rigid due to the complexity of mixed methods designs. This can be seen in the GRAMMS and NIH standards of reporting a quality mixed methods study. Both of these methods use a broad set of guidelines instead of definitive standards.

In addition to the complexity, mixed methods research is still in an early stage of existence compared to both qualitative and quantitative methodologies. With being in an early phase of its development, the complexity of MMR has created a "messy" environment. As O'Cathain et al. (2008) reported, MMR is too early in its infancy to create a standard set of evaluation criteria. Mixed methods research is not exact in its methodology as it continues to grow and adapt to a changing environment. The lack of an exact methodology is why it may be important to have many standards so that the researcher has an option of choices that best fit his or her research design. MMR is

flexible and a rigid set of guidelines may devalue this flexibility and hinder its future development.

The second viewpoint to be discussed is that of mixing and its relation to quality in MMR. Mixing allows qualitative and quantitative data to communicate with each other. In the process of communicating, mixing also allows for new information to emerge that otherwise may not be possible through the use of singular methods. As Participant 11 stated, "the goal is to bring as many methods to bare of trying to understand a certain topic." Mixing allows researchers a different way to understand meaning, context, and experience. The amount of mixing is not an exact method. The amount should be based on the purpose and questions asked by the researcher. More mixing may provide a more complex study, but may not add to its quality. Some participants felt that more mixing is better, while others focused on the why mixing occurs and if the design is appropriate.

The third viewpoint to be examined is that of paradigms within MMR. Eleven of the 12 participants were in favor of including some level of information about the paradigm used by the researcher when reporting MMR. Paradigms provide the reader an understanding of how the research was conducted and how to better understand the study. Different paradigms lead to different interpretations of a study. By providing a background to a study, this inherently provides a level of quality that may be missing if paradigms are not disclosed. Since mixed methods studies are often conducted by teams of researchers, multiple paradigms may also be included in one study. Paradigms within mixed methods have often been addressed, but they are not linked with measures of quality. By being reflexive about the role of paradigms, researchers not only provide an additional measure of quality, but also help raise the level of transparency of the study.

Currently there is a lack of literature detailing the connection between evaluating the quality of MMR to the use of paradigms. Many articles focus on the use of various paradigms or differences between paradigms in MMR, but few specifically linked to their role in evaluating quality.

The participants of this study discussed why quality is important in MMR and some of the challenges that make the evaluation of quality a difficult subject to agree on. Three major points of agreement include the idea that strict guidelines for quality may not work and in fact stunt the development of the methodology, the discussion of paradigms is important to provide a greater understanding of how and why a study was conducted, and that mixed methods studies should be transparent to allow for replicability and quality assurance. It is in these common ideas where the discussion of evaluating quality may have the best chance of becoming widely accepted among mixed methods practitioners.

Conclusion

Quality is an important factor in research and mixed methods is no different in this aspect. Experts in the field of MMR have expressed different views when it comes to quality and how to evaluate it. Researchers with different backgrounds have all agreed that although quality is important, a standardized method for evaluating it may not be the best approach. Participants unanimously agreed that rigid criteria are not productive and should be addressed with care. Agreement emerged about what can be considered evaluation criteria, including the importance of paradigms and the need for transparency. Participants also mostly agreed that some level and variety of standards can be useful to promote quality mixed methods studies.

To assist in creating a quality study, paradigms add to the quality by providing information that allows the reader to better understand the view of the researcher and increase the transparency of the study. Without paradigms, the reader can misinterpret a study or arrive to false conclusions. Understanding that limited space can make an in-depth discussion of paradigms difficult, only a few sentences can provide the necessary background to limit the problems that may stem from the absence of paradigms.

Even though there are differences in viewpoints about conducting MMR and evaluating its quality, one constant viewpoint was the idea that a study must be transparent. The idea of transparency allows the reader of a study to not only replicate a study, but also be informed enough to evaluate the overall quality of the study. Being transparent includes the researcher discussing the research questions, purpose, design (including why and how the data are mixed), and results to name just a few. This is where guidelines for quality may help researchers. Often it is unclear of what is necessary to be included in a MMR study in order to provide a level of transparency that most readers will find acceptable to ensure a high level of quality.

Several guidelines and rubrics for evaluating the quality of mixed methods exist, but there is still room for additional sources on this topic. It may be wrong to state that there is only one way to define or assess quality, but multiple sources allow for researchers to advance the discussion of quality and choose pieces that are most pertinent to particular studies. As MMR grows and progresses, so should the idea of how its quality is measured. The next phase of this research study will provide a dual-purpose guideline, one with both dichotomous and polytomous scales, for evaluating quality

based on these experts points of agreement, while remaining cognizant of the criticism towards standards.

References

Bryman, A. (2007). Barriers to integrating quantitative and qualitative research. *Journal of Mixed Methods Research, 1*(1), 8-22.

Bryman, A. (2006). Integrating quantitative and qualitative research: How is it done? *Qualitative Research, 6*(1), 97-113.

Creswell, J. W. (2009). *Research Design: Qualitative, quantitative, and mixed methods approaches* (3rd ed.). Thousand Oaks, CA: Sage.

Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research.* Thousand Oaks, CA: Sage.

Dellinger, A. B., & Leech, N. L. (2007). Toward a unified validation framework in mixed methods research. *Journal of Mixed Methods Research, 1*(4), 309-332.

Gambrel, L. E., & Butler VI, J. L. (2011). Mixed methods research in marriage and family therapy: A content analysi. *Journal of Marital & Family Therapy* .

Greene, J. C., Caracelli, V. J., & Graham, W. F. (1989). Toward a conceptual framework for mixed-method evaluation design. *Educational Evaluation and Policy Analysis, 11*(3), 255-274.

Greene, J. (2007). *Mixed methods in social inquiry.* San Francisco, CA: Jossey-Bass.

Johnson, R. B., Onwuegbuzie, A., & Turner, L. A. (2007). Toward a definition of mixed methods research. *Journal of Mixed Methods Research, 1*(2), 1-22.

Leech, N. L., Dellinger, A. B., Brannagan, K. B., & Tanaka, H. (2009). Evaluating mixed research studies: A mixed methods approach. *Journal of Mixed Methods Research, 4*(1), 17-31.

Leech, N. (2010). Interviews with the early developers of mixed methods research. In A. Tashakkori, & C. Teddlie (Eds.), *Sage Handbook of Mixed Methods in Social & Behavioral Research* (2nd ed., pp. 253-274). Thousand Oaks, CA: Sage.

National Institutes of Health. (2012). *National Institutes of Health Office of Behavioral and Social Sciences Research*. Retrieved November, 17, 2012, from Best Practices for Mixed Methods Research in the Health Sciences: http://obssr.od.nih.gov/scientific_areas/methodology/mixed_methods_research/section2.aspx

O'Cathain, A. (2010). Assessing the quality of mixed methods research. In A. Tashakkori, & C. Teddlie (Eds.), *Sage handbook of mixed methods in social & behavioral research* (2nd ed., pp. 511-555). Thousand Oaks, CA: Sage.

O'Cathain, A., Murphy, E., & Nicholl, J. (2008). The quality of mixed methods studies in health services research. *Journal of Health Services Research Policy*, *13*(2), 9298.

Pace, R., Pluye, P., Bartlett, G., Macaulay, A. C., Salsberg, J., Jagosh, J., et al. (2012). Testing the reliability and efficiency of th epilot Mixed Methods Appraisal Tool (MMAT) for systematic mixed studies review. *International Journal of Nursing Studies, 49*, 47-53.

Pedhazur, E. J., & Schmelkin, L. P. (1991). *Measurement, design, and analysis: An integrated approach.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Pluye, P., Gagnon, M.-P., Griffiths, F., & Johnson-Lafleur, J. (2009). A scoring system for appraising mixed methods research, and concomitantly appraising qualitative, quantitative and mixed methods primary studies in Mixed Studies Reviews. *International Journal of Nursing Studies, 46*, 529-546.

Schwandt, T. A. (2007). *The Sage Dictionary of Qualitative Inquiry.* Thousand Oaks, CA: Sage.

Tashakkori, A., & Teddlie, C. (Eds.). (2010). *Mixed methods in social & behavioral research.* Los Angeles, CA: Sage .

Appendix

Interview Protocol

Name: _____
ID No: _____
Interviewer: _____
Date: _____
Phone or in person: _____

Introductory Remarks

I am a doctoral student from Virginia Tech working to complete my dissertation. This study is researching the quality of mixed methods research. Since the conception of mixed methods research, there has been a debate on what components are critical to a mixed methods research study. In addition to the debate on what makes up a mixed methods research study, there has also been a lack of research on how the quality of a mixed methods research study is evaluated. The purpose of this study mixed methods study is to conduct a cross-disciplinary, multi-phase content analysis to compare multiple rubrics for evaluating articles in multiple disciplinary fields.

Our purpose of these interviews is to engage experts in the field of mixed methods research and gain insight on the key elements of a mixed methods study that should be used to create a rubric for the scoring of mixed methods research studies. The interview should last between 45 minutes and one hour. All the questions ask you to share your opinions and experiences.

All the questions that I am going to ask you are about your opinion or personal viewpoint. There are no right or wrong answers to any of these questions. A few of these questions are semi-structured, but most are open-ended.

Question 1: How are you involved in the mixed methods community? (e.g., Do you go to conferences, write scholarly articles, conduct research using MMR, etc.)

Question 2: Why is the quality of mixed methods research important in the ways it is conducted and reported?

Question 2a: On a scale of 1-4, how important is the ability to identify quality in mixed methods research? One being not at all important, two being somewhat unimportant, three being somewhat important, four being necessary. _____

Question 3: What elements of a mixed methods research study are most important? Why?

Question 4: Paradigms are often mentioned as being important in mixed methods research, do you feel that they are important? Why or why not?

Question 5: To what degree is mixing necessary in a quality mixed methods study and how much mixing should occur, please explain your answer?

Chapter 5: Article 2

Evaluating quality in mixed methods research: A guideline for quality

Timothy Burrows

Virginia Tech

Abstract

The diversity of designs and vocabulary found within mixed methods research has led to the need for more guidelines for evaluating the quality of mixed methods research. The purpose of this study was to produce a preliminary set of guidelines to evaluate mixed methods research articles, as well as to confirm the first phase of the study. The participants included nine mixed methods experts. This phase of the study utilized a short survey, including open-ended questions that provided both quantitative and qualitative results. The survey revealed the three most important standards were (1) a clearly stated purpose, (2) clearly stated results, and (3) a clearly stated research design. The results presented here are contrary to the current literature by revealing that a shorter list of standards was necessary to determine quality than previously stated in the literature. The absence of design identification, timing, and sequencing was also dissimilar to the current literature on defining the quality of mixed methods research.

Introduction

Since the mid 1980s, mixed methods research articles have become more prevalent in research journals across a wide variety of disciplines. In 1989, Greene et al. conducted a content analysis of mixed methods research articles published between 1980 and 1988 (Greene, Caracelli, & Graham, 1989). This comprehensive study included a sample of 57 empirical mixed methods evaluations (Greene et al., 1989). In 2006, Allen Bryman conducted another content analysis of mixed methods research and was able to use a sample of 262 articles (Bryman, 2006). One must assume that, in order to for Bryman to develop a larger sample size, the overall population must have also increased. The trend for mixed methods dissertations and grant funding also support this claim. From 1997 to 2008, the number of researcher identified mixed methods dissertations rose from approximately 10 in 1997 to over 700 in 2008 (Plano Clark, 2010). In addition to dissertations, the number of federally funded grants also rose between 1997 and 2007 (Plano Clark, 2010). Currently mixed methods research is not only commonly used in the health field and to create quality assessment, but also in numerous applied research fields (O'Cathain, 2010; Teddlie & Tashakkori, 2009). Even though the overall number of mixed methods research publications and funded grants are increasing at a steady rate, the quality and application of mixed methods research is largely undefined within the literature (Bryman, 2006; Gambrel & Butler, 2011; Greene et al., 1989).

This researcher addresses the issue of a largely undefined definition of quality within mixed methods research by proposing a new set of quality standards that can be used to ensure a quality mixed methods study. The standards and guideline presented in this article are based on the findings of the qualitative portion of a larger mixed methods

study, which was designed to use interview data from experts in the field of mixed methods research. The term "standards" was chosen to reflect the definition provided by Fitzpatrick, Sanders, and Worthen (2011). Standards are criteria that can be used to judge the effectiveness of what is being evaluated; in this case, mixed methods research (Fitzpatrick, Sanders, & Worthen, 2011). Burrows (2013) interviewed 12 experts, selected from the authors of the *Sage Handbook of Mixed Methods in Social and Behavioral Research* (2010) to inductively create a set of standards which were thought to be important when evaluating the quality of mixed methods research. *What is quality? A look at what makes a quality mixed methods study* (Chapter Four) provided evidence that experts in mixed methods research hold diverse views about the evaluation of quality, the use of paradigms, and transparency. The twelve experts interviewed in that study had an array of experience and came from diverse backgrounds, thus providing an in-depth look into views towards evaluating the quality of mixed methods studies, the purpose of mixing, the use and role of paradigms in mixed methods research, and the importance of transparency within a mixed methods study.

