

TRIANGULATION IN A QUASI-EXPERIMENTAL MIXED METHODS STUDY: CONVERGENCE, DIVERGENCE, AND COMPLEMENTARITY OF RESULTS

J.P. Nzabonimpa,

Monitoring and Evaluation Expert, The African Capacity Building Foundation, Harare,
Zimbabwe

H.E. Prozesky,

Centre for Research on Evaluation, Science and Technology, Stellenbosch University,
Private Bag X1, Matieland 7602, South Africa; and DST-NRF Centre of Excellence in
Scientometrics and STI Policy, Stellenbosch University, Private Bag X1, Matieland 7602,
South Africa

ABSTRACT: *Mixed methods research is empirically known to bring about converging, diverging, and/or complementary results and this remains a methodological illumination to embrace and a challenge to overcome. The existing literature deals with the review of relevant research results in mixed methods studies with limited testing whether and how results, different or similar, originate from the integrated data analysis and how additional non-quantitized or non-qualitized data are used. After matching phase-one with phase-two respondents, we merged quantitized essay and interview data (qualitative methods strand) with survey questionnaire data (quantitative methods strand), performed mixed and non-mixed data analyses to statistically compare and enrich results from both methods strands. The aim of this article is to move beyond the mixing of methods and data and triangulate results. We statistically test, empirically analyse, and offer guidance on the practical dimensions to consider while mixing qualitative and quantitative methods.*

KEYWORDS: mixed methods research, triangulation, data integration, convergence, divergence, complementarity

INTRODUCTION

In mixed methods research (MMR), more than one methodological paradigm is used to gain a deeper and broader understanding of the issues under than one method would uncover. With MMR, evaluators seek to triangulate and consequently enhance the validity of the findings by ensuring all potential methods and sources of data are used to reach valid conclusions (Flick, 2017; Gibson, 2017; Creswell, J.W. & Plano Clark, 2011; Hesse-Biber, 2010). In MMR, however, few researchers have statistically tested and empirically ascertained how results emanating from qualitative and quantitative methods converge with, diverge from, or complement each other. This article is meant to contribute to the existing gaps surrounding triangulation in MMR with the main focus laid on convergence, divergence and complementarity as the three possible triangulation outcomes (Morgan, 2019).

Survey mode (Wallace *et al.*, 2018) and data collection instrument, coupled with researcher, participant, and contextual effects, can rarely go without shaping survey responses. Resorting to triangulation in research is acknowledging, according to Caillaud, et al., (2019), that research is not a mere description but a transformative account, serving as an emporium for different knowledge systems by moving from one method to another to avoid imposition of one group's reality to another. Our interest, and the focus of this paper, is the triangulation process, and the results thereof, in mixed methods.

We contend that missing are statistical tests to examine the significance of divergence or convergence in mixed methods analysis and show how additional insights from either methods strand are used for complementarity. We believe and concur with Jick (1979) and Morgan (2019) that convergence of data from two methods strands can validate results while divergence can come in as an opportunity for supplementary explanations. Despite some insights generated by some authors (Pluye, Grad, Levine, & Nicolau, 2009; Pluye, Bengoechea, Granikov, Kaur, & Tang, 2018), empirical evidence is missing, and few mixed methods studies are statistically analysed specifically to look into the extent to which divergence and/or convergence is engendered and handled and how mixed data and results pave the way for complementarity in mixed methods.

TRIANGULATION AND METHODOLOGICAL UNDERPINNINGS

Triangulation is not new (Kern, 2018; Flick, 2017); it traces its origin from the military and navigation contexts whereby an object's exact position is located by using at least two known points the tips of which form a triangle (Morgan, 2019; Caillaud, Doumergue, Préau, Haas, & Kalampalikis, 2019). Triangulation, defined by Denzin as "the combination of methodologies in the study of the same phenomenon" can be traced back to the 1960s when Campbell and Fisk introduced the use of multiple methods in a single inquiry for validation purposes (Johnson, Onwuegbuzie, and Turner, 2007; Caillaud, et al., 2019). As such, triangulation is confined to positing that results from one method are valid if they are correlated with results obtained using another method (Caillaud, et al., 2019). As stated in Johnson et al. (2007) and at varying degrees echoed by Moon (2019), Archibald (2016), and Kern (2018), Denzin proposed in 1966 that researchers can triangulate data, investigators, theory, and methods. In addition to Denzin's typologies, other forms of triangulation focus on paradigms, space, time, population, and so forth, with the aim to enhance validity and reliability in research (Babbie & Mouton, 2001; Cohen, Manion, & Morrison, 2000; Kelle, 2006). During triangulation process, results may diverge from, converge with, and/or complement each other.

Of all types of triangulation, two categories are of interest for this article: data triangulation (generated from closed-ended survey questionnaire on the one hand, and qualitative semi-structured interviews and biographic narrative educational essays on the other hand) and methodological triangulation (with conversion of qualitative data into quantitative and merging and using data from both methods strands for analysis with and without statistical tests). With triangulation in MMR, similar constructs are measured through qualitative and quantitative methods and the evaluator compares data to cross-validate study results. In addition to convergence or cross-validation, it can be argued that triangulation provide a potential source

of additional information and evidence in the form of complementarity and divergence of results (Braun et al., 2020; Morgan, 2019; Caillaud, et al., 2019; Flick, 2017; Kern, 2018). Increasingly popular, triangulation is a methodological process (not an end-result) bringing together quantitative and qualitative methods and data to determine whether convergence, divergence and/or complementarity (the end results) are achieved through navigating research settings requiring more than one method.

With equal priority assigned to qualitative and quantitative methods strands in some MMR designs, the conceptual dimension of triangulation resides, on the one hand, at the design stage whereby the researcher identifies constructs or concepts to measure across qualitative and quantitative methods strands. On the other hand, the operational dimension of triangulation is the process by which the researcher breaks down the constructs into measurable items, bearing in mind and ensuring each itemized construct has its equivalent item in both methods strands and in each data collection instrument. Worth noting is that mixing qualitative and quantitative methods can take place at different stages in the process but, as research scholars state, data integration remains the hallmark of MMR (Moseholm & Fetter, 2017; Fetter & Freshwater, 2015; Guetterman, Fetter, & Creswell, 2015; Alexander, Eppler, & Comi, 2020).

