

UNDERSTANDING DIVERGENCE OF QUANTITATIVE AND QUALITATIVE DATA (OR RESULTS) IN MIXED METHODS STUDIES.

Pierre Pluye , Roland M Grad, Alissa Levine, & Belinda Nicolau. Understanding divergence of quantitative and qualitative data (or results) in mixed methods studies. *International Journal of Multiple Research Approaches*, 2009, 3(1):58-72. DOI: 10.5172/mra.455.3.1.58

Abstract

In mixed methods studies, novice researchers need to know that qualitative and quantitative data or results sometimes diverge. However, few studies focus on this aspect of mixed methods research. The present paper aims to review the literature on divergence of qualitative and quantitative evidence, and describe examples. The prior literature reveals four strategies for taking divergence into account: reconciliation, initiation, bracketing and exclusion. Nine examples derived from empirical studies were found, and they are described. Then, a detailed example is given of how divergence was identified and explored in a pilot study of the implementation of one electronic knowledge resource on handheld computer in an academic family medicine clinic. Finally, this worked example is described in the context of a teaching exercise for novice researchers.

Key words: bracketing, divergence, initiation, mixed methods, quality appraisal, reconciliation, triangulation

INTRODUCTION

Few evaluation studies focus on the examination of divergent qualitative and quantitative evidence, and the literature lacks exemplars on how to take such divergence into account. In this paper, we explore this divergence, which may emerge from combining qualitative and

quantitative components or studies, increasingly referred to as mixed methods. A literature review on the divergence of qualitative and quantitative data or results, a worked example on how such divergence may improve evaluation research, and a teaching exercise are provided for the reader. The objective is to propose strategies for taking divergence into account, and to illustrate these strategies using a didactic exemplar for novice mixed methods researchers.

LITERATURE REVIEW OF DIVERGENCE IN MIXED METHODS STUDIES

Mixing qualitative and quantitative evidence may commonly reveal or refer to some form of divergence. However, few mixed methods studies examine details of divergence of qualitative and quantitative data or results (Greene, Caracelli & Graham 1989; Greene 2007). The lack of studies and the frequency of divergence suggest a need for exemplars promoting the integration of qualitative and quantitative data or results with respect to their divergence. In the past, attention was largely devoted to differences among quantitative results that led researchers to search for errors, or to order studies along a hierarchy of evidence (Brewer & Hunter 2006), while qualitative data contributed to reconciling differences among quantitative results (Jick 1979). Conflicting evidence between qualitative findings and quantitative results often led researchers to dismiss or ignore qualitative findings (Patton 2002). Here, as proposed by Greene (2007: 152), we ‘acknowledge and respect the value of divergence and dissonance as generative of unanticipated insights and understandings’ in mixed methods research.

What we call divergence refers to an umbrella concept. The following terms were retrieved in literature reviews and textbooks on mixed methods research regarding differences between qualitative and quantitative data or results: conflict, contradiction, discordance, discrepancy, dissonance and inconsistency. Divergence may be revealed at the stages of data collection/analysis or interpretation of results, or may occur by design (Caracelli & Greene 1993; Greene et al. 1989; Greene 2007). While Greene and collaborators associate the concepts of convergence and divergence with two different mixed methods purposes, respectively triangulation and initiation (discussed in detail below), we believe that qualitative and quantitative data or results may range from convergence to divergence whatever the mixed methods approach or design. In addition, we believe that divergence may occur in any type of mixed methods design, like those proposed by Creswell and Plano-Clark (2007), i.e. triangulation, embedded, explanatory and exploratory designs. Taking divergence into account

constitutes a key issue for triangulation designs and their variants (convergence, data transformation, validating quantitative data and multilevel).

Literature about divergence

Literature reviews and textbooks on mixed methods research suggest four strategies are used to take into account the divergence of qualitative and quantitative data or results: reconciliation, initiation, bracketing and exclusion. These four strategies require (1) an appraisal of the quality of components of mixed methods studies (or of qualitative and quantitative studies of a mixed methods research program), (2) the comparison of qualitative and quantitative data or results, and (3) the collection-analysis of additional data when needed (Moffatt et al. 2006).

Reconciliation

Reconciliation may occur when the divergence between qualitative and quantitative data or results can be interpreted in a sense-making plausible manner, which may lead researchers to re-analyze existing data (Trend 1978). Reconciliation may also suggest a new perspective or a new framework; however, it does not lead researchers to ask a new research question, or collect and analyze additional data to further examine the new perspective or framework (in contrast to initiation, below). For example, among African-American women, the Harlem Mammogram Study examined factors associated with delays in following-up abnormal mammograms (Padgett 2004; see details in Table 1). While qualitative findings exposed women's fear of abnormal tests and frustration with waiting, quantitative results indicated that women with repeated abnormal mammograms were more likely to delay follow up. To reconcile this divergence, researchers reconceived fear and frustration as factors associated with delays in follow-up, in a counterintuitive manner.

Initiation

Initiation begins with new frameworks or perspectives that emerge from conflicting evidence between qualitative findings and quantitative results, and refers to two additional steps: (1) asking new research questions, and (2) collecting and analyzing new data to further examine the fresh perspective or framework (Caracelli & Greene 1993; Greene et al. 1989; Gaber & Gaber 1997). By way of illustration, Moffatt et al. (2006) evaluated impacts of welfare rights advice on health and social outcomes among an aged population (see Table 1). While qualitative findings suggested many different impacts, quantitative results indicated no impact. Thus, qualitative and

quantitative evidence were critically appraised and re-analyzed, and additional data collection and analysis were conducted.

Bracketing

Bracketing is appropriate when qualitative and quantitative data or results are irreconcilable (Reichardt & Gollob 1987; Mark & Shotland 1987), and suggest extreme results such as best-case and worst-case scenarios. For instance, Gaber (2000) reports an evaluation of multiple needs assessments conducted for community-based organizations (see Table 1). Qualitative findings (from multiple sources of qualitative data) and quantitative results (derived from a census) were divergent, and a plausibility bracket was developed.

