



UNIVERSITY
OF TAMPERE

This document has been downloaded from
Tampub – The Institutional Repository of University of Tampere

Post-print

Authors: Ihantola Eeva-Mari, Kihn Lili-Anne
Name of article: Threats to validity and reliability in mixed methods accounting research
Year of publication: 2011
Name of journal: Qualitative Research in Accounting and Management
Volume: 8
Number of issue: 1
Pages: 39-58
ISSN: 1176-6093
Discipline: Social sciences / Business and Management
Language: en
School/Other Unit: School of Management

URN: <http://urn.fi/urn:nbn:uta-3-487>

DOI: <http://dx.doi.org/10.1108/1176609111124694>

All material supplied via TamPub is protected by copyright and other intellectual property rights, and duplication or sale of all part of any of the repository collections is not permitted, except that material may be duplicated by you for your research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone who is not an authorized user.

“Threats to Validity and Reliability in Mixed Methods Accounting Research”

By Eeva-Mari Ihantola & Lili-Anne Kihn

School of Management, University of Tampere, Finland

Qualitative Research in Accounting and Management, Vol. 8, No. 1, 2011, pp. 39-58

<http://www.emeraldinsight.com/journals.htm?issn=1176-6093&volume=8&issue=1>

This article is (c) Emerald Group Publishing and permission has been granted for this version to appear here (<http://www.uta.fi/>). Emerald does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from Emerald Group Publishing Limited.

“Threats to Validity and Reliability in Mixed Methods Accounting Research”

Abstract

Purpose – The purpose of this study is to shed light on the threats to quality in mixed methods accounting research, wherein quantitative and qualitative approaches are combined in data collection, analysis and interpretation.

Design – Our paper is framed according to the following three perspectives: We first synthesize the threats to validity and reliability in quantitative and qualitative parts of mixed methods research using the quality standards of each. We then introduce an integrative framework of mixed methods research quality by Teddlie and Tashakkori (2003, see also and Tashakkori and Teddlie, 2008). Thereafter, we address the specific threats to quality that come to the fore when inferences from the quantitative and qualitative components of the study are combined to form meta-inferences using a legitimation framework by Onwuegbuzie and Johnson (2006).

Findings – Our analysis not only indicates a wide range of threats to the validity and reliability of mixed methods research in a range of categories, but also clarifies how the three perspectives described in this paper are linked and supplement each other.

Research limitations – Methodological research published in English over the last decade is emphasized to create an approach to assess mixed methods accounting research. The frameworks analyzed could still be studied in greater detail. Additional perspectives on the validity and reliability of mixed methods research could also be studied and developed.

Practical implications – This study furthers our understanding of such new developments in methodological research which may be of great importance to those conducting or evaluating empirical research.

Originality/value – Based on a comprehensive synthesis, this paper presents and analyzes theoretical frameworks potentially useful for scholars, students and practitioners. It focuses on both traditional and novel areas of validity and reliability in mixed methods research.

Keywords – accounting research, mixed methods, reliability, validity.

Classification: Conceptual analysis, literature analysis

1. Introduction

While empirical accounting studies have traditionally been based on either quantitative or qualitative methods, triangulation or mixing of such methods in the data collection, analysis and interpretation has also been called for (see, e.g., Creswell and Clark, 2007; Creswell, 2009; Eriksson and Kovalainen, 2008, 293; Ryan et al., 2002; Yin, 1994, 92). Such “mixed methods research” has been proposed for the following reasons: to improve validity of theoretical propositions and to obtain a more complete (less biased) picture of the phenomenon under study than it is possible with a more narrow methodological approach (Webb et al., 1966). It has also been considered useful in specifying research questions, familiarizing the scholar with the subject and/or context, and in confirming that all respondents understand the concepts and measures in a similar way. Mixed methods research has been recommended in uncharted regions where theoretical roadmaps do not yet exist, but where it is important to apply several methods to stay on firm ground to arrive safely at the destination. (For a review, see Hurmerinta-Peltomäki and Nummela, 2006, 440). Triangulation of methods can enable a case researcher to address a broader range of historical, attitudinal and behavioural issues, and to develop converging lines of inquiry that can be used to make case study findings and conclusions more convincing and accurate (Yin, 1994, 92). Triangulation in its various forms has also been considered useful in improving the reliability of a study (Lillis, 2006; Lukka, 1988, 423). Some other rationales for conducting mixed methods research are (Collins et al., 2006, 76): participant enrichment, instrument fidelity, treatment integrity, and significance enhancement.

As the above examples suggest, mixed methods research offers researchers many opportunities. In essence, qualitative data collection and/or analysis can be combined with quantitative data collection and/or analysis either concurrently or sequentially, in one or more stages in the research process and to different degrees (see Brannen, 1992, 12, 23; Bryman, 1988; Creswell, 2009, 206-208). For example, the use of a qualitative method can be used to facilitate the quantitative part of the study, or then the other way around, or then both approaches can be given equal emphasis (see Bryman, 1988). As a result, as many as 13 different strategies have been identified by Hurmerinta-Peltomäki and Nummela (2006, 447) ranging from “qualitative data analyzed quantitatively” to “qualitative and quantitative data analyzed concurrently with qualitative and quantitative research methods.”

Although mixed methods research has been presented as another step forward utilizing the strengths of both qualitative and quantitative approaches (see Creswell, 2009, 203), it has also been acknowledged that it is not a panacea for all researchers and research problems in all circumstances (see e.g., Collins et al. 2006, 69). This is not least because of the considerable resources and researchers' methodological skills it necessitates (Bryman, 1992, 69) but also because the triangulation of methodologies results in differing ontological and epistemological assumptions, which can be challenging to combine (Blaikie, 1991). Last but not least, discussions on the more complex validity and reliability issues of mixed methods research are only now emerging.

The purpose of this study is to shed light on the threats to quality in mixed methods accounting research design and analysis. We approach this issue from the following three perspectives: We first synthesize the threats to validity and reliability in quantitative and qualitative phases of a mixed methods study using the quality standards of each. This is because many researchers have recommended the use of standard procedures for both the quantitative and qualitative phases of the study (see Creswell, 2009, 219; Dellinger and Leech, 2007, 314; Onwuegbuzie and Johnson, 2006, 56; Tashakkori and Teddlie, 1998). Only some of the most salient threats to the quality of quantitative and qualitative methods are reviewed here, because these have been extensively analyzed elsewhere. We then introduce an integrative framework of mixed methods research by Teddlie and Tashakkori (2003, and Tashakkori and Teddlie, 2008). They suggest inference quality and inference transferability as umbrella terms that researchers could use in assessing validity in mixed methods research. Thereafter, we present a legitimation framework by Onwuegbuzie and Johnson (2006). It addresses the specific threats to quality that come to the fore when inferences from the quantitative and qualitative components of the study are combined to form meta-inferences.

This paper contributes to accounting research by providing an inventory of various threats to validity and reliability in a range of categories. Existing accounting literature has generally focused on validity and reliability issues in either the quantitative approach (e.g., Abernethy et al., 1999; Nazari et al., 2006; Searcy and Mentzer, 2003; Van der Stede et al., 2007) or the qualitative approach (Ahrens and Chapman, 2007; Lillis and Mundy, 2005; Lillis, 2006; Lukka and Kasanen, 1995; Lukka and Modell, 2010; Vaivio, 2008, etc.), but not on the mixed methods approach. In addition, our paper contributes to the methodological literature

by clarifying how the three perspectives analyzed in this paper are linked and complement each other.