Six examples of evaluation guidelines for mixed methods research stand out in current literature pertaining to mixed methods research. These six sources include articles (Leech, Dellinger, Brannagan & Tanaka, 2009; O'Cathain, Murphy, & Nicholl, 2008; Pluye, Gagnon, Griffiths, & Johnson-Lafleur, 2009; Schifferdecker & Reed, 2009), a textbook (Creswell & Plano Clark, 2011), and government grant reporting procedures (National Institutes of Health, 2012). These sources provide a broad picture of the present state of the evaluation of mixed methods research. Sources for evaluating quality are not limited to these six, but the six presented here are found frequently in the mixed methods

research literature at the time this article was written. As more resources are provided for evaluating mixed methods research, researchers will have a larger selection of options to evaluate the quality of their research according to the specific needs presented by the dynamic process that is mixed methods research.

As mixed methods research continues to emerge as the third methodological movement, it is important to enable researchers to evaluate the quality of the research that is presented. Best practices for conducting mixed methods research are important to the growing interest in mixed methods (Klassen, Creswell, Plano Clark, Smith, & Meissner, 2012). Researchers are enabled to evaluate the quality of mixed methods research through the use of guidelines and rubrics. Guidelines are a standard for practice, developed to specify specific criteria when evaluating standards (Fitzpatrick et al., 2011). According to Fitzpatrick et al. (2011), a rubric is similar to a guideline in that it is also used to evaluate standards, but does so using a quantitative measure to assign a value to what is being evaluated. The use of guidelines and rubrics allows mixed methods researchers to use multiple tools to both identify standards, as well as to measure standards.

This article is a report on the quantitative phase of a larger multi-phase mixed methods study to assess experts' views about standards that are appropriate to evaluate the quality of mixed methods research reports. The quantitative phase included a survey to confirm the qualitative analysis conducted in the first-phase of the study. This study has been created from the perspective of a pragmatic paradigm. The knowledge gained in this study was used to produce results that can improve the practice of mixed methods research.

This article presents the quantitative results from the second phase of a multi-phase mixed methods study that used an exploratory research design with a sequential timing and quantitative priority. The integration of methods occurred during this phase, which transformed qualitative data into polytomous variables used to evaluate the quality of mixed methods research publications and reports. In addition to transforming the qualitative data, this researcher also used connectivity to link the qualitative data to the quantitative data in the analysis. The rationale for the mixing in this study was to transform concepts and ideas pertaining to the quality of mixed methods research into a set of standards (rubric) that possessed a quantitative value. In the discussion section, findings from the second phase of the study were considered in light of results from the larger study.

Literature Review

As mixed methods research grows in popularity as a research methodology, quality is becoming a more important focus to ensure accurate results. Currently there are many views about the characteristics of mixed methods research and how they are related to determining the quality of mixed methods studies. As of now, there is not a set of guidelines that have been endorsed by the professional association. The methodological literature suggests that consensus has not been reached that identifies a set of standards that can be used to evaluate the quality of mixed methods research publications. This lack of consensus may be related to the lack of conformity about the key defining qualities of mixed methods research. Several examples are given below that illustrate the differences when defining the key elements for defining mixed methods studies.

Accounting for the quality of mixed methods research is difficult due to the many defining characteristics of mixed methods research, such as mixing and design types. Currently, there are variety of viewpoints about what a mixed methods research study should include and what its purpose should be (Creswell & Plano Clark, 2011; Greene, 2007; Tashakkori & Teddlie, 2010). Although researchers have given several purposes for mixed methods research and guidelines for what makes a mixed methods study, they all lack a way to evaluate the quality of the study. Greene (2007), Creswell and Plano Clark (2011), and Tashakkori and Teddlie (2010) have each laid out their own interpretations of the dimensions that compose a mixed methods research study, but neglect to propose a way to evaluate them.

In her book, *Mixed Methods in Social Inquiry* (2007), Greene proposed seven characteristics of mixed methods research: (1) paradigms, (2) phenomena, (3) methods, (4) status, (5) implementation: independence, (6) implementation: timing, and (7) study. *Paradigms* refers to which paradigms are used and how they are implemented in the study. *Phenomena* refers to the degree in which qualitative and quantitative methods are used to assess the phenomena being studied. *Methods* is how the various selected methods are similar to and different from one another. *Status* represents the weighting of the qualitative and quantitative components of the study. *Implementation: independence* is the degree to which the selected methods are utilized and designed. *Implementation timing* refers to when the qualitative and quantitative methods occur in the study. Finally, the *study* represents the category of research in which the overall study fits. Greene's characteristics are described as applicable, but not as absolutes, meaning that they may all apply to a mixed methods research study, but are not mandatory in its design.

Creswell and Plano Clark (2011) provide another set of characteristics of mixed methods research in their publication, *Designing and Conducting Mixed Methods Research* (Creswell & Plano Clark, 2011). Creswell and Plano Clark offer six characteristics that define a mixed methods research study. The six characteristics are (1) collecting both qualitative and quantitative data, (2) using rigorous methods for data collection and analysis, (3) mixing the data, (4) utilizing a mixed methods design, (5) conducting the study within a paradigm or worldview, and (6) reporting the findings using vocabulary consistent in the present literature (Creswell & Plano Clark, 2011).

Tashakkori and Teddlie (2010) follow Greene (2007) and Creswell and Plano Clark (2011) with the most comprehensive list of characteristics with nine in total. The first characteristic that Tashakkori and Teddlie set forth is methodological eclecticism, or the idea that the best method should be selected for the desired outcomes. The second characteristic is paradigm pluralism, the belief that various paradigms may best serve the researcher in answering the research question(s). "The third characteristic of contemporary (mixed methods research) is an emphasis on diversity at all levels of the research", from the conceptual to the empirical (Tashakkori & Teddlie, 2010, p. 9). The fourth characteristic is the emphasis on a continuum for analyzing the data rather than an either-or approach. The inclusion of both inductive and deductive logic in the same study is the fifth characteristic. The sixth characteristic is using a research question to guide the methods employed in the study. The inclusion of a specified research design is the seventh characteristic. The eighth characteristic is balance and compromise in the mixed methods research community. The final characteristic includes the use of "visual

representations and a common notation system" in the research Tashakkori & Teddlie, 2010, p. 11).

When comparing these approaches to defining the characteristics of mixed methods, several similarities and differences become apparent. Greene (2007), Creswell and Plano Clark (2011), and Tashakkori and Teddlie (2010) are similar in the regards to the expectation about transparency towards the author's paradigm and the collection of both qualitative and quantitative data. Greene (2007) and Creswell and Plano Clark (2011) directly address the inclusion of both qualitative and quantitative data, while Tashakkori and Teddlie (2010) are more open on this topic by including inductive (usually qualitative) and deductive (usually quantitative) methods. These three selections also contain similarities when it comes to the design of the study. Greene (2007), Creswell and Plano Clark (2011), and Tashakkori and Teddlie (2010) discuss the inclusion of a design, but do so on different levels. Greene (2007) refers to the relationship between the use of mixed methods to the design, while Creswell and Plano Clark (2011) and Tashakkori and Teddlie (2010) make references to specified designs. Greene (2007) and Tashakkori and Teddlie (2010) both make references to making sure the method chosen is the best to obtain the desired results. While these three groups or characteristics contain several similarities, they contain many more differences.

The degree of differences among these sets of characteristics that define a mixed methods study is greater than the similarities presented above. This can first and foremost be seen in the number of characteristics listed. Greene (2007) lists seven characteristics, while Creswell and Plano Clark (2011) list six. Tashakkori and Teddlie (2010) provide the largest set of characteristics with nine. Tashakkori and Teddlie (2010) do not

explicitly refer to mixing, while Greene (2007) and Creswell and Plano Clark (2010) do not address the relationship between the research and the mixed methods community nor diversity of the research. Greene (2007) makes an explicit reference for the inclusion of timing, a topic that is neither addressed by Creswell and Plano Clark (2011) or Tashakkori and Teddlie (2010). Greene's (2007) set of standards is also the only one that makes references to the prioritization of the qualitative and quantitative strands. Creswell and Plano Clark's (2011) standards is the only set of the three that includes an emphasis on rigor and the use of accepted terminology when conducting a mixed methods study. Tashakkori and Teddlie differ from Greene (2007) and Creswell and Plano Clark (2011) by addressing the importance of an emphasis on diversity at all levels of the research" (Tashakkori & Teddlie, 2010, p. 9). Tashakkori and Teddlie (2010) are also the only ones to include the use of a research question to guide the mixed methods study. Of the three sets of standards presented here, Tashakkori and Teddlie (2010) are the only mixed methods researchers to also include the use of a visual representation.

In addition to the differences in terms of what is addressed and what is not addressed, these three sets of characteristics also differ by the populations that are discussed within the standards. Greene (2007) and Creswell and Plano Clark (2011) present more of a list that is directed towards the individual researcher and the design of the study. Tashakkori and Teddlie (2010) make a larger connection between the researcher, the study, and the community of mixed methods as a whole. This can be seen in Tashakkori and Teddlie's (2010) inclusion of balance and compromise in the study as it relates to the mixed methods community. By balance and compromise, they mean the common components of the mixed methods study among a group of researches

conducting mixed methods research. Another example of Tashakkori and Teddlie (2010) relating to the mixed methods community is the inclusion of diversity within the study. This difference of approaches complicates the already murky waters of defining mixed methods research.

These similarities and differences in definitions of mixed methods research point to differences in the guidelines each researcher might create to evaluate the quality of a mixed methods research article. Although an overlap does exist between these sets of characteristics, it is not as large as one may think. The literature provides evidence for inconsistency in the current definitions of mixed methods research. Each definition contains aspects that are not included in other definitions found in the literature. This lack of agreement may be a key problem when designing a guideline or a rubric that will be accepted by the mixed methods community on a large scale. This problem may be countered by focusing on the similarities found in the literature instead of the differences.

Methods

This section details how a set of guidelines for evaluating the quality of mixed methods research was created, including the collection and analysis of data. For additional details on methods used prior to this phase of the study, please refer to *What is quality? A look at what makes a quality mixed methods study* (Chapter Four).

Sampling

After the interview was conducted with the 12 initial participants, they were asked if they would be willing to participate in a second phase of the study. The second phase included a short questionnaire (Appendix A) using Qualtrics, a secure server based survey software, to rate the importance of the most predominate categories found relating

to evaluating the quality of mixed methods research. In addition to questions about quality, they were also asked about their methodological background and basic demographic information. Of the 12 original participants, 11 agreed to take part in the questionnaire. Out of the 11 participants, nine (82%) fully completed the questionnaire. One of the participants explained discomfort with rating qualitative data during the interview and did not participate in the questionnaire. One participant did not fully complete the questionnaire and was not included as a full participant in this phase. The reason for the loss of the final participant was not confirmed and may have been a result of fatigue. The nine remaining participants included 7 males and 2 females. Of the nine participants, six supplied comments to the open-ended questions. Race is not addressed in this sample in order to limit any compromise of the participant's identity.

The nine participants represented multiple disciplines from both the United States and the United Kingdom. Two of the participants were faculty at universities located in the United Kingdom, one was a United Kingdom resident working with a United States based firm, and seven were American faculty members at institutions located within the United States. The experience of the participants ranged from five years to over 30 years. The participants represented both qualitative and quantitative paradigms.

This study was approved by the Virginia Tech Institutional Review Board as part of a larger study to ensure that interactions with the participants were conducted in an ethical manner. Consent for participating in this portion of the study was given verbally at the end of the interview protocol in the first phase of this study. Permission to include the participants names in the acknowledgment section of this article was given through the questionnaire.

Data Collection

This article is a description of the process of transforming the qualitative data collected during the first phase of this study into a guideline for evaluating the quality of mixed methods research. From the interviews, a list of codes was inductively produced. These codes were used to create the items for the rubric because of their specificity to evaluation guidelines. Fourteen codes were selected based on their level of occurrence in the expert interviews. A code was included if three or more participants mentioned it. The frequency of codes ranged from 12 (paradigm use) to 3 (reliability).

The participants rated the items based on a six-point Likert scale ranging from extremely important to not at all important. Extremely important was worth six points, very important was worth five points, somewhat important was worth four points, somewhat unimportant was worth three points, very unimportant was worth two points, and not at all important was worth one point. A mean for each of the 14 items was produced to create a ranking of importance from highest to lowest. The four highest and two lowest scored items are discussed in further detail in the results/findings section of this article. Items with a mean score of less than 4.0 were removed from the final guideline due to being interpreted as unimportant. This score was determined based on the total number of participants and an average rating of somewhat important.