Exclusion

Exclusion refers to three situations (Erzberger & Kelle 2003; Morse 1991): (1) qualitative evidence contradicts or is contradicted by quantitative evidence (e.g. cross-validation), (2) the results of the mixed methods study are incomplete or inadequate, and (3) one type of data or result lacks validity. Even though exclusion is more broadly defined than the three other strategies, no examples of it were found in the literature. As discussed below, this might be associated with publication bias.

Empirical studies exploring divergence

We reviewed nine empirical studies focusing on the divergence of qualitative and quantitative data or results. This review confirms what is found in textbooks and review papers: namely, a paucity of studies focusing on such divergence. To review these empirical studies, we used a snowball technique for two reasons: (1) no specific key words exist with respect to this topic; and (2) divergence is a common term that precludes building a search strategy for retrieving a workable number of potentially relevant papers within bibliographic databases.

The review followed three steps. First, we reviewed books on mixed methods research (Brewer & Hunter 2006; Creswell & Plano-Clark 2007; Greene 2007; Tashakkori & Teddlie 2003), and review papers (Gaber & Gaber 1997; Johnson, Onwuegbuzie & Turner 2007; Morse 1991; O'Cathain, Murphy & Nicholl 2007; Reichardt & Gollob 1987; Shotland & Mark 1987). This led us to identify four papers on divergence of qualitative and quantitative data or results. Second, we searched relevant 'citees and citers' articles, i.e. articles that are either cited in these papers or

that cited these papers. We searched for citers using ISI Web of Science (all databases, no limits). We selected potentially relevant articles by reading (1) titles and abstracts (exclusion of articles on ‘quantitative methods only’, or on ‘qualitative methods only’, or on ‘conflicting paradigms’, or on ‘social conflicts’), and (2) full text of retained articles when they were available via McGill libraries (inclusion of articles focusing on divergence of qualitative and quantitative data or results). This second step led us to retain three additional empirical studies. Third, we searched ISI Web of Science (all databases, no limits) for publications containing at least one divergence-related word in their title (list derived from our reading of retained papers). We combined this search with a common strategy to identify mixed methods studies (Creswell & Plano-Clark 2007). Thus, our search strategy may be presented as follows: [Mixed method* OR multiple method* OR (qualitative AND quantitative)] AND [bracket* OR conflict* OR contrad* OR discord* OR diverg* OR disson* OR discrep* OR inconsist* OR initiation OR reconcil*]. This third step led us to retain two additional empirical studies.

In sum, nine empirical studies were retained from this snowball review and are presented in Table 1. As stated by McConney, Rudd and Ayres (2002), divergence generates tension: ‘We experienced a certain sense of disquiet because each year data divergence was apparent’ (p. 132). With respect to the proposed divergence-related strategies, five articles illustrate a reconciliation between qualitative findings and quantitative results (Cox 2003; Erzberger & Kelle 2003; McConney et al. 2002; Padget 2004; Trend 1978). Three articles illustrate the initiation of new frameworks or perspectives using additional data collection and analysis (Moffatt et al. 2006; Rossman & Wilson 1985; Waysman & Savaya 1997), while one article illustrates the bracketing strategy (Gaber 2000). The fact that we found only one paper on bracketing suggests the transfer of such a procedure to mixed methods research may be problematic since bracketing has been developed in quantitative research to estimate plausible extreme results from differences (a notion related to the confidence interval concept). We found no empirical studies on exclusion. Indeed, this strategy involves a rather radical and difficult decision-making to omit certain data and results that researchers may find hard to justify or to translate into a paper. In addition, such a paper would report negative results, which may only rarely pass through the process of peer review (publication bias).

Frequency of divergence

While few empirical studies specifically address the issue of divergence between qualitative and quantitative evidence, such divergence is not rare in the field of mixed methods research. By way of illustration, O'Cathain et al. (2007) reviewed mixed methods studies in the field of health services research in England. In 6 out of 48 retained studies (12.5%), 'the opportunity to explore seemingly discrepant findings was not taken' (p. 157).

Furthermore, examining what we call Mixed Studies Reviews (MSR), i.e. concomitant reviews of qualitative, quantitative and mixed methods studies, we found four (7%) of 59 retained reviews mentioned some form of divergence (Pluye, Gagnon, Griffiths & Johnson-Lafleur 2007b). To do so, we reviewed the literature on health-related MSR: We retrieved 2,322 references in MEDLINE, selected 149 potentially relevant references, examined corresponding full-text papers, and identified 59 MSR that were scrutinized using qualitative content thematic analysis.

We sought further examples to illustrate the frequency of divergence between qualitative and quantitative evidence in the Journal of Mixed Methods Research, a specialized new journal. We searched all issues of this journal (from January 2007 to January 2008) for papers containing at least one divergence-related word in the abstract and body of the text (see above-mentioned list). Of 23 research articles, two empirical studies (8.5%) mention some divergence of qualitative and quantitative data or results, and four review papers (17%) mention or present such divergence.

A WORKED EXAMPLE OF DIVERGENCE AND RECONCILIATION

The study

In 2001, the first two authors conducted a pilot study on the implementation of one electronic knowledge resource on handheld computer. With a convenience sample of eight Family Physicians working in an academic clinic (hereinafter FPs), we combined a questionnaire and a qualitative case study to explore perceived usefulness and use of this resource. At the time of this study, electronic knowledge resources on handheld computer were a 'new' technology. Subsequently, we found 25 additional observational studies suggesting nearly one-third of searches for clinical information in such resources may have a positive impact on physicians (Pluye, Grad, Dunikowski & Stephenson 2005).

Two participants left on maternity leave: one at the halfway point of the study and a second just after the 24-week follow-up period. Participants on maternity leave were not available for interview; however, one completed the post questionnaire. Consequently, post questionnaire data at 24 weeks was obtained from seven of eight FPs. Participants were provided with a handheld computer, two drug databases and InfoRetriever®, a search engine over seven databases (e.g. a database of synopses of research-based articles selected for validity and relevance to primary care). At the time of recruitment, no participant used either InfoRetriever® or a handheld computer in clinical practice. Participants were offered training during four consecutive weekly lunchtime meetings, and invited to a one-hour booster training session halfway through the 24-week assessment period.