In the second section of this paper, we present a brief synthesis of the threats to the quality of quantitative and qualitative parts of research. Thereafter, in Section Three, we address the specific quality threats to mixed methods research based on two different frameworks (Teddlie and Tashakkori's 2003 (see also Tashakkori and Teddlie, 2008), and Onwuegbuzie and Johnson's, 2006). The fourth and final section comprises the conclusions.

2 Threats to the quality of quantitative and qualitative parts of the work

When doing mixed methods research, it is important to seek to compensate the weaknesses of one method with the strengths of another method (see Onwuegbuzie and Johnson, 2006). Below, we provide a brief synthesis of such threats to validity and reliability that should also be taken into account in the a) quantitative and b) qualitative parts of mixed methods research. Following the Ryan et al. (2002) classification, we focus on *internal* and *external validity* and *reliability* of quantitative work and, analogously, on *contextual validity*, *generalizability and transferability*, and *procedural reliability* of qualitative work. We also seek to group the threats according to the research stages, i.e., whether they are most likely to be present during research design, data collection, and/or data analysis and interpretation (cf., Onwuegbuzie, 2003). This is because mixed methods studies impact all these elements as such studies are designed to draw on the value of multiple data collection methods, both quantitative and qualitative data analysis, and interpretation drawing together multiple perspectives on evidence.

2.1 Internal (contextual) validity

Quantitative approach. Internal (contextual) validity, as it is called in quantitative (qualitative) research, is one of the most essential manifestations of validity. In quantitative accounting research, the ultimate question is whether we can draw valid conclusions from a study given the research design and controls employed (Ryan et al., 2002, 141). Internal validity asserts that variations in the dependent variable result from variations in the independent variable(s) – not from other confounding factors (Abernethy et al., 1999, 16). To an extent, it is about the logic between a piece of research and existing theory (Arbnor and

Bjerke, 1977, 217). A central question is how the theory has been built based on previous studies? (Arbnor and Bjerke, 1977, 217) In experiments, internal validity is also determined by how much control has been achieved in the study during data collection (Ryan et al. 2002, 122). The use of statistical control variables is also important in survey research.

In qualitative research, contextual validity refers to the *credibility* of case study evidence and the conclusions drawn (Ryan et al. (2002, 155-156). The primary focus of such research is to capture *authentically* the lived experiences of people and to represent them in a *convincing* text, which demonstrates that the researcher fully understands the case (see Golden-Bibble and Locke, 1993; Lukka and Modell, 2010, 3; Ryan et al. 2002, 158). The key question to ask is “did we indeed capture the phenomenon or attribute that we intended to (or we believe we captured)” (Tashakkori and Teddlie, 2003, 694).

Threats to the internal validity of quantitative work may occur throughout the research process. A good research design is always of crucial importance when pursuing high internal validity. During research design, the threats to internal validity include insufficient knowledge of, or contradictions in the logic. However, deficiencies in the later stages of research – i.e., during data collection, analysis and/or interpretation – can also lead to studies with low internal validity. During data collection, possible threats to internal validity are many including, for example, instrumentation issues (Campbell and Stanley, 1963 in Tashakkori and Teddlie, 1998, 87), order bias and researcher bias in the use of techniques (see Ongwuegbuzie, 2003). Instrumentation issues occur when scores yielded from a measure lack the appropriate level of consistency or do not generate valid scores (as a result of inadequate content, criterion and/or construct validity). Order bias occurs if the effect of the order of the intervention conditions cannot be separated from the effect of the intervention conditions. Researcher bias means that the researcher has a personal bias in favor of one technique over another. Errors in statistical testing, illusory correlation and causal error are some examples of threats during data analysis and interpretation. (See further Appendix 1).

Appendix 2 synthesizes some of the threats to contextual validity of qualitative work. *Insufficient or biased knowledge* of earlier studies and theories (see Näsi, 1979, 302) and *contradictions in the logic* (such as a mismatch between research question and study design, see Lillis, 2006, 467) threaten contextual validity during the research design phase. The

following are some of the threats to contextual validity during data collection (McKinnon, 1988, 37-41): *observer-caused effect, observer bias, researcher bias, data access limitations, and complexities and limitations of the human mind*. Finally, the threats to contextual validity during data analysis and interpretation are many ranging from *lack of descriptive validity* of settings and events (see Maxwell, 1992) to *effect size* (see Onwuegbuzie and Leech, 2007, 235-237).

2.2 External validity (*generalisability and transferability*)

External validity is a key criterion in *quantitative research* (Ryan et al., 2002, 123). It determines whether one can draw more general conclusions on the basis of the model used and data collected, and whether results may be generalized to other samples, time periods and settings. The following three typical problems may threaten the external validity of a quantitative study (Ryan et al., 2002, 123-124): population, time and environmental validity. *Population validity* refers to whether inferences can be drawn from a study of a given population. The questions analyzed concern, for example, whether a relationship between two variables also exists in the population at large and not only in the sample selected. External validity is seriously threatened, if biases or other limitations exist in the accessible population. If the sample size is inadequate and/or the sample is not random, the estimates may be meaningless, because the sample may not faithfully reflect the entire population (cf., Howell, 1995, 6-7). In such cases generalizations should not be made to the target population. *Time validity* shows the extent to which the results of a particular study at a point in time can be generalized to other time periods. If structural changes in the relationships between variables occur, the time validity of such a study will be low. *Environmental validity* indicates whether results can be generalized across settings. International generalizability is an example of a potential problem (Ryan et al. 2002, 123-124).¹

In qualitative research, generalizability is concerned with whether the research results are *transferable* (Lincoln and Guba, 1985), i.e. can be extended to a wider context (Eriksson and Kovalainen, 2008, 293-204), have *theoretical generalizability* (see, e.g. Ryan et al. 2002,

¹ Some other possible threats to external validity are: multi-treatment interference, researcher bias, reactive arrangements, order bias, matching bias, specificity of variables, treatment diffusion, pretest x treatment interaction, and selection x treatment interaction (see Onwuegbuzie, 2003).

149), *empirical applicability* (Näsi, 1979, 292), *practical usefulness* (Arbnor and Bjerke, 1977; Mäkinen, 1980), *contextual and/or constructive generalizability* (Lukka and Kasanen, 1995). Severe threats to the transferability of a qualitative study may occur due to *selective plausibility*. That is, if the researcher, for example, fails to reconnect the empirical findings of the study at hand to those of other cases and theories, and explain how the new evidence enhances our understanding of the research question (see Golden-Bibble and Locke, 1993, 600). The lack of comparison between empirical findings and previous theoretical contributions can lead to rather myopic conclusions and, in the worst case, a scholar may claim to have discovered something already demonstrated in other studies (see, Vaivio, 2008, 76).