In addition to rating the 14 items, the participants were asked if they had any additional comments about the listed standards or additional points not addressed in the questionnaire. The open-ended responses from the questionnaire on additional standards and recommendations are also presented in the results/findings and were taken into

account when producing the final guidelines. Six participants made an additional comment about the 14 items.

Quantitative Analysis

The analysis of this article was done through the rating scale and the open-ended questions of the questionnaire. Support for the evaluation of mixing was provided by the interview data. A matrix of the item scoring is provided (Table 5.1) to show a comparison between items and their levels of importance. The analysis is presented to provide a high level of transparency in the process to create the guideline presented in this study.

Trustworthiness of the qualitative analysis portion of this study was ensured through the use of an audit trail, member checking, peer review, and reflexivity (Patton, 2002). An audit trail made up of procedural memos was used to create an outline of the decision making process of the qualitative procedures. The participants reviewed their interviews for accuracy of the transcripts. Additionally, an additional qualitative expert was used to confirm the coding schemes and categories that were developed during the qualitative phase. Finally reflexive practices were used to acknowledge and minimize personal bias. As a pragmatist, I am supportive of the development of a relatively brief and applicable set of standards.

Instrument validity was ensured through Kane's (2011) criteria for validity. Kane stated that validity is created through the interpretation of the score and the plausibility of the instrument's proposed uses (Kane, 2011). The claim of this instrument, the rubric, is that it can be used to score the quality of mixed methods research articles. The evidence that supports this claim comes from the data (participants being interviewed and their responses to the protocol). The claim is warranted through a series of assumptions which

back the claim. As experts, it is assumed that the themes that arose during the interview process represented the critical parts of a mixed methods study. By identifying the themes that appeared with the greatest frequency, it was assumed that these were the most important themes amongst the participants in regards of quality. By turning these themes into items, it was then assumed that the items reflect the most important issues in judging the quality of mixed methods research. Finally, by scoring the items, the rubric was then made more concise by selecting those items which were scored the most vital by the experts. By following this logical sequence, the rubric's score of the quality of mixed methods research articles was determined to be valid and also potentially generalizable to a larger population.

Limitations

Through the use of standards for creating trustworthy qualitative results and a valid instrument for rating the quality of mixed methods study, the aim of this study was to produce a transparent process that can be replicated for further analysis. To maintain these principles, the limitations of this study are addressed. The largest limitation for this study may be one of the biggest challenges with mixed methods research, definitions of terms. The definition for "study" in this study was a written product used to disseminate information to the public via publication, or presentation. Some members disclosed their confusion on this matter. Some felt that a study was the act of creating research, while others agreed that is defined by the report. This may have altered some responses to the items on the questionnaire. The second limitation was sample size. Although the participants are all experts in the field of mixed methods, their participation may have been motivated by their own interest in the topic of quality. In addition to terminology

and sampling, my own bias as a researcher entrenched in the quality literature may influence the findings of this study. Finally, the design used the same participants from the initial interviews to confirm the initial findings so consistency in results between two phases would be expected. Many attempts were made to limit the level of bias, but due to the human nature of research, can never be completely removed.

It is important to also address that the participants of this study may have been limited by the idea that universal standards are improper due to the diversity found in mixed methods research. Many experts who were asked to participate may have opted out because they did not feel comfortable with agreeing to take part in the development of standardized guidelines for evaluating the quality of mixed methods research. As this topic becomes more of a critical issue, future participation may increase to a larger body of mixed methods researchers.

Findings/Results

Following the qualitative analysis of interviews conducted with 12 experts in the field of mixed methods research, 14 standards emerged for ensuring a high level of quality for mixed methods research. The 14 standards and their scores are presented below in Table 5.1. Table 5.1 shows the frequency of responses, the level of response by the participant, and the mean score for each standard. All items were rated as somewhat to very important, reflecting some consensus among the respondents.

The scoring of the 14 standards along with the additional comments provided by the participants led to the development of a guideline that includes 15 standards to ensure a transparent and quality mixed methods research study. One of the initial standards was removed and two additional standards were added based on the open-ended comments

discussed later in this article. This revision changed the total number of standards presented in the final guideline from 14 to 15. The collection of standards is referred to as a guideline since it specifies specific criteria, or standards, for the purpose of evaluation. The guideline contains the following standards:

- Clearly stated results of the study.
- Clearly stated research purpose statement.
- Clearly stated research design.
- Clearly stated methods of how the study was conducted.
- Clearly stated evidence of how data are mixed/integrated in support of the study's findings.
- Clearly stated research question(s) used to guide the study.
- Clearly stated explanation of how the data are mixed/integrated in the research design.
- Clearly stated evidence for the studies legitimization.
- Clearly stated reason for mixing in the study.
- Clearly stated conclusion that includes the "value added" by mixing/integrating the data.
- Clearly stated philosophical paradigm(s) by the researcher.
- Clearly stated philosophical paradigm(s) by the researcher.
- Clearly stated how the methodology fit the purpose of the study.
- Clearly stated in which phase the mixing/integration occurs.
- The study is replicable.

Quantitative Responses

Of the 14 standards that emerged from the qualitative analysis, the participants arrived at the highest level of consensus for "a clearly stated results of the study" (mean=5.9), "a clearly stated purpose statement" (mean=5.7), "a clearly stated research design" (mean=5.6), and "the overall transparency of the research study" (mean=5.6). These four standards represent the three highest scores given to the standards that emerged from the analysis of the interviews with experts. Each of these standards did not receive any scores in the unimportant side of the scale. "Clearly stated results of the study" received two selections for very important and eight selections for extremely important. Having "a clearly stated purpose statement" had three selections for being very important and six selections for being extremely important. Both standards relating to research design and transparency had one selection for somewhat important, two selections for very important, and six selections for extremely important. Based on the scoring of these items, the quality of a mixed methods research study without these four standards may be significantly degraded. The qualitative responses are presented in Table 5.1.

Table 5.1

Expert Ratings of Mixed Methods Quality Standards

Standard	Number of responses							
	Not at all important	Very unimportant	Somewhat unimportant	Somewhat important	very important	Extremely important	Total Responses	Mean **(Standard Deviation)**
Clearly stated results of the study.	0	0	0	0	1	8	9	5.9 **(0.33)**
Clearly stated purpose statement.	0	0	0	0	3	6	9	5.7 **(0.50)**
Clearly stated research design.	0	0	0	1	2	6	9	5.6 **(0.73)**
Overall transparency of the research study.	0	0	0	1	2	6	9	5.6 **(0.73)**
Clearly stated methods of how the study was conducted.	0	0	0	1	3	5	9	5.4 **(0.73)**
Clearly stated evidence of how data are mixed/ integrated in support of the study.	0	0	0	0	5	4	9	5.4 **(0.53)**
Clearly stated research questions used to guide the study.	0	0	0	2	2	4	8	5.3 **(0.89)**
Clearly stated explanation of how the data are mixed/ integrated in the research design.	0	0	1	0	3	5	9	5.3 **(1.00)**
Clearly stated evidence for the studies robustness/ trustworthiness.	0	0	0	1	4	4	9	5.3 **(0.71)**
Clearly stated purpose for mixing in the study.	0	0	0	3	1	5	9	5.2 **(0.97)**
Clearly stated conclusion that includes the "value added" by mixing/ integrating the data.	0	0	2	0	3	4	9	5.0 **(1.22)**
The study is replicable.	0	0	0	5	2	2	9	4.7 **(0.87)**
Clearly stated review of literature pertaining to mixed methods research.	0	0	2	4	1	2	9	4.3 **(1.12)**
Clearly stated paradigm(s) by the researcher.	0	0	1	5	2	1	9	4.1 **(0.84)**

Note: N = 53

The two standards with the lowest frequency of importance in determining a quality mixed methods study are "a clearly stated review of the literature pertaining to mixed methods research by the researcher" (mean=4.3) and the study are a "clearly stated paradigm(s) by the researcher" (mean=4.1) (Table 5.1). The standard about the review of the literature received two selections for being somewhat unimportant, four selections for

somewhat important, one selection for very important, and two selections for being extremely important. The standard of having a "clearly stated paradigm" received one selection of somewhat unimportant, five selections for somewhat important, three selections for very important, and one selection for extremely important. The scores presented in Table 5.1 for this item reflect the data presented in Chapter Five. Although these items were the lowest scoring, all the items scored above the minimum cut score of 4.0.

Responses to Open-Ended Questions

In addition to the scoring of the standards, supplementary comments were made about these items by the participants for evaluating the quality of mixed methods research. The participants were asked two open-ended questions about the standards that were presented to them. The first open-ended question was about a quality criteria that was not listed that they believe should have been. The second open-ended question was if there were any additional comments about the standards presented in the questionnaire. Of the nine participants, six of them responded to these items. These comments added additional information that affected the analysis by introducing two new standards that were not openly discussed in the interview phase of this study. Two new standards and the opinion that transparency was not a standard, but a result that is achieved through the other standards, are discussed below.

The comments provided for the first open-ended question were divided into three categories. These categories included general comments, standard suggestions, and terminology suggestions. The one general comment was "reflexivity is of the utmost importance." The two comments about standard suggestions are, the

"methods/methodology used should fit the purpose of the study", and a space should be provided to denote "which phases of the study were mixed methods introduced and used". The terminology comment was for the use of “the term "legitimization" rather that trustworthiness or validity for mixed research." Both the comments about additional standards align with the other participant’s viewpoints and the current literature, so therefore are also included in the final guidelines listed below. As a way to help distinguish mixed methods as its own paradigm, I have also elected to use the term "legitimization" to encompass both trustworthiness and validity.

The responses to the second open-ended question raised questions about the overlap of standards and the meaning of a term. The comment about overlap addressed two cases; the first case being "evidence for the studies robustness/trustworthiness" and "overall transparency", and the second case being, "clearly stated purpose" and "clearly stated purpose for mixing in the study." A comment was also raised about possible multiple meanings for "overall transparency of the research study." Due to the number of standards presented here and the complexity of mixed methods research, care was taken to clarify the standards in the final guideline. The final comment addressed the difficult task of rating the proposed quality standards because "so many of these are very important when evaluating a mixed methods research study."

Discussion

This article presents the key results of the second phase of this multi-phase mixed methods study. The quantitative results confirm that the analysis of the qualitative data was successful in identifying a set of standards that the mixed methods experts interviewed considered to be important. The three most important standards were "clearly

stated results" (mean=5.9), "clearly stated purpose" (mean=5.7), and "clearly stated research design" (mean=5.6). The three standards that were rated the least important were "the study is replicable" (mean=4.7), "clearly stated review of literature pertaining to mixed methods research" (mean=4.3), and "clearly stated paradigm(s) by the researcher" (mean=4.1). Additional conclusions were reached about the lack of emphasis on designs, timing, and priority.

The quantitative phase discussed in this article reveals the expert participants of this study validated the qualitative phase of this study by rating all the standards as "somewhat important" or higher. This shows an agreement amongst the participants by showing a common level of acceptance of the categories developed in the qualitative phase, regardless of whether they were discussed in-depth in their individual interviews. Eleven of the 14 items were rated as important or very important. This confirms there are areas of agreement for common standards when evaluating the quality of mixed methods research. Even if the results of this article cannot be generalized across a larger population, it is still important to note this second phase was successful in confirming the findings from the qualitative phase of this study.

The lack of variability in the ranking of the importance of the standards listed in the survey instrument is somewhat in contrast to what the participants said during the interviews. Several of the participants spoke about the problems with creating a set of standards or a rubric due to the lack of agreement towards what is important in regards of quality in mixed methods research. Nine of the 12 participants expressed reservations about creating a rubric because of its limiting and constraining qualities. Participants discussed a fear that any restrictions may limit the growth of mixed methods research.

The guideline presented in this article was compared to several common guidelines currently used to evaluate the quality of mixed methods research to illustrate these similarities and differences. The guideline presented in this article was compared with those presented by Collins, Onwuegbuzie, and Johnson (2012), Creswell and Plano Clark (2011), Leech, et al. (2009), Leech, Onwuegbuzie, and Combs (2011), NIH (2012), O'Cathain (2010), O'Cathain et al. (2008), Pluye at al. (2009), Schifferdecker and Reed (2008), and Teddlie and Tashakkori (2009).