The design

In line with Creswell and Plano-Clark (2007), this pilot study followed a triangulation design. Qualitative and quantitative data were collected and analyzed separately (Pluye & Grad, 2005). Qualitative findings and quantitative results were integrated at the interpretation stage. They were mixed by the first two authors, and in the context of a graduate studies course (see ‘Teaching exercise’ below). While the first author did not have personal knowledge of participants at the time of the interview, the second author trained and observed participants in their clinical work as a clinical colleague (observer participant).

Quantitative component

The self-perceived importance of information resources was measured pre (week 1 or 2) and post implementation (week 24) by this question: ‘Presently, what are your most important sources of information for solving clinical problems?’ Responses to Likert-type items were rated on a six-point scale ranging from ‘least important’ to ‘most important’ (questionnaire available on request). The relative importance of InfoRetriever® as a source of information for solving clinical problems was measured by comparing the difference in scores from pre to post against five other sources (textbooks, journals, specialist colleagues, FP colleagues and other).

Qualitative component

In-depth semi-structured interviews were conducted about seven months after the introduction of InfoRetriever®. Questions scrutinized the use of InfoRetriever®, usefulness, technical performance, impact on practice and critical incidents. Interviews took place in participants' clinical office, and varied in duration from 15 to 60 minutes (interview guide available on request). Results on technical performance, impact and critical incidents are published elsewhere (Pluye & Grad 2004; Pluye & Grad 2006).

To explore use and perceived usefulness, among other questions, participants were asked: 'How frequently did you use InfoRetriever®?', and 'How useful was it for you to have access to InfoRetriever®?'. Interviews were audio taped and then transcribed. The first two authors conducted a three stage thematic qualitative data analysis (Paillé 1996). First, extracts of transcripts were categorized according to themes from interview questions. Then, sub-themes were developed from the data. Finally, sub-themes were organized in 'process-outcome' tables with the use of InfoRetriever® ordered as an outcome (Huberman & Miles 1991). Consensus on the interpretation of data was obtained after nine sharing sessions. Transcripts were imported into computer assisted qualitative data analysis software for coding and editing reports at each stage of analysis (NVivo 1.3). For validation purposes, results were presented to two participants, who agreed with the sub-themes.

Integration of qualitative and quantitative components

The first two authors integrated qualitative findings and quantitative results using a matrix (research report available on request). Their interpretation is presented below, by the degree of change, with respect to the perceived importance of electronic knowledge resources, and the reported use and usefulness of InfoRetriever®.

Results

Quantitative data are presented in Figure 1. They revealed an increase in the importance of electronic knowledge resources for three participants (FP 5, 6 and 7), in that these resources became 'most important' for solving clinical problems. Questionnaire responses suggested no or only minor change in the importance of electronic knowledge resources for four other FPs: Participants 4 and 8 considered these resources as the second most important sources of information, while Participants 2 and 3 did not consider them to be important. Pre-study,

Participant 1 did not consider electronic knowledge resources as important for practice, and did not complete the post questionnaire.

Qualitative data are presented in Table 2 with regard to use and perceived usefulness of InfoRetriever® on handheld computer. Interviews revealed that Participant 3 did not use InfoRetriever® after the first week of the study. For Participant 4, there was little use, in that he used InfoRetriever® weekly during the first two months of the study, but then his usage decreased. Participants 5, 6 and 7 used InfoRetriever® almost weekly during the study period. Participant 8 used InfoRetriever® almost daily during the study period. Participants 1 and 2 were not interviewed.

Four sub-themes related to the usefulness of InfoRetriever® were identified (Table 2):

- (1) Five participants perceived InfoRetriever® as useful for clinical practice (participants 3, 5, 6, 7 and 8). For example, one participant stated ‘the most important issue is that you can very quickly access current information which helps to guide your decision-making.’
- (2) Of these five participants, four perceived InfoRetriever® as useful for clinical teaching (participants 5, 6, 7 and 8). As another participant said, ‘it is good to use with residents; I always try to sort of challenge a question.’
- (3) These four participants nevertheless expressed frustration with the lack of background information provided by InfoRetriever® (participants 5, 6, 7 and 8). One participant said, ‘Not everything is there; it is limited.’
- (4) For their part, two participants perceived InfoRetriever® to be less useful as compared to other electronic knowledge resources (participants 4 and 7). According to one of them, ‘I think other resources are more useful.’

Integration of qualitative findings and quantitative results

The first two authors interpreted qualitative findings and quantitative results as follows. They were *convergent* for three participants (5, 6 and 7 – see also Table 3, column 2, Researchers’ interpretation). Among these three participants, qualitative findings showed that InfoRetriever® was used almost weekly and felt to be useful. Quantitative results indicated that electronic

knowledge resources were the ‘most important’ sources of information six months after receiving InfoRetriever® on handheld computer.

Despite some inconsistencies, qualitative findings and quantitative results were *not divergent* for three other participants (1, 2 and 3 – see also Table 3, column 2, Researchers’ interpretation). Qualitative findings showed that Participant 3 did not use InfoRetriever® during the study period. Quantitative results indicated that Participant 3 perceived electronic knowledge resources as ‘the least important’ sources of information six months after receiving InfoRetriever®. Participant 3 nevertheless reported during the interview that InfoRetriever® may be considered as useful to practice (in general). This inconsistency between qualitative evidence (high usefulness in general) and mixed evidence (no use and low importance) was not interpreted as a divergence, since it may take time to apply a cognitive behavior (perceived usefulness) in practice (use). Indeed, Participant 3 bought a handheld computer after the study, and is currently using electronic knowledge resources. Participants 1 and 2 were not interviewed, and perceived electronic knowledge resources as unimportant sources of information.