2.3 (Procedural) reliability

Quantitative approach. Reliability generally refers to the extent to which a variable or set of variables is consistent in what it is intended to measure. When multiple measurements are taken, the reliable measures will all be consistent in their values. (Hair et al., 2006, 3). Lack of reliability refers to random or chance error. If measurement results are not reliable, it becomes more difficult and precarious to test hypotheses or to make inferences about the relations between variables in quantitative research (Kerlinger, 1964). The following issues also represent some serious threats to reliability during data collection (cf. Kerlinger, 1964, 442-443): *lack of clear and standard instructions, measurement instruments describe items ambiguously so that they are misinterpreted, abstract concepts are not measured with enough indicators of equal kind and administration conditions differ*. Fink and Kosecoff (1985, 50) also cite the following threats to reliability: *lack of pretesting, not all alternatives are provided, the questions are not presented in the proper order, the questionnaire is too long or hard to read, and the interview takes too long*. Failure to answer questions, giving several answers to the same question and comments in the margin may all indicate lack of reliability. All these issues also represent threats to reliability during data collection. *Random sources of error* – such as typos and other errors in data collecting, saving and analysis (see Alkula et al., 2002) – may threaten reliability at every stage of research process.

In qualitative research, procedural reliability is related to *consistency*, typically meaning that another person should be able to examine the work and come to similar conclusions (see

Douglas, 1971; Eriksson and Kovalainen, 2008, 292; Grönfors, 1982; Koskinen et al., 2005, 258; Ryan et al., 2002, 155). The key question to ask is: “Did we *accurately* capture/represent the phenomenon or attribute under investigation?” (Tashakkori and Teddlie, 2003, 694). Careful documenting and reporting should allow the reader to assess how the researcher has collected, produced and interpreted the data. However, there are threats to reliability also at every stage of the qualitative research process (Lillis, 2006, 472).

The following are threats to procedural reliability during data collection: *Inaccurate and unsystematic interview questions* and *inaccurate transcriptions* (Koskinen et al. 2005, 262-263). *Failure to tape-record or take notes on the spot* may increase random errors, and *not having a comprehensive research plan, a coherent set of field notes on all evidence, or fully-documented case analysis* is also problematic (cf., Ryan et al., 2002, 154-155). *Relations that develop between researchers and participants* may also threaten procedural reliability during data collection (see Lillis, 2006). The procedural reliability of qualitative research may also be impaired *if the data is not collected over a long enough period of time, additional questions are not posed to interviewees when needed* (McKinnon, 1988, 40-51), and *the researcher is not aware of informal evidence* (see, Ryan et al., 2002, 154-155).

Finally, *errors may also occur in data classification, attaching data to constructs, drawing linkages between constructs, reduction, interpretation and development of links with theory, etc. Not taking distance from preconceptions* is also problematic. (See Lillis, 2006, 470-471) All of these represent threats to procedural reliability during data analysis and interpretation.

2.4 Conclusion

In this section we analyzed threats to validity and reliability in quantitative and qualitative parts of mixed methods research using the quality standards of each. Our analysis indicates a wide range of threats to the validity and reliability of mixed methods research in the following categories (Ryan et al., 2002): internal (contextual) validity, external validity (generalizability and transferability) and (procedural) reliability. The threats may occur during the following phases of research (cf. Onwuegbuzie, 2003): research design, data collection, analysis and interpretation. While this perspective addresses important issues that need to be taken into consideration, the traditional criteria are not sufficient alone for the

purposes of mixed methods research. Other more specific threats are likely to exist in mixed methods research when quantitative and qualitative approaches are combined in research design, data collection, analysis and interpretation.

3 Specific threats to the quality in mixed methods research

An emerging field of research is considering how validity might be different for mixed methods studies than for a quantitative or a qualitative study. In particular, the following two frameworks can be mentioned: Teddlie and Tashakkori's (2003, see also Tashakkori and Teddlie's, 2008) and Onwuegbuzie and Johnson's (2006).

3.1. Integrative framework

Teddlie and Tashakkori's (2003) and Tashakkori and Teddlie's (2008) work builds on previous research on validity issues of quantitative and qualitative research. They extend it by developing new terms that can be used to discuss validity of mixed methods research. They also begin with internal validity, but call internal validity and credibility *inference quality* and divide it into *design quality* (that refers to the standards used for the evaluation of the methodological rigor of the mixed methods research) and *interpretive rigor* (that pertains to the standards for evaluating the validity of conclusions, see also Lincoln and Guba, 2000). According to Teddlie and Tashakkori's (2008) expanded framework, design quality is reflected by design suitability, design adequacy or fidelity, analytic adequacy and within-design consistency. *Design suitability* refers to whether the methods of a study are appropriate for answering the research questions and whether the design matches the research questions. *Design adequacy/fidelity* is concerned with whether the components of the design were implemented adequately. *Analytic adequacy* addresses the questions of whether the data analysis techniques are appropriate and adequate to answer the research questions. *Within-design consistency* (i.e. the consistency of the procedures/study design from which the inferences emerge) is threatened if any of the following conditions referring to contradictions in logic (cf., Teddlie and Tashakkori, 2003, 40):

- The design is not consistent with the research questions/purpose
- The observations/measures do not demonstrate validity.

- The data analysis techniques are not sufficient and appropriate for providing answers to research questions.
- The results do not have the necessary strength or frequency to warrant the conclusions.
- The inferences are not consistent with the results of data analysis.
- The inferences are not consistent with the research questions/purposes.

According to Tashakkori and Teddlie's (2008) expanded framework, interpretive rigor is indicated by the following: interpretive consistency, theoretical consistency, interpretive agreement, interpretive distinctiveness and integrative efficacy. The first one (i.e., *interpretive consistency*) has to do with the consistency of inferences with each other and with the results of data analysis. For example, does each conclusion faithfully follow the findings and do multiple conclusions based on the same results agree with each other? That is, is the type of inference consistent with the type of evidence and is the level of intensity reported consistent with the magnitude of the events or the effects that were found? *Theoretical consistency* addresses whether each inference (explanation for the results or for relationships) is consistent with current theories in the academic field and/or with empirical findings of other studies?

Interpretive agreement refers to the consistency of interpretations across scholars and participants' construction of reality. Threats to interpretive agreement exist if (cf., Teddlie and Tashakkori, 2003, 41): other scholars do not agree that the inferences are the most plausible interpretations of the findings, and (if the participants' construction of the events/relationships is important to the researcher) the interpretations do not make sense to the participants of the study. *Interpretive distinctiveness* is the degree to which the inferences are distinctively different from other possible interpretations of the results and the rival explanations are eliminated. It is not demonstrated if (Teddlie and Tashakkori, 2003, 41): the inferences are not distinctively superior to other interpretations of the same finding – i.e., if there are other plausible explanations for the findings.

Finally, *integrative efficacy* is the degree to which inferences made in each strand of a mixed methods study are effectively integrated into a theoretically consistent meta-inference. The four previous criteria related to interpretative rigor were applicable to both qualitative and

quantitative parts of research and to the meta-inferences that emerge when the inferences of the two or more parts are integrated. By contrast, integrative efficacy is unique to meta-inferences in mixed methods. It is concerned with the degree to which a mixed methods researcher adequately integrates the findings, conclusions and policy recommendations gleaned from each of the two strands (i.e., makes meaningful conclusions of them, see further Tashakkori and Teddlie, 2008).

In their framework, Teddlie and Tashakkori (2003, 38) also borrowed the term transferability from Lincoln and Guba (1985) to construct *inference transferability* as an umbrella term for the concepts of external validity (used in quantitative literature) and transferability (used in qualitative literature). In line with existing quantitative literature, they also defined the following specific types of transferability: *population transferability* (to other individuals, groups or entities), *ecological transferability* (to other contexts and settings), *temporal transferability* (to other time periods), and *operational transferability* (to other modes/methods of measuring/observing the variables/behaviours).