Similarities were found between several previously published guidelines. Similarities between Collins at al. (2012) and this guideline include the use philosophical clarity to make meta-inferences. Similarities with Creswell and Plano Clark are seen in rigorous data collection and analysis, integration to provide a better understanding, and using a defined research design. Additional similarities can be found in Leech et al. (2009). These similarities include design suitability, design quality, and legitimization. Further similarities are found in Leech et al. (2011) in regards to justified and clear methods, a clear design, the use of appropriate research questions and research purpose, legitimization, and the appropriateness of integration. The similarities with NIH (2012) include the similarities with other guidelines that are addressed within this paragraph. Many similarities are also found in O'Cathain at al. (2008); these include a justification for using mixed methods, a clear design, a clear methods section, a description of where integration occurs, and a statement about what is gained through mixing data. Final similarities presented here align with Teddlie and Tashakkori's (2009) design suitability, design fidelity, design consistency, analytic consistency, interpretive agreement, integrative efficacy, and interpretive correspondence. Although the names do not match,

the overall concepts represented in these ideas presented here embody similar goals towards defining quality in mixed methods research.

Along with several similarities, differences also were apparent when comparing the guideline presented here and those previously published. The most notable difference between the guidelines presented in this article and many of those being currently used, such as the NIH standards, was the absence of design, timing, and sequencing. The use of specific designs was not seen as important. The primary statement that accompanied the use of designs was its purpose. The important factor was if the design appropriately allowed the researcher to answer the research questions, not the specific design itself. Timing and the use of sequence was absent from the interviews with the participants, which may be due to the emphasis on these concepts by competing writers in the mixed methods literature. A heavy importance may be placed on concepts that are commonly referred to in the literature opposed to reflecting on the overall importance of these concepts as they directly relate to quality. This is not to say that the disclosure of designs, timing, and sequence does not add to quality, but was not determined to be necessary by the participants of this study. An additional point for why design, timing, and sequence are found in several guidelines may be that several of the authors of current guidelines also either developed or were closely linked to the developers of these concepts. The inclusion of these concepts may be an artifact of personal bias.

Another difference excludes the complex nature found in the works presented by Collins et al. (2012), Leech et al. (2010), and O'Cathain (2010). O'Cathain stated that it is difficult and time consuming to apply too many standards to a real-life mixed methods study (O'Cathain, 2010). The guidelines presented here address more generalizable

categories, such as clear methods, which includes the data collection and research design. The participants of this study provided information to create 15 standards, which is considerably shorter than O'Cathain's 30 pages of criteria/standards found in the 2nd edition of the *Sage Handbook of Mixed Methods in Social & Behavioral Research* (2010). The guideline here provides evidence that relatively few standards are required to create a quality mixed methods study.

In addition to the similarities amongst the experts interviewed in this study, more similarities than differences were found between the guideline presented here and those found in the literature, revealing that common standards exist in the evaluation of quality in mixed methods research. By validating the qualitative phase of this study and confirming many similarities between previously developed quality standards, evidence was given to further develop the idea of standardized, but flexible criteria for evaluating quality of mixed methods studies.

Conclusion

The purpose for this guideline is to present a high degree of transparency to the reader. All of the standards presented are individual pieces that create a transparent study. The more standards that are present in a mixed methods study, the higher the level of transparency and thus the higher the level of quality. Each of the 15 standards is to be identified as present or absent.

From the on-set of this study, I met a considerable amount of resistance from several members of the mixed methods community. This resistance came in the form of a reluctance to develop or use a set of common standards for evaluating quality in mixed methods research. Researchers focused on the high degree of differences found within

mixed methods research projects. These differences were also compounded by the multitude of designs, paradigms, and methodologies used. It is difficult to determine if the sentiments voiced by some of the participants reflect a wider sentiment among the mixed methods community or if they are idiosyncratic to members of this sample and thus a reflection of respondent bias. It is not surprising that this group of authors would hold strong views about the nature of mixed methods research and how it should be evaluated.

I believe that the primary force behind this unwillingness to accept a set of standards results from particular paradigms. Many qualitative researchers do not believe that qualitative research can or should be quantified, which puts them in direct opposition of using rubrics to measure a qualitative construct such as quality. In addition to this, the idea of standardization does not appeal to many mixed methods researchers due to a fear of losing the ability to freely create new concepts for carrying out mixed methods research. Standards were said to stunt the growth of the methodology during a messy period of development. I believe that the similarities between many guidelines and a more generic set of standards are acceptable at this time without limiting further development in mixed methods research. Standards and guidelines are not immovable objects. If harmful side effects are identified from a set of standards, the response is simple; simply change or remove the conflicting standard(s). Often our own paradigms lock us into particular mindsets and we must be able to challenge ourselves and take actions that may be uncomfortable.

Many of the participants in this study were uncomfortable with the ideas of standards, rubrics, and guidelines. As the participants voiced their reasons for

disagreement with these concepts, they often did not make any notable distinction between the three. Standards are developed to act as a common ground for evaluating something. Rubrics are designed to assign a quantitative value to a set of standards to ensure they are being met. A guideline is a set of rules to guide an individual to ensure that all standards are met in the process of evaluating a process. It appeared that the difference between these three ideas was not as important as the idea that they could each use their own processes in the evaluation of quality. Of these three concepts, the idea of using a rubric was met with the least amount of approval, again, most likely due to the large differences in paradigms.

Contrary to the literature and the opinions of many of the participants in this study, I believe that mixed methods researchers need a more structured method of evaluating quality of mixed methods research. Confusion is growing as the use of mixed methods research is growing. By not having a solid set of guidelines in place, as a mixed methods community grows larger, so will its lack of adherence to the initial principles. I do believe that this list is sufficient to accurately evaluate the quality of mixed methods research publications. This list offers the principle fundamentals of mixed methods research and is designed for expansion. It is permissible to promote more rigor, but not to reduce it.

Although the use of rubrics to evaluate the quality of mixed methods studies is not readily accepted by many experts, it does present an important issue. Currently, funding agencies are rating mixed methods studies to choose the best studies for funding. This is where practice may be ahead of the experts. If the experts do not choose to engage in the creation of rubrics to evaluate the quality of mixed methods research, who will do it? We

must ask ourselves, do we want the most qualified people to wade through these murky waters, or do we want those who may not be as skilled to make the big splash? In addition to funding agencies, experts are already ranking quality when reviewing journal articles for acceptance in peer reviewed journals.

The results of this study are based upon the consensus of the participants towards what defines quality in a mixed methods study. The consensus of standards found here may be a result of response bias or the sample who agreed to respond to participate in the study. Future research is needed to identify a larger sample to confirm if this consensus holds true for a much larger population of mixed methods researchers. The findings of this article are in direct contrast to O'Cathain's (2010) exhaustive list of what seems to be hundreds of criteria. To successfully begin a dialogue about standardizing quality, we should start with a large focus and narrow it as necessary over time.

References

Bryman, A. (2006). Integrating quantitative and qualitative research: How is it done? *Qualitative Research, 6*(1), 97-113.

Bryman, A. (2007). Barriers to integrating quantitative and qualitative research. *Journal of Mixed Methods Research, 1*(1), 8-22.

Caracelli, V., & Greene, J. (1993). Data analysis strategies for mixed-method evaluation designs. *Educational Evaluation and Policy Analysis, 15*(2), 195-207.

Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research.* Thousand Oaks, CA: Sage.

Creswell, J. W., Klassen, A. C., Plano Clark, V. L., & Smith, K. C. (2011, August). *Best practices for mixed methods research in the health sciences*. Retrieved December 30, 2011, from Office of Behavioral and Social Science Research: http://obssr.od.nih.gov/mixed_methods_research

Fitzpatrick, J. L., Sanders, J. R., & Worthen, B. R. (2011). *Program evaluation: Alternative approaches and practical guidelines.* Upper Saddle River, NJ: Pearson.

Gambrel, L. E., & Butler VI, J. L. (2011). Mixed methods research in marriage and family therapy: A content analysis. *Journal of Marital & Family Therapy*.

Greene, J. (2007). *Mixed methods in social inquiry.* San Francisco, CA: Jossey-Bass.

Greene, J. C., Caracelli, V. J., & Graham, W. F. (1989). Toward a conceptual framework for mixed-method evaluation design. *Educational Evaluation and Policy Analysis, 11*(3), 255-274.

Kane, M. (2011). Validating score interpretations and uses: Messick lecture, language testing research colloquim, Cambridge, April 2010. *Language Testing, 29*(1), 3-17.

Klassen, Creswell, Plano Clark, Smith, & Meissner. (2012). Best practices in mixed methods for quality of life research. *In Quality of Life Research*, 21, 377-380.

Leech, N. L., Dellinger, A. B., Brannagan, K. B., & Tanaka, H. (2009). Evaluating mixed methods research studies: A mixed methods approach. *Journal of Mixed Methods Research, 4*(1), 17-31.

Morse, J. M. (1991). Approaches to qualitative-quantitative methodological triangulation. *Nursing Research, 40*(2), 120-123.

Morse, J. M. (2003). Principles of mixed methods and multimethod research design. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 220-235). Thousand Oaks, CA: Sage.

O'Cathain, A. (2010). Assessing the quality of mixed methods research. In A. Tashakkori, & C. Teddlie (Eds.), *Sage handbook of mixed methods in social & behavioral research* (2nd ed., pp. 511-555). Los Angeles: Sage.

O'Cathain, A., Murphy, E., & Nicholl, J. (2008). The quality of mixed methods studies in health services research. *Journal of Health Services Research Policy, 13*(2), 92-98.

Patton, M. Q. (1990). *Qualitative research & evaluation methods* (2nd ed.). Newbury Park, CA: Sage.

Patton, M. Q. (2002). *Qualitative research & evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage.

Plano Clark, V. (2010). The adoption and practice of mixed methods: US trends in federally funded health-related research. *Qualitative Inquiry, 16*(6), 428-440.

Plano Clark, V. L., Huddleston-Casas, C. A., Churchill, S. L., O'Neil Green, D., & Garrett, A. L. (2008). Mixed methods approaches in family science research. *Journal of Family Issues, 29*(11), 1543-1566.

Pluye, P., Gagnon, M.-P., Griffiths, F., & Johnson-Lafleur, J. (2009). A scoring system for appraising mixed methods research, and concomitantly appraising qualitative, quantitative and mixed methods primary studies in Mixed Studies Reviews. *International Journal of Nursing Studies, 46*, 529-546.

Schifferdecker, K. E., & Reed, V. A. (2009). Using mixed methods research in medical education: Basic guidelines for researchers. *Medical Education, 43*, 637-644.

Small, M. L. (2011, April 15). *How to conduct a mixed mehtods study: Recent trends in a rapidly growing literature.* Retrieved November 26, 2011, from The Annual Review of Sociology: http://www.soc.annualreviews.org

Teddlie, C., & Tashakkori, A. (2009). *Foundations of mixed methods research: Integrating quantitative and qualitative approaches in the social and behavioral sciences.* Thousand Oaks, CA: Sage.

Tashakkori, A., & Teddlie, C. (Eds.). (2010). *Mixed methods in social & behavioral research.* Los Angeles, CA: Sage.

Appendix

Mixed Methods Quality

Thank you for your assistance in determining the importance of various components that have emerged from interviews as being important to evaluating the quality of a mixed methods study. Fourteen criteria for quality have emerged from the qualitative data. Each of the 14 items can be rated on a 6-point Likert scale. In addition to rating the criteria, there are three questions pertaining to demographics and consent. Your assistance will help create a set of guidelines to help increase the overall quality of mixed methods research across disciplines. If you have any questions, please contact me at tburrows@vt.edu.

The following four items of this questionnaire pertain to criteria that are important to evaluating the overall quality of mixed methods research.

Please rate the level of importance of the following criteria listed below. The following criteria were most commonly referred to during the interview process. Some of the criteria are specific to mixed methods research, while others are pertinent to all research methodologies.

Questionnaire

Standard	Response					
	Not at all important	Very unimportant	Somewhat unimportant	Somewhat important	very important	Extremely important
Clearly stated paradigm(s) by the researcher.						
Clearly stated research design.						
Clearly stated purpose statement.						
Clearly stated results of the study.						
Clearly stated research questions used to guide the study.						
Clearly stated explanation of how the data are mixed/ integrated in the research design.						

Questionnaire cont.

Standard	Response					
	Not at all important	Very unimportant	Somewhat unimportant	Somewhat important	very important	Extremely important
Clearly stated review of literature pertaining to mixed methods research.						
Clearly stated conclusion that includes the "value added" by mixing/ integrating the data.						
Clearly stated evidence for the studies robustness/ trustworthiness.						
Clearly stated evidence of how data are mixed/ integrated in support of the study.						
Overall transparency of the research study.						
The study is replicable.						

Is there a quality criteria that is not listed that you believe should be? If so, please explain.

Please identify your top two most important criteria for evaluating a mixed methods study (please choose only two). If you do not wish to do so, please skip this question.