Regarding two participants (4 and 8– see also Table 3, column 2, Researchers’ interpretation), qualitative findings and quantitative results were *divergent*. Qualitative findings showed Participant 8 used InfoRetriever® almost daily, and perceived it as useful both for clinical practice and clinical teaching. In contrast, qualitative findings showed Participant 4 used InfoRetriever® sparingly, and did not perceive it as useful for clinical practice or teaching. Despite these differences, quantitative results indicated no change in terms of relative importance of electronic knowledge resources as compared to other sources of information with respect to pre and post questionnaires completed by participants 8 and 4. Quantitative results also indicated that participants 8 and 4 considered electronic resources as ‘the second most important’ sources of information. Since both Participant 8 and Participant 4 were users of electronic knowledge resources before the study, these two participants showed how multiple sources of information compete for usage in practice. When InfoRetriever® was introduced at the start of the study, Participant 8 found a way to integrate this new electronic knowledge resource in practice, while Participant 4 already had a full toolbox. This example of reconciliation of a divergence between qualitative findings and quantitative results suggests how qualitative findings may nuance and add complexity to quantitative results.

However, the qualitative findings and quantitative results presented here must be handled with caution. Three methodological issues must be recognized with respect to the pre-post survey: (1) the inability to establish a causal relationship between the intervention and self-reported change; (2) the possibility of social desirability bias, that is, the influence of researchers' expectations on FPs' responses to questions; (3) the selection bias arising from a small convenience sample, and the lack of follow-up with two participants. In addition, data on actual use of InfoRetriever® could not be tracked in 2001; therefore, there is the possibility of recall bias in the answers to interview questions on software use. This limitation has been addressed in more recent research in which usage tracking is combined with *Computerized Ecological Momentary Assessment*, to systematically capture quantitative data on cognitive impact around the moment of use, with the *Critical Incident Technique* (Grad et al. 2005, 2008; Pluye et al. 2007a).

Given these limitations, our pilot study work is presented here to illustrate how divergence between qualitative findings and quantitative results may be reconciled, and then used to improve evaluation research. This pilot study might also contribute to a better understanding of how electronic knowledge resources on handheld computer may fit into the spectrum of information resources for clinical practice and medical education. While electronic knowledge resources on handheld computer offer rapid access to much clinical information, further research is needed to evaluate their effect on knowledge use, clinical decision-making and patient health outcomes.

A TEACHING EXERCISE FOR NOVICE MIXED METHODS RESEARCHERS

Using the above worked example, we conducted a teaching exercise in an 'Applied Mixed Methods in Health Research' course with four teams of two students. The co-instructors could be described as specializing in mixed methods (first author), qualitative methods (third author) and quantitative methods (fourth author). Those enrolled in the course are graduate students in public health and novice mixed methods researchers with backgrounds in dentistry, education and nursing.

In line with the matrix of 'result possibilities' proposed by McConney et al. (2002), co-instructors and students completed the matrix presented in Table 3. In this matrix, qualitative findings and quantitative results were summarized with participants classified in four categories

(high, moderate, low and unknown) by two criteria (1) their perceived importance of electronic knowledge resources, which fits with the degree of change with respect to this importance, and (2) reported use and usefulness of InfoRetriever® on handheld computer.

Co-instructors and students attended an oral presentation on the pilot study, which did not describe the integration of qualitative findings and quantitative results. They then examined qualitative findings and quantitative results, and also read the corresponding extracts of interviews (exercise booklet available on request). They were asked (1) to critically scrutinize the proposed matrix of ‘result possibilities’, (2) to complete the third column of the matrix by integrating qualitative findings and quantitative results, and (3) to write comments on their response sheet that would justify their interpretation of the integration of qualitative and quantitative evidence, and highlight the limitations of the exercise and the pilot study.

Co-instructors (hereinafter teachers I and II) and students (hereinafter teams A, B, C and D) agreed that qualitative findings and quantitative results focused on different and complementary aspects of the implementation of electronic knowledge resources on handheld computer.

Regarding the divergence between qualitative and quantitative evidence, results of the exercise are summarized as follows: all agreed on the convergence between qualitative and quantitative evidence concerning participants 3, 5, 6 and 7; all but Teacher I excluded participants 1 and 2 from their analysis since qualitative data were missing; disagreements regarding participants 4 and 8 illustrated two strategies for addressing the divergence between qualitative and quantitative evidence (exclusion and reconciliation). The exercise did not lead teachers or students to propose ‘bracketing’ or ‘initiation’ strategies.

With respect to Participant 4, Team B, Team C and Teacher II recognized the divergence between qualitative and quantitative evidence, and agreed with researchers’ interpretations. For example, Team B wrote the following comment: The participant ‘rated ‘moderate’ importance [quantitative data], but rarely used the tool [qualitative data].’ In contrast, Team A challenged the validity of qualitative data, and proposed to exclude it from the mixed methods analysis. Team A commented: ‘there are contradictions in the interview.’ For their part, Team D and Teacher I tried to reconcile the divergence between qualitative and quantitative evidence by challenging researchers’ interpretation of qualitative data (moderate vs. low use of InfoRetriever®). Team D commented: ‘We noticed that the qualitative data reported usage that can be considered as

‘moderate’, and we would then have concordant [mixed methods] results.’ Teacher I commented: ‘although the results regarding Participant 4 seem to be contradictory, I would argue that this participant should be classified as moderate [qualitative evidence]. In the excerpts (qualitative data) he/she says that the decrease in use of secondary databases was due to being away from work.’

With respect to Participant 8, Team B and teachers I and II recognized the divergence between qualitative and quantitative evidence, and agreed with researchers’ interpretations. For example, Team B wrote the following comment: ‘[the participant] rated ‘moderate’ importance [quantitative data], but actually used the tool almost daily [qualitative data].’ In contrast, teams A, C and D tried to reconcile the divergence between qualitative and quantitative evidence by challenging researchers’ interpretation of qualitative data (moderate vs. high use of InfoRetriever®). As stated by Team A in their comment, ‘[Based on qualitative data, we are] not aware of the daily use.’