In conclusion, these new terms developed by Teddlie and Tashakkori (2003) and Tashakkori and Teddlie (2008) make an important contribution to mixed methods research as there has been a lack of joint vocabulary for the overall assessment of validity in mixed methods research. The new umbrella terms allow us to address some more specific forms of validity in mixed methods research. In so doing, they decrease the need to rely on quantitative and qualitative terms only. However, the overall significance of the new terms will depend on how generally they become accepted.

3.2 A legitimization framework

Onwuegbuzie and Johnson (2006, 56) have been concerned that Teddlie and Tashakkori's (2003) framework may give a false impression that validation is an outcome only and, hence, not all steps in the research process are not equally important. Accordingly, the legitimization framework by Onwuegbuzie and Johnson (2006, see also Collins et al., 2006) has quite different objectives from Teddlie and Tashakkori's (2003) framework. In their framework,

Onwuegbuzie and Johnson (2006, 55-60) call validity *legitimation*.² They emphasize that legitimation is not an outcome, but a continuous, iterative and interactive process that should occur at each stage of the mixed research process, whether quantitative, qualitative or both. They also stress the need to address several types of legitimation that come to the fore as a result of combining inferences from the quantitative and qualitative components of a mixed research study to form meta-inferences. In so doing, they introduce several novel dimensions to the validity of mixed method research that have not, to the best of our knowledge, been addressed before. In particular, Onwuegbuzie and Johnson (2006) describe the following nine types of legitimation (see Table 3):

[Table 3 about here]

First, *sample integration legitimation* is “the extent to which the relationship between the quantitative and qualitative sampling designs yields quality meta-inferences” (ibid., 57). A serious threat is that *unless exactly the same individuals (or groups) are involved in both the qualitative and quantitative arms of a study, constructing meta-inferences by drawing together the inferences from the qualitative and quantitative phases can be problematic*. For example, if the group of managers interviewed is very small or different from the group of managers that has responded to a questionnaire it may not be justified for a meta-inference to include inferences from the qualitative component. The meta-inference may be poor because of the unrepresentative sample from the qualitative phase which, in turn, would affect statistical generalizability (population transferability). The situation becomes even worse if the quantitative sample is non-random and/or too small.

Second, *inside-outside legitimation* refers to “the extent to which the researcher accurately presents and appropriately utilizes the insider's view and the observer's view for purposes such as description and explanation” (ibid., 57). Certain tensions exist, because quantitative research often seeks the objective outsider view and qualitative research seeks interpretations made by insider. The basic threat to the inside-outside legitimation in mixed methods research is that these two viewpoints are not fully in balance. This occurs if the researcher,

² Collins et al. (2006) also refer to validity and legitimation as trustworthiness, credibility, dependability, plausibility, applicability, consistency, neutrality, reliability, objectivity, confirmability, and/or transferability of quantitative and/or qualitative data and interpretations stemming from them.

for example, fails to maintain a well informed and balanced perspective when collecting, analyzing, and interpreting what the whole set of qualitative and quantitative data mean (ibid. 58).³

Third, *weakness minimization legitimation* refers to “the extent to which the weakness from one approach is compensated by the strengths from the other approach” (ibid., 57). For weakness minimization legitimation it is important that the threats to the quality of the quantitative and qualitative parts of mixed methods research are carefully identified. According to this knowledge the researcher should plan, design and implement the study so that the possible threats and weaknesses from one approach can be compensated by the strengths from another approach.

Fourth, *sequential legitimation* means “the extent to which one has minimized the potential problem wherein the meta-inferences could be affected by reversing the sequence of the quantitative and qualitative phases” (ibid., 57). If a sequential mixed research design is used, *it is possible that the meta-inference is the effect of the sequencing itself*. That is, if the results and interpretations had been different if the order of the quantitative and qualitative phases had been reversed, then this would suggest that the sequencing itself was a threat to legitimation.⁴

Fifth, in Onwuegbuzie and Johnson’s (2006, 57) typology *conversion legitimation* refers to “the extent to which the quantitizing or qualitzing yields quality meta-inferences”. Counting is a very common way to quantify qualitative data. Numbers can complement and enhance narratives, but numbers must be used in ways that produce trustworthy findings. Sandelowski (2001, 230) warns about counting pitfalls associated with verbal counting, misleading counting, acontextual and overcounting. Verbal counting occurs when researcher implies

³ A strategy for obtaining a justified meta-inference is, first, to maintain a clear understanding of the meaning of qualitative and quantitative data when collecting, analyzing and interpreting it and, second, to use peer review to obtain a justified etic viewpoint, and third, use member checking or participant review to obtain a justified emic viewpoint, and finally, integrate the parts. (ibid. 58; see more on “etic” and “emic” in Currall and Towler, 2003, 522)

⁴ This threat can be assessed by changing the sequential design to a multiple wave design in which the quantitative and qualitative data collection and analysis phases alternate (Sandelowski, 2003; Onwuegbuzie and Johnson, 2006, 58).

numbers with expressions – such as a few, some, many, common, sometimes and rare – without specifying what they really mean in the research context. An example of misleading counting is using percentages to describe small samples. Acontextual counting is the case when unsubstantiated inferences are drawn from the numbers. Overcounting occurs when numbers are used just for the sake of counting that threatens developing and presenting interpretations about a target phenomenon. (Sandelowski, 2001, 236-239). All these issues can result in lower meta-inference quality.

Quantitative researchers may qualitize quantitative data via narrative profile formation, for example, by forming modal, average, holistic, comparative or normative profiles that involve constructing narrative descriptions from quantitative data. The basic threats to profile formation are over-generalizations of the observed numerical data and such representations of people (e.g., average profiles) that are unrealistic. (Onwuegbuzie and Johnson, 2006, 59)

Sixth, *paradigmatic mixing legitimation* refers to “the extent to which the researcher's epistemological, ontological, axiological, methodological and rhetorical beliefs that underlie the quantitative and qualitative approaches are successfully (a) combined or (b) blended into a usable package” (Onwuegbuzie and Johnson, 2006, 57). Combining the approaches can be problematic because of *competing dualisms of paradigmatic assumptions*: epistemological (objectivist vs. subjectivist), ontological (single reality vs. multiple realities), axiological (value free vs. value bound), methodological (deductive logic vs. inductive logic), and rhetorical (formal vs. informal writing style) assumptions. Onwuegbuzie and Johnson (2006, 59) suggest two ways of legitimation: quantitative and qualitative approaches are treated either as separate but complementary or as a continuum and compatible. Paradigmatic mixing poses threats to the legitimization of mixed research if the researcher does not make his/her paradigmatic assumptions explicit and does not conduct the research according to the stated assumptions.

Seventh, *commensurability legitimation* refers to “the extent to which the meta-inferences made reflect a mixed worldview based on the cognitive process of Gestalt switching and integration (Onwuegbuzie and Johnson, 2006, 57).” It is based on a rejection of the idea by Kuhn (1962) and others that scientific paradigms are incommensurable regarding findings, theories, language and worldviews (ibid., 2006, 59). This type of legitimation is based on the requirement that a mixed methods researcher must learn to make Gestalt switches from a

qualitative lens to a quantitative lens, going back and forth. Through this iterative process, a third well-informed viewpoint based on consideration of both qualitative and quantitative viewpoints should be created. The basic threat to this type of legitimation is *lack of cognitive and empathy training of researchers and their incapability to make Gestalt switches*.