- ❑ Clearly stated paradigm(s) by the researcher.
- ❑ Clearly stated research design.
- ❑ Clearly stated purpose statement.
- ❑ Clearly stated results of the study.
- ❑ Clearly stated research questions used to guide the study.
- ❑ Clearly stated explanation of how the data are mixed/ integrated in the research design.
- ❑ Clearly stated purpose for mixing in the study.
- ❑ Clearly stated methods of how the study was conducted.
- ❑ Clearly stated review of literature pertaining to mixed methods research.
- ❑ Clearly stated conclusion that includes the "value added" by mixing/ integrating the data.
- ❑ Clearly stated evidence for the studies robustness/ trustworthiness.
- ❑ Clearly stated evidence of how data are mixed/ integrated in support of the study.

- ❑ Overall transparency of the research study.
- ❑ The study is replicable.

If you have any additional comments about the criteria listed above or additional points not addressed here, please specify below.

The following three items pertain to demographics and consent. Your name is important for record keeping purposes and will not be released in any way without your permission.

What is your name?

What methodology were you primarily trained in?

- ❍ Qualitative
- ❍ Quantitative
- ❍ Both qualitative and quantitative
- ❍ Other _____

Do I have your permission to include your name in the acknowledgments of published and presented materials for the purpose of giving recognition and additional credibility to the study? Your name will only appear as a list of interviewed individuals. Your name will not be used in the study, publications, and/or presentations except to appear in a list of the names of the people I interviewed. No quotes will be attributed to you by name to ensure the confidentiality of your responses from the interview and this questionnaire.

- ❍ Yes
- ❍ No

Chapter 6: Conclusion

Introduction

The purpose of this mixed methods study was to conduct a cross-disciplinary, multi-phase content analysis to create a preliminary rubric to evaluate mixed methods research articles in multiple disciplinary fields. The four research questions that guided this study were:

1. What are the common evaluation criteria found in the contemporary methodological literature pertaining to the design of mixed methods research? (Qualitative & Quantitative)
2. What evaluation criteria do experts in the field of mixed methods research perceive as the most important when distinguishing top-quality research in mixed methods? (Qualitative)
3. What differences are there in the outcome of the rubric for evaluating mixed methods research identified from the literature compared to those advocated most uniformly by a panel of mixed methods research experts? (Qualitative & Quantitative)
4. What are disciplinary differences between the use of mixed methods and views about evaluating it, including the role of paradigms in mixed methods research? (Qualitative)

This concluding chapter provides a summary of the results for each research question presented in this study, as well as the results for additional analyses conducted in support of the above research questions that were not presented in previous chapters. The data presented in this chapter are not intended to be published due to the small sample sizes and lack of practical significance. In addition to results, this conclusion addresses the study's contribution to the literature, implications for practice, and possibilities for future research.

Key Results and Conclusions

This section details the key results and conclusion for each of the four research questions presented in this study. With the exception of research question four, a more in-

depth discussion of result was presented for each question in Chapters Four and Five. To provide a reference for the criteria discussed in the following sections, a copy of the final instrument, now named the Burrows Rubric for Evaluating Mixed Methods (BREMM), is given below. The results found in the section entitled Criteria Ranking, found in this chapter, contain slightly different criteria than those presented in the BREMM (Figure 6.1). This difference was due to the additional data from the follow-up survey taken by the respondents of this study. Additional information about the steps taken to refine the criteria can be found in Chapter Five. The BREMM uses the term standards instead of criteria because of its evaluative purpose. This term was selected based on the definition provided by Fitzpatrick et al. (2011), in which a standard is a criterion that is used for the purpose of evaluation.

Figure 6.1 Burrows Rubric for Evaluating Mixed Methods (BREMM)

The purpose for this rubric is to allow the reader of a mixed methods study to evaluate and rank the quality of a mixed methods study. All of the standards presented below are individual pieces of a transparent study. The more standards that are present, the higher the level of transparency, and thus the higher the level of quality. The higher the score that a study receives, the higher the level of quality. The maximum score is a 75.

	Response					
Standard	Strongly Disagree (0)	Disagree (1)	Somewhat Disagree (2)	Somewhat Agree (3)	Agree (4)	Strongly Agree (5)
The results of the study are clearly stated.						
The study's purpose statement is clearly stated.						
The research design is clearly stated.						
The methods of how the study was conducted are clearly stated.						

BREMM Cont.

Standard	Response					
	Strongly Disagree (0)	Disagree (1)	Somewhat Disagree (2)	Somewhat Agree (3)	Agree (4)	Strongly Agree (5)
Evidence of how data are mixed/integrated in support of the study's findings are clearly stated.						
The research question(s) used to guide the study are clearly stated.						
An explanation of how the data are mixed/integrated in the research design is clearly stated.						
Evidence for the study's legitimization is clearly stated.						
The reason for mixing in the study is clearly stated.						
There is a clearly stated conclusion that includes the "value added" by mixing/integrating the data.						
The researcher's philosophical paradigms are clearly stated.						
The review of literature pertaining to mixed methods research is clearly stated.						
How the methodology fits the purpose of the study is clearly stated.						
The phase in which the mixing/integration occurs is clearly stated.						
The study is replicable.						

Figure 6.1. The finalized rubric for evaluating the quality of mixed methods research studies.

Research Question One

Research question one was focused on "what are the common evaluation criteria found in the contemporary methodological literature pertaining to the design of mixed methods research"? Through an analysis of the literature presented in the literature review, several common criteria surfaced pertaining to the design of mixed methods

research. Common evaluation criteria included the inclusion of both qualitative and quantitative methods, a value added component from mixing, a method and purpose for the integration of data, and the use of appropriate methods to conduct the study.

Two principal conclusions were drawn from this result. The first conclusion is that some commonalities exist in the quality criteria identified in the literature. This gives support for the idea that a standard can be achieved to evaluate the quality of mixed methods research. The second conclusion is the focus on quality in the literature shows that it is an important topic that needs to be addressed with further research.

Research Question Two

Several key results and conclusions were found for research question two, "what evaluation criteria do experts in the field of mixed methods research perceive as the most important when distinguishing top-quality research in mixed methods"? Research question two was answered in Chapter Five, which provided a quantitative analysis of the qualitative data presented in Chapter Four. Results from the survey (n=9) indicate that respondents rated the most important criteria as "clearly stated results" (mean=5.9, standard deviation=0.31), a "clearly stated research design" (mean=5.7, standard deviation=0.48), a "clearly stated research design" (mean=5.6, standard deviation=0.48), and "the overall transparency of the study" (mean=5.6, standard deviation=0.48). Eleven of the 14 criteria had a mean score of five (very important) or greater on a six-point Likert scale. All of the 14 criteria were rated as at least somewhat important (mean>=4). Although the participants rated the 12 criteria as important, resistance came in the form of a reluctance to develop or use one set of standards with little or no flexibility for evaluating the quality of mixed methods research.

Three primary conclusions emerged from the results for research question two. The first conclusion is the idea of standardization by the creation of a single rubric to be applied to all mixed methods research publications does not appeal to many mixed methods researchers due to a fear of losing the ability to freely create new concepts for carrying out mixed methods research. The second conclusion is that even with this lack of appeal, participants agreed about the importance on 11 of the 14 criteria at the very important level, while all 14 criteria were rated as at least somewhat important. While the idea of standards did not appeal to several of the participants, some of the participants also voiced strong objections to the ideas of rubrics, guidelines, and standards all together.

Research Question Three

Three key findings were developed from the third research question, "what differences are there in the outcome of the BREMM for evaluating mixed methods research identified from the literature compared to those advocated most uniformly by a panel of mixed methods research experts"? These findings were developed during the qualitative analysis of the interview data and were presented in-depth in Chapter Four. In a surprising departure from the literature, participants did not discuss the need to identify the timing, priority, or sequence of a mixed methods study. In addition to the previous elements, the adherence to a specified set of mixed methods research designs identified in the literature was not identified as an important element of evaluating quality. Finally, the participants of this study largely agreed on 15 criteria. This list of criteria is considerably shorter than the long list of criteria O'Cathain identified in a chapter in the second edition of the *Sage Handbook of Mixed Methods in Social & Behavioral Research* (2010).

Several conclusions were drawn from research question three. Further evidence was supplied to further develop the idea of standardized, but flexible criteria for evaluating the quality of mixed methods studies. To support flexibility, participants endorsed the idea that multiple criteria are needed to address the complexity of different types of mixed methods studies. In addition to the stress placed by participants on flexibility, differences were found in opinions expressed about the appropriateness of standards and rubrics. Differences found in the viewpoints expressed about the appropriateness of standards/rubrics may be driven by personal research interests as well as personal philosophical paradigms. The differences found between existing guidelines and rubrics are fewer than their similarities, which again, supports the possibility of moving forward with a standardized measure of quality.

Research Question Four

Research question four asked, "what are disciplinary differences between the use of mixed methods and views about evaluating it, including the role of paradigms in mixed methods research"? The results of this research question are presented in detail in the disciplinary differences section below. Means were used for topical information and not to construct inferences.

The first result was that each group (qualitative group, quantitative group, mixed methods group, social science group, and research methodology/statistics group) rated each criteria with a mean score of four or greater, with the exception of the quantitative group. The quantitative group rated the "inclusion of a clearly stated literature review" with a mean score of 3.5 and a standard deviation of 0.58.

The second result reveled that 10 of the 11 participants (91%) agreed that paradigms serve an important role (minimum score of four out of six) in providing a level of quality. Both the qualitative and quantitative groups had a mean score of 4.0. The qualitative group had a standard deviation of 0.00 and the quantitative group had a standard deviation of 1.41. The mixed methods group had a mean score of 4.8 and a standard deviation of 0.84.

The third result revealed there are differences when ranking each criterion in order of importance by group. The qualitative group ranked the three most important criteria as 2 (mean= 6, standard deviation=0.00), 7 (mean= 6, standard deviation=0.00), and 8 (mean= 6, standard deviation=0.00). The quantitative group ranked the three most important criteria as 3 (mean= 6, standard deviation=0.00), 4 (mean= 6, standard deviation=0.00), and 12 (mean = 5.5, standard deviation=0.71). The mixed methods group ranked the three most important criteria as 4 (mean= 6, standard deviation=0.00), 2 (mean= 5.8, standard deviation=0.45), and 3 (mean = 5.6, standard deviation=0.55). All disciplines combined ranked the three most important criteria as 4 (mean= 5.9, standard deviation=0.31), 3 (mean= 5.7, standard deviation=0.48), and 2 (mean= 5.6, standard deviation=0.48).

The final result of research question four also exposed differences when ranking each criterion in order of importance by discipline. The research methodology/statistics group ranked the three most important criteria as 4 (mean= 5.8, standard deviation=0.50), 3 (mean= 5.5, standard deviation=0.58), and criteria 2, 8, 11, and 12 (tied with a mean=5.3, standard deviation=0.96). The social science group ranked the three most important criteria as 4, 6, and 13 (tied with a mean=6, standard deviation=0.00). The

methodology/statistics group and the social science group both list criteria 1 (mean=4.25/4.5, standard deviation=0.57/1.00) and 9 (mean=3.5/5, standard deviation=1.25/0.55) as the two least important.

Four principal conclusions were developed from research question four. The first conclusion was the quantitative group was the only group to rate the criterion, a "clearly stated review of literature pertaining to mixed methods research" as unimportant, or having a mean score of four or less. There was a general agreement about the importance of a set of criteria. This is evidence that there is an overall acceptance for a set of criteria for evaluating the quality of mixed methods research. Secondly, for practical purposes, there are no significant differences between groups. Thirdly, according to the rankings of mean scores, each of the five tested groups placed different criteria at different levels of importance. Although there is an overall agreement among the standards between groups, there is a lack of agreement on which criteria are most important.

Finally, the 91% agreement (10 of 11 participants) on the importance of the role of paradigms supports the qualitative findings that the self-identification of paradigms is important to the quality of a mixed methods study. The 91% of agreement on the importance of paradigms directly compares to the 92% (11 of 12 participants) of agreement of participants about the importance of paradigms from the interviews. The high level of agreement supports the belief that the inclusion of paradigms creates a higher level of transparency, which had an overall mean score of 5.6 and a standard deviation of 0.48.

Data Collection and Analysis Not Reported in Chapters 4 or 5

The analysis from the first part of the fourth research question was not answered in either Chapters Four or Five, but is answered in this section. Additional information about the pilot study discussed in Chapter Three is also detailed in this section. These two components were not covered in the two articles for two reasons. The first part of question four was not included in Chapter Four because of space limitations when considering journal submission guidelines. The pilot study information was not included in either chapter because it was not directly related to a research question. This information is provided following the discussion of disciplinary differences in this chapter.