DISCUSSION OF DIVERGENCE

This paper outlines one of the important features and tensions in the developing field of mixed methods, specifically in the area of evaluation research. We critically reviewed the literature focusing on the divergence between qualitative and quantitative evidence. This review led us to propose four strategies to take divergence into account: reconciliation, initiation, bracketing and exclusion. While the literature suggests divergence is not a rare phenomenon in mixed methods studies, we found only nine empirical studies to illustrate these strategies. Then, for novice mixed methods researchers, we present a worked example on divergence, and a teaching exercise. As suggested by the exercise, the divergence between qualitative and quantitative evidence is a complex issue for at least two reasons: it might lead to forced reconciliation or inappropriate exclusion, and it may not be easily recognized or acknowledged. In line with Devereux (1967), divergence-related tensions experienced by researchers may generate ‘blind spots’ that can lead to ignoring divergence.

In line with Hacking (1999), mixed methods may be conceived as a ‘mixed kind’ of methods that emerge by ‘looping effects’ between logical empiricism and constructivism, which are usually presented as competing paradigms or ‘worldviews’ in the literature (Creswell & Plano-Clark

2007; Greene 2007; Johnson et al. 2007; Pluye et al. in press; Tashakkori & Teddlie 2003). Constructivism is associated with idealism, relativism and (inter)subjectivity, while logical empiricism is associated with materialism, realism and objectivity. Constructivism is most frequently associated with inductive qualitative studies, and logical empiricism is most frequently associated with deductive quantitative studies. Indeed, what is conceptualized as 'mixed evidence' derived from 'looping effects' between qualitative and quantitative evidence, has been described in terms of iterative 'spiraling' among qualitative and quantitative data, which adds 'depth of understanding' (Caracelli & Greene 1993 p. 202). Recently, Mendlinger and Cwikel (2008) mobilized a biomedical metaphor (double helix) to represent the 'spiraling technique', i.e. an 'iterative process of going back and forth between qualitative and quantitative methods', between induction and deduction (p. 290). With respect to Participant 3, mixed evidence refers to the absence of use of InfoRetriever® (qualitative evidence) and the low importance of electronic knowledge resources (quantitative evidence) for example.

CONCLUSION

The present paper may help novice mixed methods researchers to better understand the combination of qualitative and quantitative evidence using a didactic exercise on divergence in the form of a concrete and simple worked example designed for teaching. In our experience, this didactic exercise contributed to better understanding the potential richness of mixing qualitative and quantitative evidence. The course session with this exercise was highly rated (on average 4.6 out of 5 on the weekly course evaluation), with students unanimously reporting that the exercise was what they appreciated the most. Conceptually, this exercise is relevant to more seasoned mixed methods researchers, as few evaluation studies report on how to deal with divergence between qualitative and quantitative evidence.

Acknowledgements

The Frederick and Helen Weinstein Kahn Memorial Endowment supported this study. Pierre Pluye and Belinda Nicolau hold a New Investigator Award from the Canadian Institutes of Health Research (CIHR). Pierre Pluye and Roland Grad are supported by CIHR, the 'Fonds de recherche en santé du Québec', the Department of Family Medicine (McGill University), and the Herzl Family Practice Center. A substantive contribution to the training of participants made by Dr. Howard Goldstein is gratefully acknowledged. We would also like to acknowledge the help

of Tara Bambrick, research professional, and of students in the 'Applied Mixed Methods for Health Research' course at McGill University, for their participation in the data mixing exercise.

REFERENCES

- Brewer J and Hunter A (2006) *Foundations of multimethod research*. Thousand Oaks, Sage.
- Caracelli VJ and Greene JC (1993) Data analysis strategies for mixed-method evaluation designs, *Educational Evaluation and Policy Analysis* 15(2): 195-207.
- Cox K (2003) Assessing the quality of life of patients in phase I and II anti-cancer drug trials: interviews versus questionnaires *Social Science & Medicine* 56(5): 921-934.
- Creswell JW and Plano Clark VL (2007) *Designing and conducting Mixed Methods Research* Thousand Oaks, Sage.
- Devereux G (1967) *De l'angoisse à la méthode* Paris, Flammarion.
- Erzberger C and Kelle U (2003) Making inferences in mixed methods: the rules of integration. In Tashakkori A and Teddlie C (Eds) *Handbook of mixed methods in social and behavioral research*. Thousand Oaks, Sage pp 457-488.
- Gaber J (2000) Meta-needs assessment *Evaluation and Program Planning* 23: 139-147.
- Gaber J and Gaber SL (1997) Utilizing mixed-method research designs in planning: the case of 14th street New York City *Journal of Planning Education and Research* 17: 95-103.
- Grad RM, Pluye P, Meng Y, Segal B and Tamblyn R (2005) Assessing the impact of clinical information retrieval technology in a family practice residency *Journal of Evaluation in Clinical Practice* 11(6): 576-586.
- Grad RM, Pluye P, Mercer J, Marlow B, Beauchamp ME, Shulha M, Johnson-Lafleur J and Wood-Dauphinee S (2008) Gauging the impact of email alerts in general practice. *Journal of the American Medical Informatics Association* 15(2): 240-245.
- Greene JC (2007) *Mixed methods in social inquiry* San Francisco, Jossey Bass.
- Greene JC, Caracelli VJ and Graham WF (1989) Toward a conceptual framework for mixed-method evaluation designs *Educational Evaluation and Policy Analysis* 11(3): 255-274.
- Hacking I (1999) *The social construction of what?* Cambridge, Harvard University Press.
- Huberman MA and Miles MB (1991) *Analyse des données qualitatives: recueil de nouvelles méthodes* Bruxelles, De Boeck Wesmael.