Whether results converge or diverge in MMR, the best way to investigate these issues is through the lens of repeated-measures design to find out how responses change over time as a result or not of switching to a qualitative or quantitative methods strand. By conducting a repeated-measures mixed methods study, or a fully longitudinal mixed methods design according to Plano Clark et al. (2015), we aim at investigating whether and how survey questionnaire on the one hand (i.e. quantitative methods strand) and interviews and essays (i.e. qualitative methods strand) on the other produce similar and/or significantly diverging results when used to measure the same concepts. As Creswell and Plano-Clark (2011) recommend, we corroborate, directly compare, and relate two sets of data and results from the same individuals who participated both in the quantitative and qualitative methods strands. We ask, to what extent do interview and essay results converge with or diverge from survey questionnaire results over a time-spaced mixed methods study? In this paper, besides what Nzabonimpa (2018) calls un-quantitizable and un-qualitizable data and their all-importance in complementing mixed results, we show, using statistical tests and narrative data, how triangulation of data and methods is applied in practice

Divergence, convergence, and complementarity in MMR

Methodological strengths or weaknesses can lead to divergence (Pluye, *et al.* 2009) or convergence of results when both qualitative and quantitative methods are mixed in a single study for cross-validation purposes. Divergence is defined as the lack of harmony between data and results, what other authors refer to as conflict, contradiction, discordance, discrepancy, dissonance and inconsistency between qualitative and quantitative data and results (Morgan, 2019; Moffatt, White, Mackintosh, & Howel, 2006; Pluye *et al.*, 2009). According to Pluye *et al.* (2009) divergence can occur during data collection and analysis or interpretation of results in any type of mixed methods design. On the contrary, as implied in Pluye *et al.* (2009), convergence can be defined as harmony, consistence between quantitative and qualitative data and results. Complementarity is achieved when data and results from one methods strand are

not inherently traceable in the other methods strand, yet useful to enrich study findings from the other methods strand.

As Bergman (2011) posits, and Creswell and Tashakkori (2007) imply, the professed added-value of mixed methods is not without a methodological hitch. On the premise that each methods strand has its own strengths and weaknesses, Bergman (2011) cogently argues: “*the method effect introduced by the QL [qualitative] component is not cancelled out by the introduction of a QN [quantitative] component because the latter introduces its own method effect*”. Methods-specific effects are more likely to lead to divergence than to convergence and, according to Bergman (2011), the use of a qualitative method unearths the limitations of a quantitative one, and vice versa. Each method strand may also uncover the strengths of the other. Complementary, converging or diverging data and results emanating from methodological strengths or limitations can be a natural consequence of triangulating research methods.

Convergence, divergence, and complementarity of data and results in mixed methods may stem from underlying epistemological and ontological assumptions and are underpinned by and relate to research questions. The diverging results may rightly reflect a controversy between two paradigms – a deterministic one intent on searching for causes to explain social phenomena, and a more humanistic or discursive one focused on searching for reasons to understand those phenomena. The use of the survey questionnaire presupposes knowledge of the social world and its underlying causes, i.e. that the researcher can compile an exhaustive list of categories from which the respondent chooses. These categories may make sense to the researcher but might not be how the survey respondents view their social realities – the realities which can be, arguably, investigated qualitatively. The convergence of results may originate from similar constructs under inquiry as different qualitative and quantitative measures of the same constructs can yield similar results (Jick, 1979; Nzabonimpa, 2016; Nzabonimpa 2018). In the case of the topic investigated in the study reported here, i.e. gendered choice of school subjects, some categories may very well be incorporated into a survey questionnaire, but also explored and included in qualitative measures such as interviews and essays.

The complementarity, divergence, and/or convergency of results do not come from the vacuum. They are brought about by methods effects which constitute a common methodological occurrence that most researchers attempt to counter to empirically maximise the reliability and validity of the data they collect. Following Mouton (1996) and Sheldrake (1998), these methods effects are confounding variables that include researcher, participant, and context effects. These tripartite confounding variables are known to affect study results, more so in mixed methods where data and results are brought together through triangulation. Some of these effects may be minimized by a mixed methods design that follows single- or double-blind approaches (Sheldrake, 1998). In the study of which some of the findings form the basis for this article, the researcher was not directly interacting with survey and essay respondents, but he was visible in their minds, thus a critical point that may shape data. To some extent, we note, alongside Wallace, Cesar, and Hedberg (2018), that the social presence of the researcher is more visible in face-to-face interviews than in self-administered questionnaire and can consequentially trigger various socially desirable responses. Researcher effects remain one of

the sources of harmony or disharmony between results especially in mixed methods repeated-measures design where such influence is felt twice: during the qualitative and quantitative methods strands.

With regard to participant effects, it is commonly known that study participants behave differently when they know that they are objects of investigation by a researcher (Cohen *et al.* 2000; Mouton, 1996). The researcher is usually considered a stranger “intruding” the privacy and interrupting the day-to-day lives of the study participants. This results in reactivity or reactive effects (Bryman, 2012), or participant effects (Mouton, 1996). Such atypical behaviours may reduce the level of reliability of the data collected. The survey questionnaire as a data collection instrument has been labelled such an intrusion into the life of the respondents, as it consumes their time, encroaches upon their privacy, and may lead to respondents’ over- or under-rating themselves (Cohen *et al.*, 2000). In face-to-face interviews, respondents are more likely to tell the researcher what they think he/she wants to know or would not stain their perceived or actual self-image. With respect to interviews, Cohen *et al.* (2000) point out that they are particularly problematic when mutual trust is not established, and social distance between interviewer and interviewee is extensive and hierarchical. Whether resulting from researcher or participant effects, data are to some extent a mirror of the instrument that generate them. Importantly, both researcher and study participants operate in a specific context which, in turn, cannot go without its thumbprint on data and results.

Context comes in with spatio-temporal factors in the research settings (Mouton, 1996) that may influence data. This is more likely to raise concern in a mixed method repeated-measures design, as was the case in the study of which data are used in this paper. In such a sequential design, it is also likely that the method that is employed first, and the responses it elicits, may affect the responses provided thereafter. Therefore, research studies requesting respondents to recall their past behaviours recommend shorter recall periods (Napper, Fisher, Reynolds, & Johnson, 2010). But a study that seeks to investigate the effects of methods on the reliability of self-reported data would arguably require a longer recall period, to ensure that respondents do not recall answers from the previous round of data collection. This ideal is almost impossible to achieve. As Ondaatje posits in Brockmeier (2010), human beings are inseparable from their past experiences; they retrieve and live with permanent recurrence of past narratives. A sequential mixed methods design may also be subject to repetition, or what Feldman and Lynch (1988) call “rehearsal effects”, due to data spatio-temporal distance/proximity and researcher absence/presence effects, just to name a few.

DATA AND METHODS

All research design and methods from Nzabonimpa (2016) are followed, and a subsample of the respondents that participate twice at Phase 1 and Phase 2 in a mixed methods is used. Both methods strands are equally prioritized. This allows us to investigate results from the repeated-measures design component of Nzabonimpa’s research. Within its (quasi-)experimental approach, data collection instruments (a survey questionnaire on the one hand and interviews and biographic narrative essays on the other) are switched at Phase 2 among the same sample.