Disciplinary Differences

This section refers to the question, "what are disciplinary differences between the use of mixed methods and views about evaluating it"? The answer for this question was surprisingly simple, in that there was essentially none of any practical significance. For purposes of this question, discipline was defined in two ways. The first is by the methodological discipline in which the participants practice. The second definition for discipline was by disciplinary groupings. These were chosen over the standard definition of discipline to reduce the risk of revealing the identity of the participants. According to the first definition, the three disciplines include qualitative, quantitative, and mixed methods. According to the second definition, the two disciplines are social science and research methods/statistics.

The approach taken to conducting a mixed methods study is not based on methodology, but instead driven by the research question and the purpose of the study.

This approach may provide an explanation for why disciplinary differences may not have been significant. During the interviews, participants continually emphasized that the approach taken to conduct mixed methods studies in driven by the research question(s) and the purpose of the study. The approach taken to the mixed methods study changes based on these two factors. This leads to some circumstances when the research question may cause quantitative researchers to conduct primarily qualitative studies and vice-versa. This is perhaps best seen in a response given by Participant 2. Participant 2, a primarily quantitative researcher stated,

> Whether it's pure qualitative or pure quantitative, it has got to be an appropriate match with the kind of problem or question it's asking, and it's got to be faithfully implemented, and it's got to withstand the scrutiny

According to the respondents, regardless of methodological preferences, the methods employed match the research question and fit the purpose. This means that mixed methods researchers have to be flexible in their choice of methods. Participant 6, a primarily qualitative researcher echoed these same views by stating,

> A lot of people will sometimes do research by recipe, they're doing it because they feel like they're supposed to or they're doing it because they want to show that they can do this kind of thing. To me, that's worthless; what they need to do is…they use the methods that support their purpose and as long as the methods and methodologies are true to the purpose of the study and the purpose of the study is valid, then to me that's where it all matters.

Participant 6 also reiterated the qualitative theme that it is the purpose that drives the methods used in a study, nothing else. These are just two examples of the opinion of what methods should be used when conducting mixed methods research. It is not the researcher's principle discipline, but the purpose and the research questions.

Criteria ranking.

The discipline groups were operationalized in two ways. The methodological group (n=8) was defined through a self-identifying item on the follow-up survey. The question asked the participants to choose which methodology they most closely related to. The choices were qualitative, quantitative, or both qualitative and quantitative. The academic group (n=9) was operationalized by the participant's PhD degree given in the second edition of the *Handbook of Mixed methods in social & behavioral research* (2010). The paragraphs below discuss the differences in the ranking of the importance of the criteria. The criteria with the highest means were the most important. The difference in the sample size between the two types of groups was due to one participant answering all the questions of the survey with the exception of what primary methodology he or she identified with. This led to one additional respondent in the analyses that included the social science and research methodology/statistics groups.

There was only one criterion that was among the top five ranked by all three methodological groups. The inclusion of a "clearly stated purpose statement" was included in each group's top five most important criteria. The highest ranked criterion for the qualitative participant was a "clearly stated research design" with a maximum score of six. The qualitative participant also scored "purpose for mixing", "methods", and "transparency" with a score of six. The fifth most important criterion for the qualitative

participant was the "purpose statement" with a score of five. The highest ranked criterion for the quantitative group (n=2) was a “clearly stated purpose statement” (mean=6.0, standard deviation=0.00). Tied for most important criterion for the quantitative group was "results of the study", also with a mean of 6.0 and a standard deviation of 0.00. The quantitative group ranked "evidence of mixing" and "transparency" as tied for second (mean=5.5, standard deviation=.071). The fifth most important criterion for the quantitative group was the "research question" (mean=5.0, standard deviation=0.00). The highest ranked criterion for the mixed methods group (n=5) was “clearly stated results of the study” (means=6.0, standard deviation=0.00). The second most important criterion was the "research design" (mean=5.8, standard deviation=0.45). The third most important criterion was a tie between the "purpose statement" and "explanation of mixing", both with a mean of 5.6 and a standard deviation of 0.55. The fifth most important criterion for the mixed methods group was the "methods" (mean=5.4, standard deviation=0.70).

The analysis addressing the research question about disciplinary differences was also done by comparing the mean scores on each criterion by academic discipline. Like the methodology group, the academic group also provided a unique list of top-five most important criteria for evaluating the quality of mixed methods research. Although the rankings were different, both the social science (n=5) and research methodology/statistics (n=4) groups rated “clearly stated results of the study” as the most important criterion. (means=6.0/5.8, standard deviation=0.00/0.50). The social science group rated the second most important criteria as "explanation of mixing" and "transparency", both with a mean score of 6.0 and a standard deviation of zero. Tied for third, the social science group rated "research design" and "purpose statement" with a mean of 5.8 and a standard

deviation of 0.45. The research methodologist/statistics group rated "purpose statement" as second most important with a mean score of 5.5 and a standard deviation of 0.58. "Research design" was ranked third with a mean score of 5.3 and a standard deviation of 0.96. The "methods" and "robustness and trustworthiness" were tied for fourth most important criteria with a mean of 5.3 and a standard deviation of 0.96.

While interesting, the results about the ranking of the 14 criteria by methodological and academic background are not meaningful for practical purposes. One reason for this is that respondents ranked all the criteria as important. This suggests that in future uses of the rubric, the criteria presented in the BREMM should not be weighted because of the difference of opinions about which criteria are more important. This is evidence that there is some consensus among mixed methods researchers from different disciplinary backgrounds, including both methodological and academic backgrounds, to what is important when evaluating the quality of mixed methods research.

BREMM Pilot Study

Following the creation of the BREMM, which was created using the results described in Chapter Five, additional steps were taken to document the validity of the instrument. A two-phase pilot study was conducted of the BREMM. The first phase was an item analysis of the items (criteria) presented in the BREMM. The second phase was an exploratory Factor Analysis (EFA) that tested the unidimensionality of the quality construct. The purpose for this pilot study of the BREMM was to demonstrate the procedures that could be used with a larger sample population to validate the instrument with statistical significance at a later date. The second phase is not reported here because of its small sample size and lack of significance.

Phase 1 of the Pilot Study: Procedure.

To collect data for the pilot study, a convenience sample of five volunteers was identified to use the rubric to rate 15 mixed methods research articles. The 15 article were randomly selected from a larger sample of 93 articles that contained the keywords "mixed methods" and "qualitative and quantitative". The five raters were doctoral students who completed at least one mixed methods course at the graduate level. In addition to having mixed methods coursework, each rater also had experience using a rubric to evaluate quality. Once the raters agreed to participate, I met with them to train them about the use of the rubric. This training was done on an individual basis, at the same time the rubrics and articles were distributed. The approximately 15 minute training sessions consisted of an overview of the BREMM, the operational definitions for each criterion, and how to score the criteria. The reviewers were then given 15 identical articles and asked to review as many of them as they could over a seven day period. During this time frame, the five students rated 45 articles using the BREMM. Two raters evaluated all 15 articles, while the remaining three each evaluated five articles. The 16 criteria in the BREMM were rated on a 6-point Likert scale, with each criterion being scored from "strongly disagree" to "strongly agree." The scores of the rated articles were then collected and used to conduct the item analysis and EFA. The scores for the first three articles are presented in Table 6.1. Some blanks are present because two of the raters did not rate the fifteen articles sequentially.

Table 6.1

Scores for the First Three Articles When Implementing the BREMM

Rater	Article 1 scores	Article 2 scores	Article 3 scores
Rater 1	42	24	53
Rater 2	Did not rate	Did not rate	53
Rater 3	Did not rate	29	Did not rate
Rater 4	54	38	57
Rater 5	51	52	70

Note. The articles were scored using the BREMM. The maximum score for each article was 75.

The scores provided by the raters may reflect the level of subjectivity that is used when evaluating quality. Of the first three articles, article three scored the highest amongst all of those who rated it.

Cross Research Question Integration

The first phase of this study revealed a reservation to use standards, guidelines, and/or rubrics to evaluate the quality of mixed methods research. Although the qualitative data explained this reservation in detail, a list of 14 criteria was created. The 14 criteria were the most frequently discussed topics pertaining to the quality of mixed methods research. In the second phase of this study, all the respondents ranked each of the 14 criteria as being important. Of the 14 criteria, 11 of them were rated as very important or higher.

Support for the importance of the 14 criteria identified from the interviews with mixed methods methodologists was achieved even though participants expressed a general lack of support for the idea of a single evaluation rubric or guideline being applied to all mixed methods research. Some differences were found in the mean ranking for the 14 criteria based on the participant's methodology and academic backgrounds. The failure to find meaningful differences among the respondents to the survey may be attributed to the general or broad nature of the way the criteria were worded. More

narrowly framed criteria of those that reflect a personal research agenda, such as specific designs, timing, and sequence are likely to have been more controversial.

Contribution of the Literature

This study detailed a process to evaluate the quality of mixed methods research. When looking for ways to evaluate the quality of mixed methods, researchers often turn to the NIH standards as a reference. The NIH standards for evaluating the quality of mixed methods research were primarily derived from several sources found in the mixed methods literature. These sources included both Creswell and Plano Clark's (2011) suggested criteria to evaluate a mixed methods study and O'Cathain et al.'s (2008) guidance for Good Reporting of a Mixed Methods Study (GRAMMS). Both of these guidelines were produced deductively from data gathered through a review of the literature. This study differs from the previous studies because it was conducted by means of a systematic empirical research project, instead of a literature review. This study also differs from other guidelines presented in the NIH standards because it was designed for a generalized audience. Other guidelines such as O'Cathain et al.'s were tailored towards particular groups. In the cases of O'Cathain et al. and NIH, their guidelines were designed for use by disciplines associated with health care.

Several of the criteria presented by Creswell and Plano Clark (2011) and O'Cathain et al. (2008) were found within the criteria presented in this study. Of the five criteria identified by Creswell and Plano Clark (2011), three are found in the criteria presented in this study. Those criteria include a clearly stated explanation of how the data are mixed in the research design, a clearly stated conclusion that includes the added value achieved through mixing, and evidence for the study's robustness and trustworthiness. Of

the six criteria presented by O'Cathain et al., four were similar to those presented in this study. Those four criteria were the use of research questions that guide the study, a clearly stated methods section, a description of how the data are mixed in the research design, and again, the inclusion of the added value gained through mixing. The criteria presented here are not exclusive to the guideline presented in this study, but do appear in the literature that focuses on the quality of mixed methods research.

Of the eight criteria presented by the NIH guidelines for approaching a mixed methods study in the consideration for grant funding, six were in alignment with the 15 criteria presented in this study. Only the NIH criteria of a study's feasibility and software support were not included in this study. This study confirmed the existing framework presented by the NIH by emphasizing several similar criteria.

This analysis suggests a possible gap exists between the views of experts and leading textbooks in the field of mixed methods on the topic of research designs. Creswell and Plano Clark (2011) have devoted considerable space to six different research designs in their textbook. The participants of this study did not share the view that the use of previously prescribed mixed methods designs was important. Instead, a purposeful and well explained research design is what was repeatedly emphasized as being important. Creswell and Plano Clark (2011) also identified O'Cathain et al.'s (2008) use of priority and sequence as an important element to assess in judging quality. Like specific designs, the timing and sequence was also not discussed by the participants as being important to a study's overall quality. The exclusion of these concepts is consistent with the participants' emphasis on the importance of leaving room for continued innovation and growth in the design of mixed methods studies.

The one criterion presented in this study that was not found in another set of criteria, was the level of transparency. A study's transparency is obtained through an explicit and detailed description of the procedures and rationale for the use of mixed methods. A study's transparency is evaluated through the collective of the criteria presented in the guideline presented in Chapter Five.

Implications

The future of this research will depend on the use of the guideline and BREMM by the public and additional research that explores topics that were unable to be included in this study. The guideline and rubric presented in this study are available to the public for the practical use of evaluating the quality of mixed methods research. Good uses for these tools include its use as a guide for conducting content analyses, meta-analyses, paper and projects reviews, and the synthesis of mixed methods within and across specific disciplines. The guideline presented in this study allows for researchers to ensure that a study possesses the components necessary to create a high degree of transparency across all phases of a study, which equates to quality. The BREMM provides a way to measure quality and further the use of mixed methods research across disciplines. Both the guideline and rubric were created from data obtained from noted experts in the field of mixed methods research and possess an inherent level of acceptability across qualitative and quantitative methods. Additional criteria have been identified in the literature, but not emphasized by the experts I consulted.

Below are listed several ways by which the guideline and BREMM can be improved. Additional research is needed to develop scales about dimensions of quality for individual criteria. Further clarification is needed for some terms, such as "mixing".

Differences among methodologists in the definition of terms may introduce complexity when using these tools. The scale must also be refined with more clear instructions for their use. When using the BREMM, the researcher must acknowledge the source of the instrument.