- Jick TD (1979) Mixing qualitative and quantitative methods: triangulation in action *Administrative Science Quarterly* 24(4): 602-611.
- Johnson RB, Onwuegbuzie AJ and Turner LA (2007) Toward a definition of mixed methods research *Journal of Mixed Methods Research* 1(2): 112-133.
- Mark MM and Shotland RL (1987) Alternative models for the use of multiple methods *New Directions for Program Evaluation* 35: 95-100.
- McConney A, Rudd A and Ayres R (2002) Getting to the bottom line: a method for synthesizing findings within mixed-method program evaluations *American Journal of Evaluation* 23: 121-140.
- Mendlinger S and Cwikel J (2008) Spiraling between qualitative and quantitative data on women's health behaviors: a double helix model for mixed methods *Qualitative Health Research* 18: 280-293.
- Moffatt S, White M, Mackintosh J and Howel D (2006) Using quantitative and qualitative data in health services research - what happens when mixed method findings conflict? *BMC Health Services Research* 6, 28 doi:10.1186/1472-6963-1186-1128.
- Morse JM (1991) Approaches to qualitative-quantitative methodological triangulation *Nursing Research* 40: 120-123.
- O'Cathain A, Murphy E and Nicholl J (2007) Integration and publications as indicators of 'yield' *Journal of Mixed Methods Studies* 1: 147-163.
- Padget D (2004) *Mixed methods, serendipity, and concatenation: the qualitative research experience* Belmont: Wadsworth pp 273-288.
- Paillé P (1996) De l'analyse qualitative en général et de l'analyse thématique en particulier *Recherches Qualitatives* 15: 179-194.
- Patton M (2002) *Qualitative research and evaluation methods* Thousand Oaks, Sage.
- Pluye P and Grad RM (2004) How information retrieval technology may impact on physician practice: an organisational case study in family medicine *Journal of Evaluation in Clinical Practice* 10(3): 413-430.
- Pluye P and Grad RM (2005) A mixed methods study diary *The AMIA Student Working Group News* 2(3): 3-4.

- Pluye P, Grad RM, Dunikowski L and Stephenson R (2005) The impact of clinical information retrieval technology on physicians: a literature review of quantitative, qualitative and mixed-method studies *International Journal of Medical Informatics* 74(9): 745-768.
- Pluye P and Grad R (2006). Cognitive impact assessment of electronic knowledge resources: a mixed methods evaluation study of a handheld prototype. *AMIA 2006 Symposium Proceedings* pp. 634-638.
- Pluye P, Grad RM, Mysore N, Knaapen L, Johnson-Lafleur J and Dawes M (2007a) Systematically assessing the situational relevance of electronic knowledge resources: a mixed methods study *Journal of the American Medical Informatics Association* 14(5): 616-625.
- Pluye P, Gagnon MP, Griffiths F and Johnson-Lafleur J (2007b) *Mixed studies reviews in health sciences* Mixed Methods Conference, Cambridge UK.
- Pluye P, Nadeau L, Gagnon MP, Grad RM, Johnson-Lafleur J and Griffiths F (in press). 'Les méthodes mixtes pour l'évaluation des programmes' in Ridde V & Dagenais C, *Théories et pratiques en évaluation de programme: manuel d'enseignement*, Montréal, Presses de l'Université de Montréal.
- Reichardt CS and Gollob H (1987) Taking uncertainty into account when estimating effects, *New Directions for Program Evaluation* 35: 7-22.
- Rossmann GB and Wilson BL (1985) Numbers and words: combining quantitative and qualitative methods in a single large-scale evaluation study *Evaluation Review* 9: 627-643.
- Shotland RL & Mark MM (1987) Improving inferences from multiple methods, *New Directions for Program Evaluation* 35: 77-94.
- Tashakkori A and Teddlie C (2003) *Handbook of mixed methods in social and behavioral research*, Thousand Oaks, Sage.
- Trend M (1978) Reconciliation of qualitative and quantitative-analyses - Case-study, *Human Organization* 37(4): 345-354.
- Waysman M and Savaya R (1997) Mixed method evaluation: a case study, *Evaluation Practice* 18(3): 227-237.

Table 1. Nine empirical studies on divergence of qualitative and quantitative evidence

First author(s) Date	Research objective & methods	Divergence of qualitative and quantitative data or results	Strategy to take divergence into account
PART I		RECONCILIATION	
Cox (2003)	This study examined ‘patients’ experiences of phase I and II anti-cancer drug trial participation’ (p. 921). ‘Fifty-five patients consented to be interviewed (and filled in two quality of life questionnaires) about their trial experience’ (p. 923).	‘This paper demonstrates how different methods of collecting data (...) can lead to alternative conclusions (...). Data obtained from the quality of life questionnaires interestingly revealed no statistically significant differences in any of the scores over time while in-depth interviews uncovered something of the psychological, emotional and social impact of taking part in a clinical trial from the perspective of the patient’ (p. 921). ‘The patients seemed to be minimizing their problems on the quality of life assessment forms’ (p. 931).	Reconciliation (plausible interpretation): ‘One reason for the mismatch of quality of life scores with the interview data could be that the questionnaires asked patients to rate how they have been feeling over the last week, whereas the interviews allowed for a much broader coverage of time and also for a deeper description of the issue being discussed. Another reason could be that ratings were made before the interview and were based on what came to mind in that short rating interval. Ratings are often more accurate when made after a reflected or communicated exploration of the issue’ (p. 931).
Erzberger & Kelle (2003)	The Special Collaborative Center 186 ‘focuses on the relationship among social structures, social change, life course patterns, and individual biographies during the modernization process in Germany’ (p. 467). A study on the transition between education and job in the former East-Germany combined a quantitative survey (N=551) of academics with qualitative interview of a sub-sample (N=21).	According to the quantitative data, ‘the system of state control over individual career paths and trajectories worked very well (...). The qualitative data provided a totally different picture of the transition process (...), and revealed that individual actors were indeed able to influence their individual careers to a remarkable extent if they were creative enough (...). The qualitative data revealed that the simple and straightforward picture produced by the quantitative data was incorrect and misleading’	Reconciliation (new conceptual framework): As compared to quantitative results, the qualitative findings were seen as a significant ‘counterevidence’. This divergence was ‘reconciled by a theoretical redefinition of the sociological function’ of bureaucracies and individual behaviors (p. 478). The employment bureau ‘was no longer seen as a distribution agency; instead it was seen as an institution for the legitimization of individual action’ (p. 478). There was no additional data collection and analysis.