As presented in Nzabonimpa (2016, 2018), the theoretical population for his study consists of girls and boys attending high schools. At the first stage of sampling, pragmatic and theoretical considerations, as well as relevant within-school characteristics (Seawright & Gerring, 2008; Gerring, 2006) guide the purposive selection of three typical high schools: a girls-only school (GOS), a boys-only school (BOS), and a mixed-sex school (MSS). At the second stage of sampling, study participants are randomly selected. From a sample size of 360 (120 respondents at each of the three schools) drawn for Phase 1, only half, also randomly selected from phase-one sample, participate in Phase 2. In total, 141 cases (see Figure 1) are matched and therefore constitute the basis upon which the comparative analysis of results is performed.

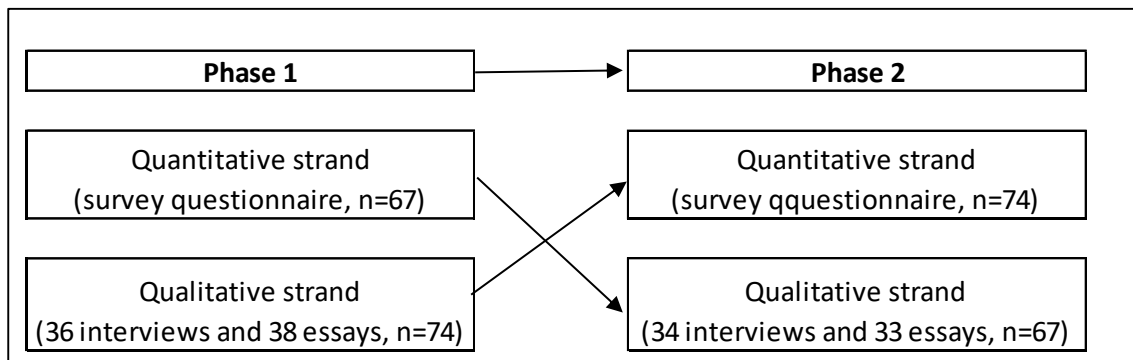


Figure 1: Diagram depicting the switching of methods strands and data collection instruments
Source: Adapted from Nzabonimpa (2016)

It is worth noting that a semi-single-blind approach is followed: the study participants are not aware of the study conditions (treatment or control) they belong to and neither is the researcher. Participants are unaware of the planned second phase of the study. At Phase 1, they are, on the one hand, randomly assigned to survey questionnaire, and on the other hand, to interviews, or essay writing. Data collection instruments are switched: half of the phase-one survey questionnaire respondents participate in face-to-face, semi-structured interviews and the other half in writing educational narrative essays, while half of both phase-one interview and essay participants fill in the survey questionnaire at Phase 2. While interviews and essays constitute different modes of collecting qualitative data, we keep both interview and essay data together as strategies to garner data of the same methods strand. The study that generated the data under use for this article was ethically reviewed and approved by two duly authorized bodies (Nzabonimpa, 2016).

The quantitative measurement of key concepts

The survey questionnaire is designed based on established and empirically validated questionnaire items used in relevant research, and which are available in the public domain. This is done to attain the levels of objectivity, validity and reliability already established elsewhere (Cohen *et al.*, 2000). However, to ensure that it is customised to the research context, the questionnaire is pretested, and relevant analysis performed to assess whether the instrument validly and reliably measures the concepts under inquiry. The translated version is also reviewed by two local language speakers (one translator and one gender researcher). Standardised scales are adapted for use in the study, i.e. Science Motivation Scale (Thoe, Thah,

& Fook, 2010); Modified Fennema–Sherman Mathematics Attitude Scale (Doepken, Lawskey, & Padwa, s.a.); and Students' Motivation Towards Science Learning (Tuan, Chin, & Shieh, 2005).

In addition to the above scales, other items are included to measure concepts reportedly underlying and/or theoretically influencing gendered educational choice, especially of science subjects. The items are worded in the first person to allow self-administration. The survey questionnaire comprises items to collect socio-demographic information (sex, age, school grade, occupation of the household head, highest educational level attained in the family, and some household assets). Questions are included on educational experiences and choices, such as favourite school subjects, self-rating of performance in school subjects, reasons motivating, or factors theoretically associated with choice of school subjects.

With regard to the role of gender in the choice of school subjects, various items are included to determine how respondents rate, from their own perspective, boys' and girls' school performance, and gender-related factors they perceive as affecting performance in science. In addition, respondents are asked questions about their future plans: career interests and plans after completion of high school education, social expectations, gender stereotypes, and career or professional goals. The survey questionnaire garners views on gender performance in science, with some items on perceived obstacles to girls' and boys' education in science, and the gender-typed nature of school subjects. Finally, the survey questionnaire includes questions on factors affecting (positively or negatively) personal motivation, self-confidence, perception of and value attached to science, as well as the extent to which survey respondents think teachers, classmates, and teaching/learning aids (can) affect their performance. The dependent variable is dichotomous, i.e. the choice of science vs. non-science school subjects, though no logistic regression modelling is reported in this article.

Qualitative measurement of key concepts

The qualitative questions are formulated based on the concepts and issues that are measured and included in the survey questionnaire as outlined above. In order to collect socio-demographic and educational data in the qualitative methods strand as in the survey questionnaire, the qualitative interview schedule and essay guide include in a tabular format questions on gender, age, school level, and grades. For each key question asked quantitatively, an equivalent or similar question is asked qualitatively. For example, the scale quantitatively measuring self-confidence in science has a qualitative counterpart in the interview and essay guides – a question about how confident respondents are or feel they are in science (See Table 1 below). Adhering to what Pluye *et al.* (2018) call *mixed instrumentation*, the data collection instruments are cross-fertilized at design stage with the ultimate goal to correlate and/or compare quantitative and qualitative responses and therefore the results emerging from each methods strand. In order to collect qualitative data using the essays, participants are asked to write biographic-educational narrative essays about their educational experiences, choices of school subjects and factors presumably associated with and influencing educational choices.