The Future of this Research

This study, like most research projects, brought forth more questions than it answered. The questions raised from this study provide years of additional research to both improve on the findings discussed within this study and to answer new questions that emerged. From this study, five major research goals have emerged to be further studied over the next decade. The first goal is to conduct a pilot study with the appropriate sample size to determine significantly significant results to the analyses presented in this section. The second goal is to use the remaining qualitative data to provide a more in-depth look at the future of mixed methods and views towards the meaning of research in terms of differences between qualitative, quantitative, and mixed methods disciplines. The third goal is to conduct further interviews and conduct a qualitative analysis of why the differences between different standards occur based on methodology and discipline. The fourth goal is to create a database on the use of the BREMM to learn how it adapts to mixed methods research as it changes. The final goal is to identify how the guidelines and rubric presented in this study have helped or hindered the teaching of mixed methods in the classroom environment.

The future research will likely be complicated and long-term, but the goal of all these projects is to strengthen mixed methods research and to help further establish it as the third primary research methodology. Although many factors will affect the feasibility

of these studies, with the assistance of additional researchers and funding, all are achievable within the next decade. In addition to strengthening and expanding mixed methods research, I also strive to introduce an acceptance of at least a minimum standard and reduce the negative connotation towards the use of rubrics and less subjective ways to evaluate the quality of mixed methods research.

References

Alise, M. A., & Teddlie, C. (2010). A continuation of the paradigm wars? Prevalence rates of methodological approaches across the social/behavioral sciences. *Journal of Mixed Methods, 4*(2), 103-126.

Bazeley, P. (2009). Editorial: Integrating data analyses in mixed methods research. *Journal of Mixed Methods Research, 3*(3), 203-207.

Becker, H. (1970). *Sociological work: Method and substance*. Chicago, Il: Aldine.

Blakie, N. (2010). *Designing social research* (2nd ed.). Cambridge, Sussex, England: Polity.

Bryman, A. (2007). Barriers to integrating quantitative and qualitative research. *Journal of Mixed Methods Research, 1*(1), 8-22.

Bryman, A. (2006). Integrating quantitative and qualitative research: How is it done? *Qualitative Research, 6*(1), 97-113.

Caracelli, V. T., & Greene, J. C. (1993). Data analysis strategies for mixed-method evaluation designs. *Educational Evaluation and Policy Analysis, 15*(2), 195-207.

Creswell, J. W. (2009). *Research Design: Qualitative, quantitative, and mixed methods approaches* (3rd ed.). Thousand Oaks, CA: Sage.

Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research.* Thousand Oaks, CA: Sage.

Creswell, J. W., Klassen, A. C., Plano Clark, V. L., & Smith, K. C. (2011, August). *Best practices for mixed methods research in the health sciences.* Retrieved December 30, 2011, from Office of Behavioral and Social Science Research: http://obssr.od.nih.gov/mixed_methods_research

Denscombe, M. (2008). Communities of practice: A research paradigm for the mixed methods approach. *Journal of Mixed Methods Research, 2*(3), 270-283.

Dewey, J. (1925). *Experience and nature.* Whitefish, MT: Kessinger.

Engelbrecht, J., Harding, A., & Du Preez, J. (2007). Long-term retention of basic mathematical knowledge and skills with engineering students. *European Journal of Engineering Education, 32*(6), 735-744.

Feilzer, M. Y. (2009). Doing mixed methods research pragmatically: Implications for the rediscovery of pragmatism as a research paradigm. *Journal of Mixed Methods Research, 4*(1), 6-16.

Fitzpatrick, J. L., Sanders, J. R., & Worthen, B. R. (2011). *Program evaluation: Alternative approaches and practical guidelines.* Princeton, NJ: Pearson.

Gambrel, L. E., & Butler VI, J. L. (2011). Mixed methods research in marriage and family therapy: A content analysis. *Journal of Marital & Family Therapy*. doi: 10.1111/j.1752-0606.2011.00260.x

Gill, J., Sharp, R., Mills, J., & Franzway, S. (2008). I still wanna be an engineer! Women, education and the engineering profession. *European Journal of Engineering Education, 33*(4), 391-402.

Greene, J. C., Caracelli, V. J., & Graham, W. F. (1989). Toward a conceptual framework for mixed-method evaluation design. *Educational Evaluation and Policy Analysis, 11*(3), 255-274.

Greene, J. (2007). *Mixed methods in social inquiry.* San Francisco, CA: Jossey-Bass.

Greene, J. (2000). Understanding social paradigms through evaluation. In N. K. Denzin, & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (2nd ed., pp. 981-1000). Thousand Oaks, CA: Sage.

Greene, J., & Henry, G. T. (2005). Qualitative-quantitative debate in evaluation. In S. Mathison (Ed.), *Encyclopedia of evaluation.* (pp. 345-350). Thousand Oaks, CA: Sage.

Harrits, G. S. (2011). More than method? A discussion of paradigm differences within mixed methods research. *Journal of Mixed Methods Research, 5*(2), 150-166.

Howell, D. C. (2010). *Statistical methods for psychology.* Belmont, CA: Wadsworth Cengage Learning.

Johnson, R. B., Onwuegbuzie, A., & Turner, L. A. (2007). Toward a definition of mixed methods research. *Journal of Mixed Methods Research, 1*(2), 1-22.

Journal of Mixed Methods Research. (2011, December). *Submitting a manuscript*. Retrieved January 20, 2012, from Journal of mixed methods research: http://www.sagepub.com/journals/Journal201775/manuscriptSubmission.

Kane, M. (2011). Validating score interpretations and uses: Messick lecture, language testing research colloquim, Cambridge, April 2010. *Language Testing, 29*(1), 3-17.

Kilgore, D., Jocuns, A., Yasuhara, K., & Atman, C. J. (2010). From beginning to end: How engineering students think and talk about sustainability across the life cycle. *International Journal of Engineering Education, 26*(2), 305-313.

Klassen, A., Creswell, J., Plano Clark, V., Smith, K., & Meissner, H. (2012). Best practices in mixed methods for quality of life research. *In Quality of Life Research,* 21, 377-380.

Lackoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind.* Chicago, IL: University of Chicago.

Larson, D., & Ahonen, A. M. (2001). Active learning in a Finnish engineering university course. *European Journal of Engineering Education, 29*(4), 521-531.

Leech, N. L., Dellinger, A. B., Brannagan, K. B., & Tanaka, H. (2009). Evaluating mixed research studies: A mixed methods approach. *Journal of Mixed Methods Research, 4*(1), 17-31.

Lou, S.J., Liu, Y., Chaung, S.Y., & Tseng, K.H. (2011). Effectiveness of on-line STEM project-based learning for female senior high school students. *International Journal of Engineering Education, 27*(2), 399-410.

Mark, M. M., Henry, G., & Julnes, G. (2000). *Evaluation: An integrated framework for understanding, guiding, and improving polisies and programs.* San Francisco, CA: Josey-Bass.

Maxwell, J. A. (2004) Causal explanation, qualitative research, and scientific inquiry in education. *Educational Researcher, 33*(2), 3-11.

Maxwell, J. A. (2010) Realism as a stance for mixed methods research. *Sage handbook of mixed methods in social & behavioral research* (2nd ed., pp. 145-167). Thousand Oaks, CA: Sage.

Maxwell, J. A. (2010). Using numbers in qualitative research. *Qualitative Inquiry, 16*(6), 475-482.

Maxwell, J., & Mittapalli, K. (2010). Realism as a stance for mixed methods research. In A. Tashakkori, & C. Teddlie (Eds.), *Sage handbook of mixed methods in social & behavioral research* (2nd ed., pp. 241-271). Thousand Oaks, CA: Sage.

Mertens, D. M. (1999). Inclusive evaluation: Implications of a transformative theory for evaluation. *American Journal of Evaluation, 20*(1), 1-14.

Mertens, D. M. (2010). Philosophy in mixed methods: The transformative paradigm as illistration. *International Journal of Multiple Research Approaches, 4*(1), 9-18.

Mertens, D. M. (2007). Transformative paradigm: Mixed methods and social justice. *Journal of Mixed Methods Research, 1*(3), 212-225.

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from person's responses and performances as scientific inquiry into score meaning. *American Psychologist, 50*(9), 741-749.

National Institutes of Health. (2012). *National Institutes of Health Office of Behavioral and Social Sciences Research*. Retrieved November, 17, 2012, from Best Practices for Mixed Methods Research in the Health Sciences: http://obssr.od.nih.gov/scientific_areas/methodology/mixed_methods_research/section2.aspx

National Science Foundation. (1997, August). *User-friendly handbook for mixed methods evaluations*. Retrieved December 30, 2012, from National Science Foundation: http://www.nsf.gov/pubs/1997/nsf97153/start.htm

O'Cathain, A., Murphy, E., & Nicholl, J. (2008). The quality of mixed methods studies in health services research. *Journal of Health Services Research Policy, 13*(2), 92-98.

Osterlind, S. J. (2010). *Modern measurement: Theory, principles, and applications of mental appraisal* (2nd ed.). Boston, MA: Pearson.

Pace, R., Pluye, P., Bartlett, G., Macaulay, A. C., Salsberg, J., Jagosh, J., et al. (2012). Testing the reliability and efficiency of th epilot Mixed Methods Appraisal Tool

(MMAT) for systematic mixed studies review. *International Journal of Nursing Studies, 49*, 47-53.

Patton, M. Q. (2002). *Qualitative research & evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage.

Pedhazur, E. J., & Schmelkin, L. P. (1991). *Measurement, design, and analysis: An integrated approach.* Hillsdale, NJ: Lawrence Erlbaum.

Plano Clark, V. (2010). The adoption and practice of mixed methods: US trends in federally funded health-related research. *Qualitative Inquiry, 16*(6), 428-440.

Plano Clark, V. L., Huddleston-Casas, C. A., Churchill, S. L., O'Neil Green, D., & Garrett, A. L. (2008). Mixed methods approaches in family science research. *Journal of Family Issues, 29*(11), 1543-1566.

Pluye, P., Gagnon, M.-P., Griffiths, F., & Johnson-Lafleur, J. (2009). A scoring system for appraising mixed methods research, and concomitantly appraising qualitative, quantitative and mixed methods primary studies in Mixed Studies Reviews. *International Journal of Nursing Studies, 46*, 529-546.

Raju, P. K., Sankar, C. S., & Xue, Y. (2004). A curriculum to enhance decision-making skills of technical personnel working in teams. *European Journal of Engineering Education, 29*(3), 437-450.

Reddy, M., & Andrade, H. (2010). A review of rubric use in higher education. *Assessment & Evaluation in Higher Education, 35*(4), 435-448.

Rorty, R. (1999). *Philosophy and social hope.* London: Penguin.

Rossman, G. B., & Wilson, B. L. (1985). Numbers and words: Combining quantitative and qualitative methods in a single large-scale evaluation study. *Education Review, 9*(5), 627-643.

Sagebiel, F., & Dahmen, J. (2006). Masculinities in organizational cultures in engineering in Europe: Results of the European Union project WomEng. *European Journal of Engineering Education, 31*(1), 5-14.

Sandelowski, M. (2001). Real qualitative researchers do not count: The use of numbers in qualitative research. *Research in Nursing & Health*, *24*(3), 230-240.

Sandelowski, M., Voils, C. I., & Knafl, G. (2009). On quantitizing. *Journal of Mixed Methods Research, 3*(3), 208-222.

Schifferdecker, K. E., & Reed, V. A. (2009). Using mixed methods research in medical education: basic guidelines for researchers. *Medical Education, 43*, 637-644.

Schultz, N., & Christensen, H. P. (2004). Seven-step problem-based learning in an interaction design course. *European Journal of Engineering Education*, *29*(4), 533-541.

Schwandt, T. A. (2008). Educating for intellegent belief in education. *American Journal of Evaluation, 29*(2), 139-150.

Scriven, M. (1967). The methodology of evaluation. *Curriculum evaluation, 1*, 39-83. (R. E. Stake, Ed.) Chicago, IL: Rand McNally.

Scriven, M. (1996). Types of evaluation and types of evaluator. *Evaluation Practice, 17*, 151-162.

Sevian, H., & Gonsalves, L. (2008). Analysing how scientists explain their research: A rubric for measuring the effectivenss of scientific explanations. *International Journal of Science Education, 30*(11), 1441-1467.

Small, M. L. (2011, April 15). *How to conduct a mixed mehtods study: Recent trends in a rapidly growing literature.* Retrieved November 26, 2011, from The Annual Review of Sociology: http://www.soc.annualreviews.org

Smith, A., & Dengiz, B. (2010). Women in engineering in Turkey - a large scale quantitative and qualitative examination. *European Journal of Engineering Education, 35*(1), 45-57.

Talmage, H. (1982). *Encyclopedia of educational research* (5th Edition ed.). (H. Mitzel, Ed.) New York, NY: Free Press.