(p. 477).

McConney et al. (2002)	The objective of this evaluation research was 'to determine the effectiveness of the SOS Model School program as implemented in a pilot program in three elementary schools' (p. 129). This evaluation combined a quasi-experimental design with four qualitative methods: school site visit, case study, focus group interviews, and open-ended teacher survey. In terms of quantitative assessment, student performance was collected from state assessments, and data collection instruments were developed to collect school-based data on students.	'Data (findings) were consistently positive or negative regarding program effectiveness depending on the type or source of data examined. The large-scale standardized state assessment data, and the school-wide quantitative data both provided consistently neutral or negative findings on program effectiveness. On the other hand, the site interview, focus group, and case study (primarily qualitative) data provided consistently neutral-to-positive messages about program effectiveness' (p. 133).	Reconciliation (data re-analysis): 'This dilemma led us to seek out and subsequently develop a method of defensibly synthesizing findings from mixed-method evaluations' (p. 133). For each school, each type of data (findings) is rated in terms of effectiveness using a score between -150 to +150, and scores are synthesized into an 'overall program effectiveness' rating.
Padgett (2004)	'The Harlem Mammogram Study (...) was funded (...) to examine factors that influence delay in response to abnormal mammogram among African-American women living in New York City' (p. 275). It combined quantitative structured questionnaire (N=212), and qualitative interviews with a sub-sample of women (N=45).	'Qualitative [data] analysis revealed the fear and frustrations of enduring painful tests and waiting for the results in the women's own words' (p. 276). Intrigued, we returned to the quantitative data and found that women who had a history of repeated abnormal mammograms (29% of the sample) were 2.5 times more likely to delay follow up' (p. 277).	Reconciliation (new perspective): 'If we had not heard a possible explanation for this [follow up delayed] in the qualitative interviews [fears and frustrations], this odds ratio would have seemed counterintuitive. After all, such women are assumed to be at higher risk and thus more compliant with recommendations' (p. 277). There was no additional data collection and analysis.
Trend (1978)	Social experiments tested 'the concept of using direct cash allowance payments to help low-income families obtain decent housing on the open market' (p. 345). Experimental methods and participant observation were combined.	Qualitative data 'depicted staff overwork and the heavy-handed interference of a contracting agency', while quantitative data indicated that managers of this agency achieved results, and cannot be dismissed as incompetent or inappropriate (p. 349)	Reconciliation (data re-analysis): The final interpretation 'The solution was to overturn the existing explanations by offering a third. This required no brilliance, some ingenuity, and a good amount of tenacity' (p. 352).
PART II		INITIATION	
Moffatt et al. (2006)	A pilot study evaluated 'whether welfare rights advice has an impact on health and social outcomes among a population	'Separate analysis of the quantitative and qualitative data revealed discrepant findings. The quantitative	Initiation (additional data collection and analysis): 'Six ways of further exploring these data were considered: (i) treating the

	aged 60 and over. Qualitative and quantitative data were collected contemporaneously. Quantitative data were collected from 126 men and women (...) within a randomized controlled trial. (...) Qualitative data were collected from a sub-sample of 25 participants purposively selected to examine the perceived impact of welfare rights advice' (p. 1).	data showed little evidence of significant differences of a size that would be of practical or clinical interest, suggesting that the intervention had no impact on these outcome measures. The qualitative data suggested wide-ranging impacts, indicating that the intervention had a positive effect' (p. 1).	methods as fundamentally different; (ii) exploring the methodological rigour of each component; (iii) exploring dataset comparability; (iv) collecting further data and making further comparisons; (v) exploring the process of the intervention; and (vi) exploring whether the outcomes of the two components match. Conclusion: The study demonstrates how using mixed methods can lead to different and sometimes conflicting accounts and, using this six step approach, how such discrepancies can be harnessed to interrogate each dataset more fully' (p. 1).
Rossman & Wilson (1985)	The aim of the study was to learn 'about the perceived usefulness of [Regional Educational Service Agencies (RESAs)] by local school people' (p. 634). A first survey of school administrators was combined with extreme case (defined from the survey) qualitative studies based on interviews with teachers, school and district administrators, and RESAs' staff, and with a second survey involving RESAs' staff.	Quantitative results identified extreme cases (survey #1), and qualitative findings 'revealed surprising variations in the ability of school administrators to select outside agencies for new information' (p. 638). Then, additional quantitative results (survey #2) showed that school 'administrators received considerably more services from their RESA than did teachers' (p. 637).	Initiation (additional data collection and analysis): The second survey data 'elaborated the interview data, providing a richness of detail about differences between teachers and administrators that the qualitative data alone could not provide' (p. 637).
Waysman & Savaya (1997)	This research evaluation aims 'to look back and plan ahead' SHATIL's activities, 'based on feedback from client organizations' (p. 2). 'SHATIL is a nonprofit Israeli agency that provides direct assistance to other nonprofit community-based organizations' (p. 2). The first qualitative phase consisted of focus groups and personal interviews with SHATIL staff and clients (general issues). Then, a survey questionnaire was conducted on specific issues derived from the qualitative data. A second qualitative phase	'Some of the focus group participants expressed feelings of being patronized by SHATIL staff, who at times had conveyed to them the message 'We know what's good for you better than you do.' Findings from the quantitative measure, however, revealed that only a small minority of clients (15%) shared this sentiment. If we had included only the qualitative component in the study design, we might have overestimated the prevalence of this finding' (p. 4).	Initiation (additional data collection and analysis): 'This inconsistency forced us to reconcile these apparent contradictions by raising a new research question: can we characterize the organizations for whom this issue is of concern?' (p. 5). The additional qualitative data collection-analysis 'revealed that the problem had been raised primarily by minority organizations. In response to this finding, SHATIL initiated a search for ways to increase the cultural sensitivity of service delivery' (p. 5).

consisted of focus groups with clients ‘on one particular issue (sources of satisfaction and dissatisfaction)’ (p. 2).