Quantitative methods strand					Qualitative methods strand																																																			
Survey questionnaire					Interview	Essay																																																		
<p><i>The following statements say something about your self-confidence in science. To what extent do you agree or disagree with these statements? Circle ONE ANSWER for EACH STATEMENT.</i></p> <table border="1"> <thead> <tr> <th></th> <th><i>Strongly agree</i></th> <th><i>Agree</i></th> <th><i>Disagree</i></th> <th><i>Strongly disagree</i></th> </tr> </thead> <tbody> <tr> <td>I am sure that I can learn science.</td> <td>4</td> <td>3</td> <td>2</td> <td>1</td> </tr> <tr> <td>I don't think I could succeed in science.</td> <td>4</td> <td>3</td> <td>2</td> <td>1</td> </tr> <tr> <td>Science is hard for me.</td> <td>4</td> <td>3</td> <td>2</td> <td>1</td> </tr> <tr> <td>I am confident that I can understand the advanced concepts in science.</td> <td>4</td> <td>3</td> <td>2</td> <td>1</td> </tr> <tr> <td>I am sure of myself when I take science exams.</td> <td>4</td> <td>3</td> <td>2</td> <td>1</td> </tr> <tr> <td>I am not the type to do well in science.</td> <td>4</td> <td>3</td> <td>2</td> <td>1</td> </tr> <tr> <td>I am confident I will do well in science tests.</td> <td>4</td> <td>3</td> <td>2</td> <td>1</td> </tr> <tr> <td>Science has been my worst subject.</td> <td>4</td> <td>3</td> <td>2</td> <td>1</td> </tr> <tr> <td>I think I could handle more advanced science.</td> <td>4</td> <td>3</td> <td>2</td> <td>1</td> </tr> </tbody> </table>						<i>Strongly agree</i>	<i>Agree</i>	<i>Disagree</i>	<i>Strongly disagree</i>	I am sure that I can learn science.	4	3	2	1	I don't think I could succeed in science.	4	3	2	1	Science is hard for me.	4	3	2	1	I am confident that I can understand the advanced concepts in science.	4	3	2	1	I am sure of myself when I take science exams.	4	3	2	1	I am not the type to do well in science.	4	3	2	1	I am confident I will do well in science tests.	4	3	2	1	Science has been my worst subject.	4	3	2	1	I think I could handle more advanced science.	4	3	2	1	<p><i>Now, I am going to ask you about your motivation and self-confidence in science. These questions have to do with your personal motivation and how self-confident you are in science. There is no right or wrong answer.</i></p> <p>How self-confident are you in science subjects? Why?</p>	<p><i>You are requested to write an essay sharing your experiences on gender and (science) education in secondary schools. This is not an exam. There is no right or wrong answer. I am only interested in your thoughts and ideas on your educational experiences. Please write an essay which covers the following topics (in other words your essay will be structured around the proposed headings)</i></p> <p>What (if anything) motivates you about science subjects, and how you view science subjects</p>
	<i>Strongly agree</i>	<i>Agree</i>	<i>Disagree</i>	<i>Strongly disagree</i>																																																				
I am sure that I can learn science.	4	3	2	1																																																				
I don't think I could succeed in science.	4	3	2	1																																																				
Science is hard for me.	4	3	2	1																																																				
I am confident that I can understand the advanced concepts in science.	4	3	2	1																																																				
I am sure of myself when I take science exams.	4	3	2	1																																																				
I am not the type to do well in science.	4	3	2	1																																																				
I am confident I will do well in science tests.	4	3	2	1																																																				
Science has been my worst subject.	4	3	2	1																																																				
I think I could handle more advanced science.	4	3	2	1																																																				

Most subjects I can handle OK, but I just can't do a good job with science.	4	3	2	1		Your self-confidence in science subjects
I can get good grades in science.	4	3	2	1		
I become anxious when it is time to take a science test.	4	3	2	1		
I'm not good at science.	4	3	2	1		
I am confident that I can understand the difficult topics in science.	4	3	2	1		

Table 1. Sample question across survey, interview and essay data collection instruments

Note: The above self-confidence survey scale is adapted and drawn from Thoe et al. (2010)

Data collection

The essays and the survey questionnaire are self-administered and completed simultaneously in the classroom or the venue allocated for this activity at the three sampled schools. Survey respondents are thus given a questionnaire to fill in, while essay respondents are given an essay guide and some sheets of paper on which to write their biographic educational narrative essays. For interview data collection, two interviewers (one male and one female) conducted qualitative interviews in the local language with respondents of the same sex, using a semi-structured interview guide. The timing of the second data collection is dictated by the intention to minimize possible recall bias and by the three-month period of school terms. Phase-two data collection is therefore conducted three months after the completion of the first phase. As empirically evidenced elsewhere, the passage of time affects memory, a factor of forgetting which affects the recall performance (Portrat, Barrouillet, & Camos, 2008).

Data processing, merging, and analysis

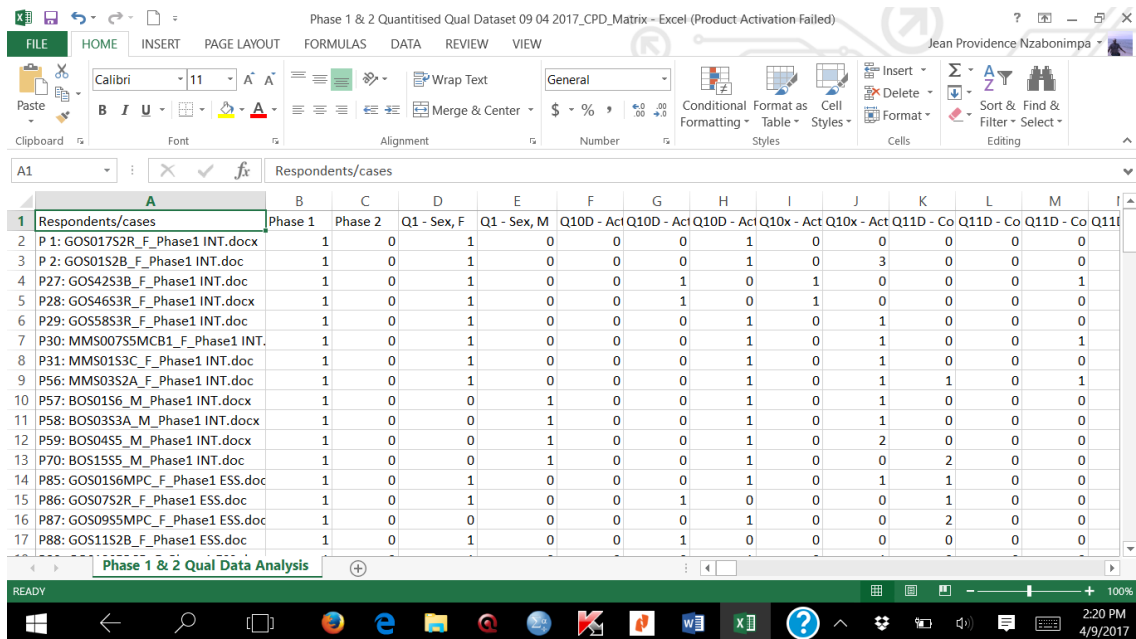
A detailed account of data processing and merging for this article can be found in Nzabonimpa (2016, 2018). What Moseholm and Feters (2017) call “matching approach”, or “mixed instrumentation” according to Pluye *et al.* (2018), is implemented with data collection instruments having related items to elucidate data from and on the same concepts. Qualitative and quantitative data collection instruments are concurrently developed with the intention to

match and merge them on questions of interest, as stated in Moseholm and Fetters (2017) and Bazeley (2009). While Nzabonimpa (2016) performed separate analyses for each methods strand within and across phases, this article, without losing sight of non-quantitizable and non-qualitizable data, focuses on merged and paired qualitative and quantitative data in a repeated-measures design. At methodological level, integration is achieved through sampling and matching cases and themes or domains (Moseholm and Fetters, 2017). At data level, integration is done through data conversion mainly by quantizing qualitative data and dichotomizing data from both methods strands with the intention to merge and triangulate them. Further details on qualitizing and quantizing are included in Nzabonimpa (2018). Integration is also done through discussion, interpretation, and reporting, triangulating quantitative-strand results with qualitative-strand results.