Eighth, *multiple validities legitimation* is “the extent to which addressing legitimation of the quantitative and qualitative components of the study result from the use of quantitative, qualitative, and mixed validity types, yielding high quality meta-inferences” (Onwuegbuzie and Johnson, 2006, 57). For example, when addressing legitimation of the quantitative (qualitative) component, the relevant quantitative (qualitative) validity criteria need to be addressed and achieved and during integration of these components the relevant mixed legitimation types need to be addressed and achieved (ibid., p. 59). Multiple validities legitimation may suffer from *threats to the quality of quantitative and qualitative parts of the study*. Therefore, it is important that researchers pay attention to the internal and external validity of the quantitative part of the study and to the contextual validity, generalizability and transferability of the qualitative part of study and then use the mixed method validity criteria to combine these parts.

Ninth, *political legitimation* refers to “the extent to which the consumers of mixed methods research value the meta-inferences stemming from both the quantitative and qualitative components of a study” (Onwuegbuzie and Johnson, 2006, 57). The challenge of politics refers to the *tensions emerging as a result of combining qualitative and quantitative approaches* including, *any value or ideologically based conflicts when different quantitative and qualitative researchers collaborate in a mixed methods study, the contradictions and paradoxes when qualitative and quantitative data are compared and contrasted, and the difficulty in persuading consumers of mixed methods research to value the meta-inferences stemming from both the qualitative and quantitative findings* (ibid. 59-60; Onwuegbuzie and Leech, 2009, 107).

3.3 Conclusion

Teddlie and Tashakkori’s (2003) and Tashakkori and Teddlie’s (2008) integrative frameworks build on the previous work (see Section 2), but extend it by developing new umbrella terms, such as inference quality and inference transferability, that researchers can

alternatively use in conducting, and evaluating validity of, mixed methods research. Their contribution to the existing research is primarily conceptual in nature in creating “a bilingual nomenclature”. In our view their conceptualizations not only present inference quality as an outcome (cf., Onwuegbuzie and Johnson, 2006, 56), but also as a process that requires methodological rigor and consistency during the procedures from which the inferences emerge (i.e., during data design, collection, analysis and interpretation, see Section 3.1 above).

Onwuegbuzie and Johnson’s (2006) legitimation framework also stresses that researchers need to pay attention to the internal and external validity and credibility threats. However, their framework is more comprehensive in encompassing both the method specific and the integrative perspectives as well as several new forms of validation that are very specific for mixed methods research. According to Dellinger and Leech (2007), these two frameworks can be used to complement each other as mixed methods elements of construct validation. In our view, they can also be used to complement each other as mixed methods elements of internal validation.

4 Discussion and conclusions

The purpose of this study was to shed light on the threats to quality in mixed methods accounting research, wherein quantitative and qualitative approaches are combined in data collection, analysis and interpretation. Our analysis indicates that the quality of mixed methods accounting research can be currently evaluated from at least the following three perspectives: using the validity and reliability standards of each approach, an integrative framework (by Teddlie and Tashakkori, 2003 and Tashakkori and Teddlie, 2008), and/or a legitimation framework (by Onwuegbuzie and Johnson, 2006, see also Collins et al., 2006). As a result, a wide range of threats to validity and reliability in a range of categories was identified and synthesized based on the three perspectives. While not all the threats are likely to materialize in a single study, one should, nevertheless, take them into consideration.

Our analysis reveals how the three perspectives described in this paper are linked and supplement each other. First, the traditional validity and reliability standards of quantitative and qualitative research appear to lay an important foundation for carefully conducted mixed methods research during research design, data collection, data analysis and interpretation

stages. This is because mixed methods studies impact all these elements. Second, Teddlie and Tashakkori's (2003) and Tashakkori and Teddlie's (2008) integrative frameworks provide new vocabulary to discuss validity and credibility of mixed methods research. While their integrative framework appears useful in bridging qualitative and quantitative concepts (and paradigms), its value will depend on how broadly it becomes accepted by researchers. If it comes to be generally accepted, it can help to reduce confusion among scholars conducting and evaluating mixed methods research. Third, Onwuegbuzie and Johnson's (2006) legitimation framework is more extensive and appears currently to be the most promising in addressing those validity threats that are very specific for mixed methods research. Taken together, these three perspectives can be unified (cf., Dellinger and Leech, 2007) to form an even more comprehensive perspective of the validity and reliability threats of mixed methods studies.

Unfortunately, given the many threats to the quality of mixed methods research, our analysis indicates that the use of mixed methods research does not automatically lead to more valid and/or reliable research. Even if the validity and reliability is good during the data collection stages, there may be other issues during data analysis and interpretation. In addition, even if the quality of the quantitative and qualitative parts of the research is excellent, problems may still occur in validating the meta-inferences of mixed methods research. Consequently, mixed methods research is more complex than conducting single method studies. Mixed methods research should not be used as an end in itself. The researcher should consider thoroughly the rationale and purpose for mixing quantitative and qualitative approaches. Mixed methods research should be used only when it is likely to provide superior answers to research questions and the best methodological fit (cf. Collins et al. 2006, 69). In carefully conducted studies, it should, however, be possible to enhance the credibility and authenticity of case study findings by supportive quantitative evidence, to reduce observer bias and illusory correlations by the need to match evidence from multiple data sources and naturally test procedural reliability or reproducibility through within-study triangulation.

This study has certain theoretical and practical implications. Among the theoretical implications is this: Based on a comprehensive synthesis, this paper presents and analyzes theoretical frameworks potentially useful for scholars, students and practitioners. It focuses on both the traditional and novel areas of validity and reliability in mixed methods research. The practical implication of this paper is that it furthers our understanding of such new

developments in methodological research which may be of great importance to those conducting and/or evaluating empirical research.

This study has certain limitations. First, methodological research published in English over the past decade has been emphasized to create an approach to mixed methods accounting research. Second, although we have aimed to provide a comprehensive synthesis of the various threats to the quality of mixed methods research, our list may not be exhaustive. Third, as the literature on evaluating mixed methods research is now only emerging, we have introduced some new frameworks in this paper. These frameworks could still be studied in greater detail. Additional perspectives on the validity and reliability of mixed methods research could also be studied and developed.