Tashakkori, A., & Teddlie, C. (Eds.). (2003). *Handbook of mixed methods in social & behavioral research* (1st ed.). Thousand Oaks, CA: Sage.

Tashakkori, A., & Teddlie, C. (Eds.). (2010). *Mixed methods in social & behavioral research (*2nd ed.*).* Thousand Oaks, CA: Sage .

Teddlie, C., & Tashakkori, A. (2006). A general typology of research designs featuring mixed methods. *Research in the Schools, 13*(1), 12-28.

Tully, D., & Jacobs, B. (2010). Effects of a single-gender mathematics classrooms on self-perception of mathematical ability and post secondary engineering paths: An Australian case study. *European Journal of Engineering Education, 35*(4), 455-467.

Tyson, N. d. (2009). *The pluto files: The rise and fall of America's favorite planet.* New York, NY: W.W. Norton.

Wolcott, H. F. (1994). *Transforming qualitiative data: Description, analysis, and interpretation.* Thousand Oaks, CA: Sage.

Wolf, K., & Stevens, E. (2007). The roles of rubrics in advancing and assessing student learning. *The Journal of Effective Teaching, 7*(1), 3-14.

Woolley, C. M. (2009). Meeting the mixed methods challenge of integration in a sociological study of structure and agency. *Journal of Mixed Methods Research, 3*(1), 7-25.

Appendix A

VirginiaTech

Office of Research Compliance
Institutational Review Board
2000 Kraft Drive, Suite 2000 (0497)
Blacksburg, VA 24060
540/231-4606 Fax 540/231-0959
email irb@vt.edu
website http://www.irb.vt.edu

MEMORANDUM

DATE: May 14, 2012

TO: Elizabeth Creamer, Timothy Burrows

FROM: Virginia Tech Institutional Review Board (FWA00000572, expires May 31, 2014)

PROTOCOL TITLE: Designing a Preliminary Rubric to Evaluate Mixed Methods Research

IRB NUMBER: 12-346

Effective May 14, 2012, the Virginia Tech Institution Review Board (IRB) Administrator, Carmen T Green, approved the New Application request for the above-mentioned research protocol.

This approval provides permission to begin the human subject activities outlined in the IRB-approved protocol and supporting documents.

Plans to deviate from the approved protocol and/or supporting documents must be submitted to the IRB as an amendment request and approved by the IRB prior to the implementation of any changes, regardless of how minor, except where necessary to eliminate apparent immediate hazards to the subjects. Report promptly to the IRB any injuries or other unanticipated or adverse events involving risks or harms to human research subjects or others.

All investigators (listed above) are required to comply with the researcher requirements outlined at:

http://www.irb.vt.edu/pages/responsibilities.htm

(Please review responsibilities before the commencement of your research.)

PROTOCOL INFORMATION:

Approved As:	**Expedited, under 45 CFR 46.110 category(ies) 6,7**
Protocol Approval Date:	**May 14, 2012**
Protocol Expiration Date:	**May 13, 2013**
Continuing Review Due Date*:	**April 29, 2013**

*Date a Continuing Review application is due to the IRB office if human subject activities covered under this protocol, including data analysis, are to continue beyond the Protocol Expiration Date.

FEDERALLY FUNDED RESEARCH REQUIREMENTS:

Per federal regulations, 45 CFR 46.103(f), the IRB is required to compare all federally funded grant proposals/work statements to the IRB protocol(s) which cover the human research activities included in the proposal / work statement before funds are released. Note that this requirement does not apply to Exempt and Interim IRB protocols, or grants for which VT is not the primary awardee.

The table on the following page indicates whether grant proposals are related to this IRB protocol, and which of the listed proposals, if any, have been compared to this IRB protocol, if required.

Appendix B

Interview Protocol

Interview Protocol

Name:_____ ID No:_____
Interviewer: _____
Date: _____

Introductory Remarks

I am a doctoral student from Virginia Tech working to complete my dissertation. This study is about seeing what consensus there is about the criteria for evaluating the quality of mixed methods research publications. The purpose of this study mixed methods study is to conduct a cross-disciplinary, multi-phase content analysis to compare the outcomes of multiple rubrics for evaluating articles in multiple disciplinary fields.

I am using one-on-one phone interviews is to engage experts in the field of mixed methods research about quality measures. I want to get your views about the key elements of a mixed methods study that should be used to create a preliminary rubric for the scoring of mixed methods research studies. The interview should last between 45 minutes and one hour. All the questions that I am going to ask you are about your opinion or personal viewpoint. There are no right or wrong answers to any of these questions. All questions but one are open-ended. Do I have your permission to continue with the recorded interview?

Question 1: What has been your engagement with mixed methods research and the mixed methods community? (e.g., Do you go to conferences, write scholarly articles, conduct research using MMR, etc.)

Question 2: Why is the quality of mixed methods research important in the ways it is conducted and reported?

Question 2a: On a scale of 1-4, how important is the ability to identify quality in mixed methods research? One being not at all important, two being somewhat unimportant, three being somewhat important, four being necessary. _____

Question 3: When you look at a manuscript of a mixed methods publication, what things do you look for to evaluate its quality?

Question 4: Paradigms are often mentioned as being important in mixed methods research. What role, if any, do paradigms play in evaluating the quality of mixed methods research? Do you feel that they are an important part of an evaluation rubric? Why or why not?

Question 5: To what degree is mixing or some strategy for integrating the qualitative and quantitative strands necessary to the quality of a mixed methods study? Do you believe that mixing should occur in all phases of the study? Please explain your answer.

Question 6: Can you provide the name of any mixed methods researchers who you believe should be interviewed to provide additional insight about the topics discussed in this interview?

I would like to follow-up this interview during the beginning of the fall semester (2012) and ask you to rate a list of criteria that emerged most frequently from the interviews. The process of rating the criteria will be carried out through an email attachment. A list of criteria will be given and then you will be asked to place a number next to each criteria in order of importance (one being the most important, two the next important, and so forth).This process should only take 10 minutes of your time. Are you willing to participate in this final component of the study?

Thank you for taking the time to assist with my research, as I know you time is very valuable. As a recognized expert in your field, your opinions on the matters discussed in this interview are most valuable and will lead to the highest quality data.

Appendix C

E-mail One to Participants

Dear (insert name),

As someone who is considered to be a recognized expert in mixed methods research (MMR), I am writing to ask if you would agree to be interviewed between May 10 and June 30th, 2012 on your views about what criteria are important to evaluate the quality of MMR. During the interview, I will be asking you questions about how you currently evaluate quality in MMR, the role of paradigms as they pertain to quality in MMR, the role mixing plays in evaluating quality in MMR, and the overall importance of quality in MMR. The interview should take approximately one hour. This study is important to the field of mixed methods because it will provide a rubric that will provide a standardized method to evaluate the overall quality of MMR that may be utilized by both students and practitioners in the field.

If you are able to participate in this study, please respond to this email. I will send you further details, including information about setting up an interview time.

For additional information about me and this study, including the research design, purpose statement and research questions, please click on the link below.

https://virginiatech.qualtrics.com/SE/?SID=SV_bfs1YloXDcbGoPq

Thank you for your time and consideration,

Timothy Burrows
Doctoral Candidate
Educational Research & Evaluation
Graduate Assistant
Office of Assessment and Evaluation
Virginia Tech
tburrows@vt.edu

Appendix D

E-mail Two to Participants

Dear (insert name),

I appreciate your willingness to participate in this study to design a rubric to evaluate mixed methods research articles. Please indicate below what dates you are available to be interviewed. I will return your email with a Meeting Wizard/Doodle offering a number of time options for the dates you specified. Once we have set the time, I will follow-up by sending the IRB approved consent form for your approval. If you are not available during the May to late June time frame, please put an "x" in the "Other" box and fill in dates that you might be available for an interview.

Available times for interviews

May 13-19	May 20-26	May 27-June 2	June 3-9	June 10-16

June 17-23	June 24-30	Other

Before you agree to the interview I can confirm that:

- With your permission an acknowledgement will be made to you for your support in this research in all publications.
- With your permission the interview will be recorded.
- A transcript of the interview will be sent to you after it is transcribed.
- Your confidentiality will be maintained at all times and no comments will be ascribed to you by name in any written document or verbal presentation without your written permission.
- You will be free to withdraw from the research at anytime and/or request that your transcript not be used.
- A copy of the interview questions will be sent to you before the interview.
- A copy of the manuscript pertaining to the interview data will be made available to you upon request (spring 2013).

If you have any queries concerning the nature of this research, please contact me or my advisor, Dr. Elizabeth G. Creamer, Professor, Educational Research and Evaluation, Virginia Tech.

Yours sincerely,

Timothy Burrows
Doctoral Candidate
Educational Research & Evaluation
Graduate Assistant
Office of Assessment and Evaluation
Virginia Tech
tburrows@vt.edu

Appendix E

Participant Consent Form

VIRGINIA POLYTECHNIC INSTITUTE AND STATE UNIVERSITY
Information for Participants in Research Projects Involving Human Subjects

Title of Project: Shaping Global Perspective: A Preliminary Rubric Design to Evaluate Mixed Methods Research Articles

Investigators: Timothy Burrows & Elizabeth Creamer

I. Purpose of this Project
The purpose of this mixed methods study is to conduct a cross-disciplinary, multi-phase content analysis to create a preliminary rubric to evaluate mixed methods research articles in multiple disciplinary fields. We are asking you, an expert in the mixed methods field to participate in a recorded telephone or computer audio conferencing interview (Skype). This research is being conducted to create a rubric that can be used to evaluate mixed methods research. Your participation in this study is voluntary and important to the success of this study.

II. Procedures
1) The participants will be contacted via telephone or computer (Skype) according to a time provided by the participant during the recruitment e-mail. The participant will be given a link to a copy of the protocol in an email prior to the interview so they have had time to fully understand the questions or to clerify any questions they may have. I will follow the protocol, which begins with a verbal consent and permission to leave the study at any time for any reason. This interview will last between 45 minutes and 1 hour. The interview participant will participate in the phone interview from a location of his or her choice.
2) A a follow-up, the participant will be asked to review the test rubric items developed from the interview process and rate them in order of importance based on their opinion. The review and rating of the rubric items should take between 5-10 minutes. The rating will be done through an e-mail.

III. Risks
Participation in this project does not place you at more than minimal risk of harm. It is possible that participants may feel uncomfortable discussing personal opinions about others researchers. If you agree to participate, you may choose to stop the journal entry at any time, or leave any item unanswered.

IV. Benefits of Project
This study will help to clarify which components, and how they are addressed, of mixed methods research are critical to creating a quality mixed methods research study. We

hope that other researchers and students will benefit from the findings of this study. As a participant, you agree that no promise or guarantee of benefits has been made to encourage you to participate in this study.

V. Extent of Anonymity and Confidentiality
The results of this study will be kept strictly confidential. No one beyond the investigation team will be able to connect any data with your name. The information you provide will be reported in publications in aggregate form to protect your identity and to guard against possibly connecting you to your responses.

In addition to our research team, it is possible that the Virginia Tech Institutional Review Board (IRB) may view this study's collected data for auditing purposes. The IRB is responsible for the oversight of the protection of human subjects involved in research.

VI. Compensation
You will not be compensated for your time in participating in this study.

VII. Freedom to Withdraw
Your participation in this study is entirely voluntary. You are free to withdraw or refuse to answer any question, at any time, for any reason without negative consequences.

VIII. Approval of Research
This research has been approved, as required, by the Institutional Review Board (IRB) for Research Involving Human Subjects at Virginia Polytechnic Institute and State University.

IX. Researchers' Contact
Should you have any pertinent questions about this research, its conduct, or research subjects' rights you may contact:

Timothy Burrows
Doctoral Candidate
Educational Research & Evaluation
Graduate Assistant
Office of Assessment and Evaluation
Virginia Tech
tburrows@vt.edu

Dr. Elizabeth G. Creamer
Program Leader, Education Research and Evaluation (EDRE)
Professor, Educational Research and Evaluation
School of Education
Virginia Tech
211 E. Eggleston
Blacksburg, VA 24061
Work phone: (540) 231-8441

http://web.me.com/egcreamer/EGCreamer.com/Consulting_Bio.html
http://web.me.com/egcreamer/Traveling-Foodie.com
CO-PI, NSF Funded Project, Appalachian Information Technology Services (AITES)

David Moore, Ph.D.
Chair, Institutional Review Board
Virginia Tech
Office of Research Compliance (Attn: IRB)
2000 Kraft Drive, Suite 2000 (0497)
Blacksburg, VA 24060
540-231-4991 / moored@vt.edu

X. Participant's Permission
I have read and understand the Informed Consent and conditions of this project. I have had all my questions answered. I hereby acknowledge the above and give my voluntary consent for participation in this project.

Participant Signature _____ Date _____