In an iterative manner, while remaining open to emerging insights and revisiting original data (Gibson, 2017), both qualitative and quantitative data are dichotomized to allow for merging and comparison (Bazeley, 2009) and they thus serve as a rich emporium to make sense of statistical findings. Using ATLAS.ti, qualitative data are quantized (Creswell & Plano Clark, 2011; Sandelowski, 2009; Nzabonimpa, 2018), by identifying in the textual data the presence (coded as 1) or absence (coded as 0), mention (coded as 1) or non-mention (coded as 0), yes (coded as 1) or no (coded as 0), agreement (coded as 1) or disagreement (coded as 0) to relevant statements or themes, that match, where appropriate, with the variables in the questionnaire. The resulting binary-coded data are exported to SPSS for merging with the survey questionnaire data which, where appropriate, are also dichotomized (e.g. strongly agree and agree are coded as 1 while strongly disagree and disagree are coded as 0). In addition, some responses to the scaled items in the survey questionnaire are qualitized by identifying, through factor analysis, the underlying dimensions of a concept measured by a set of items.

As reported in Nzabonimpa (2016, 2018), the method of constant comparison analysis, keyword-in-context, and classical content analysis proposed by Onwuegbuzie *et al.* (2010) or what Sapsford and Jupp (2006) refer to as “ethnographic analysis” are used. An additional coder reviews thematic categories, called coding frame by O’Connor and Joffe (2020) that are developed by the researcher to create categories consisting of similar or identical themes encapsulating the views of study participants (Castro, Kellison, Boyd, & Kopak, 2010). Identical themes are then validated for inclusion in a thematic matrix. Similar thematic categories are reviewed, discussed, and harmonised by agreeing to the correct coding. Differing themes are subjected to further scrutiny and discussion in order to develop them into agreed-upon themes for cross-paradigm correlational analysis of phase-one to phase-two data.

Using various modalities of ATLAS.ti, response quotes are interactively identified and labelled, thus producing thematic categories that can be linked up with one or more quantitative “themes”. In other words, one chunk of the narrative at times embodies more than one theme. This leads to the methodological attempt of generalising qualitative findings solely based on theoretical generalisation supported by the strength of the insights, not the numbers of occurrences (Kelle, 2006). The illustration (Figure 2) below indicates how codes and primary documents (i.e. individual interview or essay cases) are cross-tabulated to generate binary variables from the concepts qualitatively measured. The file is finally imported into SPSS.



Source: Drawn from Nzabonimpa, 2018

As the study includes a matched and repeated-measures design, McNemar's test is used in the analysis. Within this repeated-measures sample, if any respondent's unique identifying code is not found or misaligned in both phases, the corresponding case data are discarded from the paired analysis. The merged quantitative and quantitized qualitative data are analysed as a single dataset. Although there are different ways to run mixed methods data analysis (Wolf, 2010; Pluye *at al.*, 2018), we opt for statistical analyses to answer our research questions but also use content analysis to illuminate mixed methods results. McNemar tests is used to test the null hypothesis that the probabilities of changing phase-one responses are equal to the probabilities of changing phase-two responses. As SPSS version 20 does not give the McNemar chi-square value, the formula below is used to generate the value wherever significant differences are observed.

$$McNemar's \chi^2 = \frac{(c-b)(c-b)}{c+b}$$

where c represents yes/agreement to, mention/presence of a theme of a dichotomized variable at Phase 1 whereas b represents yes/agreement/mention/presence for the same variable at Phase 2. Throughout the study, p-values are expressed as $p < .05$, $p < .01$, $p < .001$ for statistically significant differences, or $p > .05$ for non-significantly different results.

RESULTS

The two-phase convergent parallel mixed methods data shed light on whether data generated at phase one converge with, or diverge from, data generated at phase two when the same sample switch data collection instruments. The percentages of phase-one respondents, as Field (2009) suggests, are compared with the percentages of phase-two respondents on a single, dichotomous variable to determine whether there is over time any significant change in response patterns. Results are dually presented: (i) from phase-one survey to phase-two essay and interview, and (ii) from phase-one essay and interview to phase-two survey responses from the same respondents. The statistical test results are presented in Appendix 1.

Phase-one survey responses (68.7%) are significantly higher than phase-two essay and interview responses (47.8%) on science as a favoured school subject. The same trend is observed with significantly more phase-two survey respondents (77.0%) reporting science as their favourite subject than phase-one essay and interview participants (. Significant change of phase-one to phase-two responses is observed regarding the type of school liked most (single-sex vs. mixed-sex school). Phase-one survey responses (70.1%) are more likely to report liking mixed-sex school than phase-two essay and interview responses (35.8%). The same trend is found between the phase-two quantitative and phase-one qualitative methods strands with the latter garnering fewer responses than the former (31.1% vs. 68.9%).

Structural factors as reason influencing school subject choice are more likely to be reported higher using survey questionnaire than using the essay and interviews. The phase-one survey questionnaire responses reporting structural factors as reason underlying their choice (62.7%) reduce significantly to 20.9% when essays and interviews are used at phase two. When essays and interviews are used at phase one and survey questionnaire at phase two among the same sample, similar findings are noted with essay and interview responses being significantly lower (23.0%) than survey responses (70.3%). As school choice is not always straightforward, it is not only dependent on structural factors, but also on personal factors. The likelihood to identify personal factors and considerations as informing the choice of school subjects is significantly differently reported between phase one and phase two. The responses in agreement that personal factors and considerations influence school-subject choice reduce significantly from 97.0% for phase-one survey to 53.7% for phase-two essay and interview responses, and significantly increase from 48.6% for phase-one essay and interview to 95.9% for phase-two survey responses.

Science as the anticipated field for tertiary studies is also significantly differently reported between survey questionnaire and essays and interviews. The phase-one survey responses are exceptionally lower (38.8%) than phase-two essay and interview responses (91.0%). The results pattern is still the same with phase-two survey responses (91.9%) being greater than

phase-one essay and interview responses (40.5%). Additional to structural factors, other findings from the qualitative methods strand indicate that there are unwritten cultural rules in the Rwandan society that require respect and compliance with parental and elders' guidance, and teaching personnel as the source of educational advice. While no significant difference is observed on the positive views and attitudes towards the role of women in society when responses are garnered with phase-one survey questionnaire (25.4%) and phase-two essay and interviews (38.8%), significant difference is found with fewer phase-one essay and interview (6.8%) participants than phase-two survey questionnaire (21.6%) respondents reporting positive views and attitudes towards the role of women in society.