PART III		BRACKETING	
Gaber (2000)	A meta-needs assessment (evaluation of multiple needs assessments) was conducted to help community-based organizations ‘in their development of state-wide needs assessments’ (p. 142). It combined census data, focus groups and documents: ‘74 documents were received, cataloged and analyzed’ (p. 143). In these documents data were qualitative and quantitative.	‘Divergence between the census data and the needs assessment analysis did not assume that the identified need was less significant than those needs when the two data slices converged. Instead, the divergence of data highlighted that more research was needed to flesh-out what was going on for a particular need’ (p. 144).	Bracketing: When ‘the census data diverged from the needs assessment analysis, either a new explanation was determined (initiation) or a plausibility bracket was developed’ (p. 144). Initiation is only mentioned as potentially needed (no additional data collection and analysis): ‘For example, if the needs assessments identified a growing need, but the census data showed that the population in need was decreasing, a tentative hypothesis could be that the particular population in need may be experiencing further social or geographic isolation which warrants more research’ (p. 144).

Table 2. Qualitative findings: Use and usefulness of InfoRetriever® on handheld computer

	Participan t 3	Participan t 4	Participan t 5	Participan t 6	Participan t 7	Participan t 8
<i>Self-reported Use</i>	<i>Almost no use</i>	<i>Little use</i>	<i>Almost weekly</i>	<i>Almost weekly</i>	<i>Almost weekly</i>	<i>Almost daily</i>
Perceived usefulness of InfoRetriever ® for clinical practice	Useful	Not useful	Useful	Useful	Useful	Useful
Perceived usefulness of InfoRetriever ® for clinical teaching		Not useful	Useful	Useful	Useful	Useful
Need for background information not found in InfoRetriever ®			Yes	Yes	Yes	Yes
InfoRetriever ® less useful than other databases		Yes			Yes	

Note: Participants 1 and 2 were on maternity leave and not available for interview.

Table 3. Teaching exercise: Matrix of ‘result possibilities’ proposed to students

Implementation of electronic knowledge resources on handheld computer	Researchers’ interpretation of quantitative data (baseline & post)	Researchers’ interpretation of qualitative data (interviews):	Mixing qualitative and quantitative data
	<ul style="list-style-type: none"> <i>Importance of databases</i> 	<ul style="list-style-type: none"> <i>Reported use</i> <i>Perceived usefulness</i> 	
	Appendix 2	Appendix 3	
High	Participant 5 Participant 6 Participant 7	Participant 5 Participant 6 Participant 7 Participant 8	?
Moderate	Participant 4 Participant 8		?
Low	Participant 1 Participant 2 Participant 3	Participant 3 Participant 4	?
Unknown		Participant 1 Participant 2	?

Instructions

- Step 1. Read and discuss the qualitative and quantitative data and results (appendices*)
- Step 2: Complete the last column
- Step 3: Outline your interpretation, potential limitations and conclusion

*Note: Appendices were (1) abstract and methods; (2) quantitative results; and (3) qualitative findings with corresponding extracts of interviews.

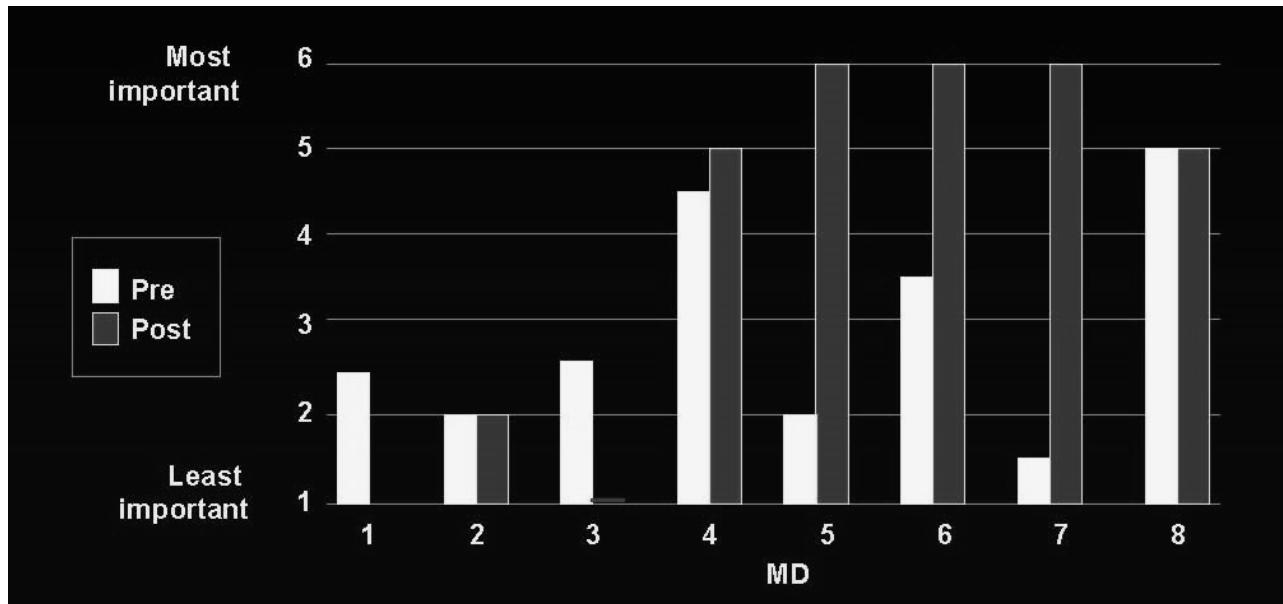


Figure 1: Quantitative results: The relative importance of electronic knowledge resources (including InfoRetriever®)

- Vertical axis: Relative importance of electronic knowledge resources compared to other sources of information
- Horizontal axis: MD - Family physician (participant number)