References

- Abernethy, M.A., Chua, W.F., Lockett, P.F. and Selto, F.H. (1999), "Research in managerial accounting: Learning from others' experiences", *Accounting and Finance*, Vol. 39, pp. 1-27.
- Ahrens, T. & Chapman, C.S. (2006), "Doing qualitative field research in management accounting: positioning data to contribute to theory". *Accounting, Organizations and Society*, 31, pp. 819-841.
- Alkula, T., Pöntinen, S. and Ylöstalo, P. (2002), *Sosiaalitutkimuksen kvantitatiiviset menetelmät*, (Quantitative methods of social research), 4th edition. WSOY, Helsinki.
- Arbnor, I. and Bjerke, B. (1977), *Företagsekonomisk metodlära*, (Methods for business economics), Lund, Sweden: Studentlitteratur.
- Blaikie, N.W.H. (1991), A critique of the use of triangulation in social research. *Quality & Quantity*, 25, pp. 115-136.
- Brannen, J. (1992), "Combining qualitative and quantitative approaches: and overview", in Brannen, J. (Ed.), *Mixing Methods: Qualitative and Quantitative Research*. Aldershot: Avebury.
- Bryman, A. (1988), *Quantity and Quality in Social Research*. London: Unwin Hyman.
- Bryman, A. (1992), *Quantitative and qualitative research: further reflections on their integration*, Aldershot: Avebury.
- Campbell, D.T. and Stanley, J.C. (1963), *Experimental and quasi-experimental designs for research*, Chicago: Rand McNally.
- Collins, K.M.T., Onwuegbuzie, A.J. and Sutton, I.L. (2006), "A Model Incorporating the Rationale and Purpose for Conducting Mixed-Methods Research in Special Education and Beyond", *Learning Disabilities: A Contemporary Journal*, Vol. 4, No. 1, pp. 67-100.
- Creswell, J.W. (2009), *Research Design. Qualitative, Quantitative, and Mixed Methods Approaches*, Third Edition, Thousand Oaks: Sage Publications Inc.
- Creswell, J.W. and Clark, V.L.P. (2007), *Designing and Conducting Mixed Methods Research*. Thousand Oaks, CA: Sage Publications.
- Currall, S.C. and Towler, A.J. (2003), "Research methods in management and organizational research: Toward integration of qualitative and quantitative techniques", in Tashakkori,

- A. and Teddlie, C. (Eds.), *Handbook of mixed methods in social and behavioural research*, pp. 513-526. Thousand Oaks, CA: Sage.
- Dellinger, A.B. and Leech, N.L. (2007), "Toward a unified validation framework in mixed methods research", *Journal of Mixed Methods Research*, Vol. 1, No. 4, pp. 309-332.
- Douglas, J.D. (1971), *Understanding Everyday Life*, Routledge, London.
- Eisner, E.W. (1991), *The Enlightened Eye: Qualitative Inquirey and the Enhancement of Educational Practice*, New York: Macmillan.
- Eriksson, P. and Kovalainen, A. (2008), *Qualitative Methods in Business Research*, Sage, London.
- Fink, A. and Kosecoff, J. (1985), *How to conduct surveys?* Newbury Park, CA: Sage Publications.
- Golden-Bibble, K. and Locke, K. (1993), "Appealing works: An investigation of how ethnographic texts convince", *Organization Science*, Vol. 4, pp. 595-616.
- Greenwald, A.G., Pratkanis, A.R., Leippe, M.R. and Baumgardner, M.H. (1986), "Under what conditions does theory obstruct research progress", *Psychological Review*, 93, pp. 216-229.
- Grönfors, M. (1982), *Kvalitatiiviset kenttätutkimusmenetelmät*, (Qualitative field research methods), WSOY, Helsinki.
- Hair, J.F., Black, W.C., Babin, B. J., Anderson, R. E. and Tatham, R. L. (2006), *Multivariate Data Analysis*. Sixth Edition, Upper Saddle River, NJ: Pearson Prentice Hall.
- Howell, D.C. (1995), *Fundamental statistics for the behavioral sciences*, Third edition, Duxbury Press, Belmont, California.
- Hurmerinta-Peltomäki, L. and Nummela, N. (2006), "Mixed methods in international business research: a value-added perspective", *Management International Review*, 46, 4, pp. 439-459.
- Kerlinger, H. (1964), *Foundations of Behavioral Research*, Holt, Rinehart and Winston, Inc., New York.
- Koskinen, I., Alasuutari, P. and Peltonen, T. (2005), *Laadulliset menetelmät kauppatieteissä*, (Qualitative methods in business economics), Vastapaino, Jyväskylä.
- Kuhn, T.S. (1962), *The Structure of Scientific Revolutions*, University of Chicago Press, USA.
- Lather, P. (1993), "Fertile obsession: validity after poststructuralism", *Sociological Quarterly*, 34, pp. 673-693.
- Lillis, A., & Mundy, J. (2005), "Cross-Sectional Field Studies in Management Accounting Research – Closing the Gaps between Surveys and Case Studies", *Journal of Management Accounting Research*, 17, pp. 119-141.
- Lillis, A. (2006), "Reliability and validity in field study research", in Hoque, Z. (Ed.), *Methodological Issues in Accounting Research: Theories and Methods*, Piramus, London, pp. 461-475.
- Lincoln, Y.S. and Guba, E.G. (1985), *Naturalistic Inquiry*, Sage, Beverly Hills, CA.
- Lukka, K. (1988), *Budjettiharhan luominen organisaatiossa*, (Budgetary Biasing in Organizations, English summary), Academic thesis, Publications of the Turku School of Economics, Series A-5:1988, Turku.
- Lukka, K. (2005), "Approaches to Case Research in Management Accounting: The nature of empirical intervention and theory linkage", in Jönsson, S. and Mouritsen, J. (Eds.), *Accounting in Scandinavia – The Northern Lights*, Liber and Copenhagen Business School Press, Kristianstad, pp. 375-399.
- Lukka, K. and Kasanen, E. (1995), "The problem of generalizability: Anecdotes and evidence in accounting research", *Accounting, Auditing and Accountability Journal*, 8, pp. 71-90.