With regard to de-gendering science subjects, the survey questionnaire garners significantly more responses than the interviews and essays. About 97.0% of the phase-one survey respondents vs. 59.7% of phase-two interviews and essays participants and 59.5% of phase-one interviews and essays vs. 95.9% of phase-two survey respondents mention their gender-neutral views vis-à-vis science subjects. All phase-one survey responses indicate that non-science subjects are meant for both girls and boys while the same viewpoint is held by 61.2% of phase-two interviews and essays. Significant difference is found on non-science subjects being reportedly meant for all, regardless of gender, with 52.7% vs. 94.6% of phase-one interview and essay participants and phase-two survey respondents, respectively. Whether science subjects are meant for boys or girls, no difference is found between quantitative and qualitative methods strands. However, when non-science subjects are labelled as subjects for boys, phase-one survey responses (37.3%) are significantly more in agreement than phase-two interview and essay responses (4.5%), yet the same study sample.

Between the survey questionnaire and interview/essay responses, no significant differences are found when non-science subjects are labelled as girls' subjects. On whether both girls and boys can equally perform in science subjects, phase-one survey responses (88.1%) are significantly more in agreement than phase-two qualitative responses (32.8%). The reverse is also true with phase-two survey responses (89.2%) being significantly higher than phase-two essay and interview responses (31.1%). The same results trend is also observed on non-science subjects. Significantly more responses are garnered from survey questionnaire than essays and interviews, regardless of whether it is administered during Phase 1 or Phase 2 (86.6% vs. 29.9%; 28.4% vs. 87.8%).

No significant differences are found on whether boys perform better than girls in science and non-science subjects. On the contrary, girls are significantly reported as better performers in science subjects than their male counterparts: phase-one survey responses (41.8%) are significantly higher than phase-two essay and interview responses (17.9%) and phase-one interview and essay responses are significantly fewer (10.8%) than phase-two survey questionnaire responses (29.7%). The same pattern is found, reporting girls as better performers in non-science subjects: with 59.7% of phase-one survey respondents and 14.9% of phase-two interview/essays participants, and 14.9% of phase-one interview/essay vs. 41.9% of phase-two survey respondents. Other factors contributing to success in science [such as (i) being hard-working, (ii) having ambition, (iii) concentrating on science subjects, and (iv) being talented

or naturally intelligent in science] are also examined and the same response patterns emerge with significantly higher responses observed in the survey than in the essays and interviews.

From respondents' own and parental perspective, the following interests after completion of secondary education are significantly differently reported between survey and essay/interview responses from Phase one to Phase two: (i) full-time university study, (ii) seeking employment, and (iii) starting own business. No phase-two essay or interview participant mentioned any interest in part-time work and study but about 26.9% of phase-one survey respondents expressed such interest after completion of high school. However, with regard to interest in part-time work and study, significant difference is found between phase-one essay and interview and phase-two survey responses (4.1% and 39.2%, respectively).

Concerning the field of interest for their career (i.e. public services, medical services, arts and humanities, business and commerce, and science and technology), it is also observed that responses significantly change from phase-one survey questionnaire to phase-two essay and interviews, and vice versa. Regarding social expectations towards married people, significant difference is found on having children, with 41.8% of phase-two survey responses and 26.9% of phase-one essay and interview responses. Phase-one survey (41.8%) and phase-two essay and interview (26.9%) responses do not significantly differ in terms of social expectations towards married people. On equality between spouses, significant differences are found. About 80.6% of phase-one survey responses and 22.4% of phase-two essay and interview responses, and 23.0% of phase-one essay and interview responses and 75.7% of phase-two survey responses are in support of spousal equality. The views and attitudes towards gender equity are also interrogated and significant differences are found with more responses garnered with phase-one survey questionnaire (86.6%) than phase-two essays and interviews (50.7%). A similar response pattern is observed with fewer responses collected from phase-one essays/interviews (25.7%) than from phase-two survey questionnaire (86.5%).

With respect to life goals oriented to family and professional life, significant differences are observed with 91.0% for phase-one survey responses and 52.2% for phase-two essay and interview responses, and 39.2% for phase-one essay and interview responses and 97.3% for phase-two survey responses. Prospects for a career in science and technology are significantly differently reported with 73.1% for phase-one survey responses and 38.8% for phase-two essay and interview responses, and 40.5% for phase-one essay and interview responses and 68.9% for phase-two survey responses. Running one's own business as a career goal is also significantly reported differently by both methods strands, with 61.2% for phase-one survey responses and 22.4% for phase-two essay and interview responses, and 14.9% for phase-one essay and interview responses and 56.8% for phase-two survey responses.

Though with limited response rate, significant differences are found between phase-one survey respondents (23.9%) and phase-two essay and interview participants (1.5%) as regards interest in a political career. The same response pattern is also observed with significantly lower phase-one essay/interview (2.7%) than survey (27.0%) responses. Career in public administration is also significantly differently reported with 40.3% of phase-one survey and 7.5% of phase-two essay and interview responses, and 2.7% of phase-one essay and interview and 39.2% for

phase-two survey responses. On career as a senior executive, significantly different responses are garnered with 55.2% of phase-one survey and 19.4% of phase-two essay and interview responses, and 20.3% of phase-one essay and interview and 60.8% of phase-two survey responses. On career in social sciences and humanities, survey responses are significantly higher than responses from the qualitative methods strand with 26.9% of phase-one survey responses and 9.0% of phase-two essay and interview responses, and 1.4% for phase-one essay and interview and 37.8% for phase-two survey responses. Regardless of the methods strand, no significant differences are found on home-based and cultural stereotypes against girls' success in science: 59.7% of phase-one survey and 64.2% of phase-two essay and interview responses; 60.8% of phase-one essay and interview and 71.6% of phase-two survey responses.

Obstacles to girls' success in science are also investigated and they are significantly differently reported across methods strands. Sexual harassment at school is significantly reported with higher phase-one survey (62.7%) than the phase-two essays and interviews (23.9%) responses, and the same trend is observed with 31.1% of phase-one essay and interview and 70.3% of phase-two survey responses. Men as dominant players in science both numerically and in leadership is differently reported: phase-one survey responses (53.7%) are drastically higher than phase-two essay and interview responses (6.0%), and phase-one essay and interview responses (21.6%) are significantly fewer than phase-two survey responses (58.1%).

Further questions about the gendered perception of science lead to significant differences with 73.1% of phase-one survey and 38.8% of phase-two essay and interview responses, and 37.8% for phase-one essay and interview and 81.1% for phase-two survey responses reporting science as easier for boys than for girls. The same response pattern is found on girls reportedly being naturally indifferent to science: with 43.3% vs. 10.4% for phase-one survey and for phase-two essay/interview responses, respectively, and 16.2% for phase-one essay and interview responses and 36.5% for phase-two survey responses. Unlike the aforementioned response patterns, some significant differences are found unidirectionally on boys reportedly as the more welcome in science than girls, with 13.5% for phase-one essay and interview and 31.1% for phase-two survey responses. Phase-one survey (25.4%) and phase-two essay/interview (19.4%) responses do not significantly differ on whether boys are more welcome in science than girls. On whether only exceptional girls make it a success in science, significant difference is noted from phase-one quantitative (34.3%) to phase-two qualitative (10.4%) responses. Phase-one essay/interview (14.9%) to phase-two survey (24.3%) responses do not significantly differ on whether only exceptional girls study science. On the contrary, significant differences are found on whether girls can be trusted to do well in science as boys, with agreement noted up to 82.1% of phase-one survey and 4.5% of phase-two essay and interview responses, and of phase-one essay and interview (9.5%) and phase-two survey (91.9%) responses.

With respect to factors contributing to success in science, self-confidence is reported by all phase-one survey and phase-two essay respondents as the springboard for success in science but phase-one essay/interviews (55.4%) and phase-two survey (98.6) responses are significantly different. The results on prospects offered by science are similar across phases and methods strands. Significant differences are noted when science is regarded as self-rewarding, with 71.6% of phase-one survey and 40.3% of phase-two essay and interview

responses, and 59.5% of phase-one essay and interview and 79.7% of phase-two survey responses. When science is reported as bearing an instrumental value for success, phase-one survey and phase-two essay and interview responses do not significantly differ. However, phase-one essay/interviews and phase-two survey responses are found significantly different with 74.3% of phase-one essay and interview and 89.2% of phase-two survey responses.

School-based negative factors affecting performance in science are significantly differently reported across methods strands and phases: 83.6% of phase-one survey vs. 43.3% of phase-two essay and interview responses, and 48.6% of phase-one essay and interview vs. 77.0% of phase-two survey responses. Unlike the school-based negative factors, positive factors driving performance in science are similarly reported across phases and regardless of methods strands.

All in all, results from qualitative methods strand are revisited from time to time to cross-validate findings from mixed data analysis. In the aftermath of the 1994 genocide against the Tutsis, a period characterised by revival and aspirations towards national reconstruction, rebuilding pride in one's country, or rendering one's family proud, as well as self-reliance, have become important social expectations, thus mentioned across social discourse. Such narratives are non-existing in the quantitative methods strand but richly elucidate insights in the qualitative methods strand. Apart from ticking the boxes and understand the trends and the magnitude of the issues under investigation, much could be lost in understanding educational and career choices from a gender perspective without complementary data on (i) educational experiences lived across personal trajectories, (ii) resistance against and/or surrender to a patriarchal order, as signified by the tone of the language used, (iii) participants' sense of responsibility to compensate for what they perceive to be deficiencies in their parents' educational background, (iv) the views vis-à-vis career paths which break traditional boundaries as a determining factor to address gender stereotypes, (v) motives underlying (and therefore often concealed by) educational preferences and choices, (vi) biological and therefore sex and gendered differences which learners have to navigate in educational settings, and (vii) the merits and demerits of mixed-sex vs. single-sex schools.

DISCUSSION

The bulk of the results in this article show a varying degree of divergence and convergence as well as complementarity between phase-one and phase-two qualitative and quantitative methods-specific responses garnered from the same respondents. Thus, despite the measurement of the same constructs, i.e. the use of two sets of methods-specific data collection instruments, significantly dissimilar responses are generated in mixed methods. A limited number of non-statistically significant differences in the response patterns between Phase 1 and Phase 2 are found, providing evidence that convergence of responses can be achieved when two different methods-specific instruments are used. Also, additional insights are garnered to complement the mixed-methods findings.

As regards divergence, more responses are consistently garnered with the survey questionnaire than with essays and interviews. The results significantly diverge in terms of the rate and intensity of responses garnered on issues under inquiry. In most cases, but not always, the

survey questionnaire generates significantly higher response rates – usually at least twice higher than those produced by the essays/interviews. It can therefore be surmised that results on most issues under investigation are more covered by the survey questionnaire than by the interviews and essays, as some issues are not spontaneously mentioned by qualitative-strand participants, thus contributing to low response rate in the latter methods strand. This is associated with the emphasis of qualitative research methods on participants' own perspectives, while the structured survey questionnaire reflects the researchers' concerns – a product of the researcher's ruminations about the object of enquiry, of decisions researchers have to make about what they expect to find (Bryman, 2012). In fact, mixed methods are often employed by researchers who want to both gain access to the perspectives of the people, and to explore specific issues in which they are interested (Bryman, 2012). Most of the phase-one and phase-two response rates significantly differ. To a large extent, whenever significant differences between phase-one qualitative and phase-two quantitative results are observed, they are also observed between phase-one quantitative and phase-two qualitative results. This provides evidence of the significant effects of each set of methods-specific data collection instruments.

As regards convergence, results are observed across the two methods strands testifying that similar results can be achieved. While it is observed in fewer convergence than divergence instances, non-significant results are found. An issue frequently responded to in the quantitative methods strand is also relatively frequently raised in the qualitative methods strand, but the responses are consistently higher in the quantitative than in the qualitative methods strand. To understand the common denominator(s) between methods strands, retrieved from factor analysis are the qualitatively-generated dimensions underlying or affecting choice of school subjects which are compared with and found their equivalent themes in the qualitative findings. Such triangulation process provides empirical support for converging conclusions between methods strands about the determining role certain factors play in the school-subject-choice process, i.e. social and educational networks (whether at school or in the family), self-confidence (or the lack thereof) in science, as well as a burgeoning list of other structural and personal factors.

As regards complementarity, the qualitative methods strand provides additional insights into aspects of the participants' lives that are not captured in the quantitative methods strand. It seems therefore, particularly in the case of Rwanda, that issues of personal and therefore sensitive nature do not lend themselves easily to reliable measures when only a single, closed-or open-ended data collection instrument is used. These findings seem to indicate that, methodologically, items measuring personal beliefs pose what Mouton (1996) refers to as ontological constraints, as they are not easily or at least reliably measured using either qualitative or quantitative data collection instruments. Study participants have their own ideas and views which they can empirically contribute to research findings, As Braun et al. (2020) imply, respondents prefer to express themselves by providing details on specific situation. Such details add value and context to understanding issues under inquiry.