- Lukka, K. and Modell, S. (2010), "Validation in interpretive management accounting research", *Accounting, Organizations and Society*, Vol. 35, No. 4, pp.462-477.
- Maxwell, J.A. (1992), "Understanding and validity in qualitative research". *Harvard Educational Review*, 62, pp. 279-300.
- McKinnon, J. (1988), "Reliability and validity in field research: some strategies and tactics", *Accounting, Auditing and Accountability*, No. 1, pp. 34-54.
- Mäkinen, V. (1980), *Yrityksen toiminnan tutkimisen lähestymistavoista. Toiminta-analyyttisen tutkimusstrategian kehittelyä*, (Researching actions of business enterprises. Developing action-oriented research strategy), University of Tampere, Publications of the Department of Economics and Law, A:1—17:1980, Tampere.
- Nazari, J., Kline, T. and Herremans, I. (2006), "Conducting Survey Research in Management Accounting", in Hoque, Z. (Ed.), *Methodological Issues in Accounting Research: Theories and Methods*, Spiramus, London, pp. 427–459.
- Näsi, J. (1979), *Yrityksen suunnittelun perusteet. Käsitteellismetodologiset rakenteet ja tieteenfilosofinen tausta*, (Fundamentals of business planning. Conceptual methodological constructions and philosophical background), Tampereen Yliopisto, Yrityksen taloustieteen ja yksityisoikeuden laitoksen julkaisuja A:1-15:1979, Tampere.
- Onwuegbuzie, A. J. (2003), "Expanding the framework of internal and external validity in quantitative research". *Research in the Schools*, Vol.10, No. 1, pp. 71-90.
- Onwuegbuzie, A.J. and Johnson, R.B. (2006), "The Validity Issue in Mixed Research", *Research in the Schools*, Vol. 13, No. 1, pp. 48-63.
- Onwuegbuzie, A. J. and Leech, N. (2007), "Validity and qualitative research: an oxymoron?" *Quality & Quantity*, 41, pp. 233-249.
- Onwuegbuzie, A.J. and Leech, N.L. (2009), "Conclusion - Lessons learned for teaching mixed research: A framework for novice researchers", *International Journal of Multiple Research Approaches*, Vol. 3, No. 1, pp. 105-107.
- Patton, M. (2002), *Qualitative research and evaluation methods*, Thousand Oaks, Sage.
- Pluye, P., Grad, R.M., Levine, A. and Nicolau, B. (2009), "Understanding divergence of quantitative and qualitative data (or results) in mixed methods studies", *International Journal of Multiple Approaches*, Vol. 3, No. 1, pp. 58-72.
- Ryan, B., Scapens, R.W. and Theobald, M. (2002), *Research Method & Methodology in Finance & Accounting*, 2nd edition. Thomson, London.
- Sandelowski, M. (2001), "Real qualitative researchers don't count: The use of numbers in qualitative research", *Research in Nursing & Health*, Vol. 24, No. 3, pp. 230-240.
- Sandelowski, M. (2003), "Tables or Tableaux? The challenges of writing and reading mixed methods studies", in Tashakkori, A. and Teddlie, C. (Eds.), *Handbook of mixed methods in social and behavioural research*, pp. 321-350, Thousand Oaks, CA: Sage.
- Schwandt, T.A. (2001), *Dictionary of Qualitative Inquiry* (2nd ed.), Thousand Oaks, CA: Sage.
- Searcy, D.L., and Mentzer, J. (2003), "A framework for conducting and evaluating research", *Journal of Accounting Literature*, 22, pp. 130-167.
- Silverman, D. (2008), *Doing Qualitative Research*, Second Edition, Los Angeles: Sage Publications.
- Spearman, C. (1904), "“General intelligence” objectively determined and measured", *American Journal of Psychology*, 15, pp. 201-293.
- Tashakkori, A. and Teddlie, C. (1998), *Mixed Methodology: Combining Qualitative and Quantitative Approaches*, Thousand Oaks, CA: Sage.
- Tashakkori, A. and Teddlie, C. (2003), "The past and future of mixed methods research: from data triangulation to mixed model designs", in Tashakkori, A. and Teddlie, C. (Eds.),

- Handbook of mixed methods in social and behavioural research*, pp. 671-701, Thousand Oaks, CA: Sage.
- Teddlie, C. and Tashakkori, A. (2003), "Major issues and controversies in the use of mixed methods in the social and behavioural sciences", in Tashakkori, A. and Teddlie, C. (Eds.), *Handbook of mixed methods in social and behavioural research*, pp. 3-50, Thousand Oaks, CA: Sage.
- Tashakkori, A. and Teddlie, C. (2008), "Quality inferences in mixed methods research", in Bergman, M. (Ed.) *Advances in Mixed Methods Research: Theories and Applications*. London: Sage, pp. 101-119.
- Vaivio, J. (2008), "Qualitative management accounting research: rationale, pitfalls and potential", *Qualitative Research in Accounting and Management*, Vol. 5, No. 1, pp. 64-86.
- Van der Stede, W.A., Young, S.M. and Chen, C.X. (2007), "Doing Management Accounting Survey Research", in Chapman, C.S., Hopwood, A.G. and Shields, M.D. (Eds.), *Handbook of Management Accounting Research*, Vol. 1, Elsevier, Oxford, pp. 445-478.
- Webb, E.J., Campbell, D.T., Schwartz, R.D. and Sechrest, I. (1966), *Unobtrusive Measures: Non-Reactive Research in the Social Sciences*. Chicago: Rand McNally.
- Yin, R. (1994), *Case Study Research Design and Methods*. Second Edition. Applied Social Research Methods Series, Vol. 5, Sage Publications, Thousand Oaks.

Table 1. Examples of threats to the quality of mixed methods research

Legitimation type ¹	Examples of threats
<p><i>Sample Integration</i> The extent to which the relationship between the quantitative and qualitative sampling designs yields quality meta-inferences.</p>	Mismatch between quantitative and qualitative samples.
<p><i>Inside-Outside</i> The extent to which the researcher faithfully presents and appropriately utilizes the insider's view and the observer's views for purposes such as description and explanation.</p>	The imbalance between insider's and outsider's views (e.g. the researcher has failed to maintain a well informed and balanced perspective when collecting, analyzing, and interpreting what the whole set of qualitative and quantitative data mean).
<p><i>Weakness Minimization</i> The extent to which the weakness from one approach is compensated by the strengths from the other approach.</p>	Careless assessing of threats to and weaknesses from quantitative and qualitative parts of research Deficiencies in compensating the weaknesses by the strengths.
<p><i>Sequential</i> The extent to which one has minimized the potential problem wherein the meta-inferences could be affected by reversing the sequence of the quantitative and qualitative phases.</p>	The sequencing itself would be a threat if the results and interpretations would be different if the order of the quantitative and qualitative phases was reversed.
<p><i>Conversion</i> The extent to which the quantizing or qualifying yields quality meta-inferences.</p>	Counting pitfalls associated to verbal counting, misleading, a contextual and overcounting. Over-generalizations and representations of people that are unrealistic.
<p><i>Paradigmatic mixing</i> The extent to which the researcher's epistemological, ontological, axiological, methodological and rhetorical beliefs that underlie the quantitative and qualitative approaches are successfully (a) combined or (b) blended into a usable package.</p>	Competing dualisms of paradigmatic assumptions: the researcher does not make her/his paradigmatic assumptions explicit and does not conduct the research according to the stated assumptions.
<p><i>Commensurability</i> The extent to which the meta-inferences made reflect a mixed worldview based on the cognitive process of Gestalt switching and integration.</p>	Lack of cognitive and empathy training of researchers and their inability to make Gestalt switches.
<p><i>Multiple Validities</i> The extent to which addressing legitimation of the quantitative and qualitative components of the study result from the use of quantitative, qualitative, and mixed validity types, yielding high quality meta-inferences.</p>	Threats to the quality of quantitative and qualitative parts of the study.
<p><i>Political</i> The extent to which the consumers of mixed methods research value the meta-inferences stemming from both the quantitative and qualitative components of a study.</p>	Value or ideologically based conflicts when different quantitative and qualitative researchers collaborate in a mixed methods study. The contradictions and paradoxes when qualitative and quantitative data are compared and contrasted. The difficulty in persuading consumers of mixed methods research to value the meta-inferences stemming from both the qualitative and quantitative findings.

¹ Onwuegbuzie and Johnson (2006, 57)

Appendix 1. Examples of threats to the internal validity of quantitative research work.

Stages of research process:	Examples of threats to internal validity:
Research design	<ul style="list-style-type: none"> - <i>Insufficient knowledge of, or contradictions in the logic</i> between research question, theory, hypotheses, statistical tests and analysis.
Data collection	<ul style="list-style-type: none"> - <i>History</i>: the occurrence of events or conditions that are unrelated to the treatment but that occur during the study to a group of individuals and produce changes in the outcome measure. - <i>Maturation</i>: the possibility that a difference between the pre- and post-tests may be the result of the physical or psychological maturation of the participants rather than of differences in the independent variable. - <i>Testing</i>: can cause changes in the participant's scores obtained in the second administration as a result of having taken a pre-intervention test. - <i>Instrumentation</i>: is problematic when scores yielded from a measure lack the appropriate level of consistency or do not generate valid scores (as a result of inadequate content, criterion and/or construct validity). - <i>Statistical regression</i>: extreme scores move toward the mean on subsequent measures, when participants are selected on the basis of an extreme attribute (such as high or low performance) on some pre-intervention measure. - <i>Differential selection</i> of participants (i.e., selection bias): pertains to substantive differences between two or more of the comparison groups prior to the implementation of the intervention. - <i>Mortality</i> (subject attrition): refers to the situation where participants selected either fail to take part in the research study at all or do not participate in every phase of the research. This, in turn, may or may not produce a bias. - <i>Selection interaction effects</i>: occur when any of the above mentioned threats to internal validity interact with the differential selection of participants to produce an effect that resembles the intervention effect.¹ - <i>Implementation bias</i>: differential selection of people who apply an innovation to the intervention groups. - <i>Sample augmentation bias</i>: not all people receive the intervention for the complete duration of the study. - <i>Behavior bias</i>: a strong personal bias in favor of or against the intervention prior to the beginning of the study. - <i>Order bias</i>: the effect of the order of the intervention conditions cannot be separated from the effect of the intervention conditions. - <i>Observational bias</i>: lack of adequate sampling of behaviors. - <i>Researcher bias</i>: the researcher has a personal bias in favor of one technique over another. - <i>Matching bias</i>: variables not used to match the groups may be more related to the observed findings than is the independent variable. - <i>Treatment replication error</i>: data collected do not reflect the correct unit of analysis. - <i>Evaluation anxiety</i>: anxiety experienced when one's behavior or achievements are being evaluated. - <i>Multiple-treatment interference</i>: carryover effects from an earlier intervention makes it difficult to assess the effectiveness of a later treatment. - <i>Reactive arrangements</i> (reactivity, participant effects): changes in subjects' responses which may occur as a direct result of the awareness of participating in a research. (E.g., the presence of interviewees or equipment during a study may alter the typical responses). - <i>Treatment diffusion</i> (seepage effect): different intervention groups communicate with each other so that some of the treatment seeps out into another intervention

	<p>group causing overlapping (rather than distinctly different) interventions</p> <ul style="list-style-type: none"> - <i>Time x treatment interaction</i>: different intervention times affect participants' responses to the intervention - <i>History x treatment interaction</i>: the interventions being compared experience different events that affect group members' responses to the intervention in different ways.²
Data analysis and interpretation	<ul style="list-style-type: none"> - <i>Statistical regression; mortality; observational, researcher or matching bias; treatment replication error</i> (see above) - <i>Restricted range</i>: lacking the knowledge that virtually all parametric analyses represent the general linear model, researchers may artificially categorize variables in non-experimental design using ANOVA, although it results in relevant variance being discarded.³ - <i>Non-interaction seeking bias</i>: the presence of interactions is not assessed when testing hypotheses. - <i>Errors in statistical testing</i>: e.g., in significance testing, violated assumptions of statistical tests, multicollinearity, misspecification error and/or lack of (or incorrect) reporting of effect sizes. - <i>The use of distorted graphics</i> in checking model assumptions. - <i>Illusory correlation</i>: identification and interpretation of relationships that are not real but statistical artifacts. - <i>Causal error</i>.⁴ - <i>Confirmation bias</i>: the tendency for interpretations and conclusions based on new data to be overly consistent with preliminary hypotheses.⁵ - <i>Positive manifold</i>: a high positive correlation between different tests of cognitive ability.⁶

1 The first eight threats are based on Campbell and Stanley (1963 in Tashakkori and Teddlie, 1998, 87)

2 See Onwuegbuzie (2003) on the next 14 threats.

3 Kerlinger (1964)

4 Onwuegbuzie (2003) and Ryan et al. (2002, 123)

5 Greenwald et al. (1986)

6 Spearman (1904)

Appendix 2. Examples of threats to the contextual validity of qualitative research work.

Stages of research process:	Examples of threats to contextual validity:
Research design	<ul style="list-style-type: none"> - <i>Insufficient or biased knowledge of earlier studies and theories</i> - <i>Contradictions in the logic¹ such as a mismatch between research question and study design²</i>
Data collection	<ul style="list-style-type: none"> - <i>Observer-caused effect</i>: subjects in the field may seek to appear different from their usual selves to the researcher. - <i>Observer bias</i>: insufficient sample of behaviours or words is collected and “interpretation gaps” closed with the researcher’s own values, projections and expectations. - <i>Researcher bias</i>: personal biases or <i>a priori</i> assumptions that s/he is not able to bracket. - <i>Data access limitations</i>: the researcher is on site for a limited period of time only and his/her access to certain documents, events or people may be restricted. - <i>Complexities and limitations of the human mind</i>: subjects may consciously seek to mislead or deceive the researcher or their statements and reports are affected by natural human tendencies and fallibilities.³ - <i>Serious reactivity</i>: changes in informants’ responses that result from being excessively conscious of participating in a study.⁴
Data analysis and interpretation	<ul style="list-style-type: none"> - <i>Lack of descriptive validity</i> of settings and events - <i>Lack of interpretive validity</i> of statements about the meanings or perspectives held by participants - <i>Lack of explanatory or theoretical validity</i> about causal processes and relationships - <i>Lack of generalizability⁵</i> (e.g. lack of inability to generalize to theory) - <i>Issues in ironic validity</i> (i.e., ability to reveal co-existing opposites of the same phenomenon) - <i>Issues in paralogical legitimation</i> (ability to reveal paradoxes) - <i>Issues in rhizomatic validity</i> (ability to map and not merely describe data) - <i>Issues in voluptuous validity</i> (the extent to which the researcher’s level of interpretation exceeds her/his knowledge base stemming from the data)⁶ - <i>Confidential information⁷</i> (i.e., problems in treating confidential information in writing case reports) - <i>Poorly executed inductive analysis⁸</i> - <i>Lack of alternative interpretations of the data</i> - <i>Difficulty in interpreting the typicality of instances and findings⁹</i> - <i>All data is not analyzed and treated equally</i> regardless of whether it fits the theory¹⁰ - <i>Lack of structural corroboration</i> (utilization of multiple types of data to support or to contradict the interpretation)¹¹ - <i>Confirmation bias</i> (i.e., interpretations and conclusions based on new data are overly congruent with a priori hypotheses¹² - <i>Illusory correlation</i> (a tendency to identify a relationship when no

	<p>relationship actually prevails)</p> <ul style="list-style-type: none"> - <i>Causal error</i> (providing causal explanations and attributions for observed behaviors and attitudes without attempting to verify such interpretations) - <i>Effect size</i>¹³ (the use of effect sizes qualitzes empirical data by helping data analysts to determine the meaningfulness of behavior and words)
--	---

1. The first two threats are based on Näsi, (1979, 302)
1. See Lillis (2006, 467)
- 3 See McKinnon (1988, 37-41) on the above mentioned five issues.
- 4 Koskinen et al. (2005, 262-263)
- 5 See further Maxwell (1992) on these four forms of validity.
- 6 The above four issues are based on Lather (1993)
- 7 Ryan et al. (2002)
- 8 Koskinen et al. (2005, 262-263)
- 9 Silverman, (2008, 210-211)
- 10 Lillis (2006, 467)
- 11 Eisner (1991)
- 12 Greenwald et al. (1986)
- 13 See Onwuegbuzie and Leech (2007, 235-237).