In addition to the complementary information garnered mostly with the qualitative strand, it is observed that obstacles to girls' success in science are empirically pre-empted and therefore non-exhaustively included by the researchers in the questionnaire. Nonetheless, a number of

other obstacles (e.g. peer pressure, gossiping, poor learning strategies, lack of perseverance, poverty, unwanted pregnancies, scapegoating, just to name a few) are identified as obstacles through interviews and essays. Although such intriguing issues uncovered with qualitative methods may not occur as frequently as other issues pertaining to educational choice, they may be deemed insightful, and therefore attract the attention of the researcher on an equal footing as, or more important than, those reflected in survey questionnaire she/he imposed onto the respondents. It seems this methodological dilemma may not be resolved by quantitizing the qualitative data, but by specifically emphasising the pertinence of data collected to answer the research questions. Without the qualitative narratives concerning national reconstruction, national pride, helping families, and setting examples, just to name a few, some of the educational choices schoolchildren were making in 2012 could not have been traced back or linked to societal responses to the aftermath of the 1994 genocide against the Tutsis.

METHODOLOGICAL IMPLICATIONS AND CONCLUSIONS

Although flagged for eliminating in mixed methods terminologies (Morgan, 2019; Fetters and Molina-Azorin, 2017), we argue that methodological triangulation is a process with which we use data to ascertain whether divergence, convergence, and/or complementarity is empirically achieved, thus cross-validating study results. This article aimed at examining the convergence, divergence, and complementarity resulting from triangulation of methods and ensuing data, which constitutes one of the objectives of using mixed methods (Bryman, 2012), as it empirically sheds light on and enhances confidence in the results produced. In particular, it contributes to current mixed methods discourse on the integration challenge (Fetters and Freshwater, 2015; Moseholm and Fetters, 2017), by providing an exemplar of a fully integrated equal-weight, repeated-measures mixed methods design. It shows that beyond convergence and divergence of results from mixed data, un-quantitizable (i.e. qualitative data that are not translatable into numerical values) and un-qualitizable raw data (quantitative data that are not amenable to qualitative themes) remain indispensably the key to understanding some of the phenomena or insights (Nzabonimpa, 2018).

Scholarship on social research methodology (Bryman, 2012; Creswell & Plano Clark, 2011; Helle-Valle & Borchgrevink, 2018) suggests that when research participants are given the opportunity to formulate their own perspective in their own words and stories, results are different from those obtained when the same participants are impelled, as respondents, to select their response from a predetermined and pre-coded set of choices. As discussed in this article and reiterated by Helle-Valle and Borchgrevink (2018), the qualitative methods strand illuminates some issues under inquiry with insightful details unavailable in the quantitative methods strand; this is complementarity, not divergence. The results that differ between methods strands offer an important insight to the mixed methods researcher and serve as a reminder that there is no single way to measure and evaluate social realities. On a more practical note, the fact that the qualitative interviews and essays uncovered various responses that are not captured in the survey questionnaire supports the view that mixed methods, specifically a convergent parallel design (Creswell & Plano Clark, 2011), is an effective approach to instrument development and complementarity.

This article identifies a number of other methodological issues that have effects on mixed methods study results and suggests ways to address these. First, methods-specific instruments affect not only the response patterns, but also the number of responses garnered, suggesting that the difference between responses from qualitative and quantitative methods may lie primarily in the amount of data generated, not in data patterns as the traditional bipolarisation of these methods would lead one to believe. Evidence from this article shows that when qualitative-strand participants agree with or mention an issue, although to a significantly lesser extent, the quantitative-strand respondents also do so to a larger extent (although this applies to varying degrees between the two methods strands). It is worth noting that mixed analysis is more productive when approached in an iterative and looping rather than in a linear manner, and the latter may even be counterproductive.

Secondly, the data collection instruments are consistently influential in determining the response rate, but the qualitative-strand responses uncovered additional insights needed to empirically understand issues under inquiry. Thirdly, the findings show that, for triangulation purposes, equal priority should not be assigned to methods strands. Equal weight of both methods strands is expected to generate unequal responses, thus unfairly capturing and depicting what the traditionally small-sample-sized qualitative methods can empirically contribute. It is indispensable to navigate back and forth between narrative and numerical data while conducting the mixed data analysis to avoid de-contextualising qualitative data through quantitizing. The quantitative data are fleshed out with qualitative insights generated through the researcher's interpretation. Mixing equal-priority sequential strands is complex, costly, and overburdens both the participants and the researcher (Plano Clark *et al.*, 2015), but for the purposes of complementarity and instrument development, convergent parallel mixed methods designs warrant incisive results. Fourthly, the rationale of mixed methods is highlighted, with an emphasis on complementarity (how results from one strand lend support to the other strand) and cross-validating what is found in one methods strand with findings from the other. Some results are nonetheless challenging to interpret, but in combination, the methods strands' results shed light on an otherwise murky picture.

The narratives prevailing across the country in the post-genocide period are synchronously revisited and referenced in the qualitative strands. Findings suggest that insights from quantitative and qualitative methods strands are interwoven in the pool of mixed data. Such insights lead to understanding the explicit and the implicit, in the responses on gendered choice of school subjects, thus greatly enhancing our understanding of the issue under consideration. For collectively and widely shared gendered attitudes and views, results are convergent or complementary, but analysis of data on individual and therefore idiosyncratic experiences, unknown to others, lead to divergent, contradictory results, primarily because of the under-reporting or non-existence of such existential experiences in the quantitative methods strand.

All in all, assigning equal priority to both methods strands in convergent parallel mixed methods designs does not necessarily ensure equality between the strands, since existing research (Alise & Teddlie, 2010) points to unequal methodological contribution, as the sample size, data conversions and analyses do not call for the same strategies for each strand. In this article, the mixed methods approach improves the quality of interpretation, provides additional

evidence, and validates findings, regardless of whether survey garners higher responses than the interviews and essays. The quantitative strand provides the opportunity to critically assess qualitative results, and its quantitative measurements and statistics allow us to delineate finer differences between responses and to provide more precise estimates of the degree of relationship between concepts. Unlike the quantitative strand, the qualitative strand portrays a picture involving the participants' voices and views, seeing the school-subject choice and related issues through the eyes of participants (Bryman, 2012) or involving in absentia participants in making sense of data and findings using their narratives. When qualitative data are reduced to fixed, coded options in order to be merged with survey data, their original qualitative or narrative formats can still be referred to in an iterative and looping manner during the analysis and interpretation.

FURTHER RESEARCH

A suggestion stemming from this article for improved MMR is to investigate whether the order in which data collection instruments are administered in sequential mixed methods designs, with a control group built in the design, affects the extent to which results converge or diverge. Such research is warranted to inform the sequencing of methods strands in MMR. Currently in the advanced stages, undertaken by the same authors.

REFERENCES

References

- Alexander, E., Eppler, M. J., & Comi, A. (2020). Data Integration: A Real-Time, Participant-Driven, and Visually Supported Method. *Journal of Mixed Methods Research*, 155868982090229. <https://doi.org/10.1177/1558689820902294>
- Archibald, M. M. (2016). Investigator Triangulation: A Collaborative Strategy With Potential for Mixed Methods Research. *Journal of Mixed Methods Research*, 10(3), 228–250. <https://doi.org/10.1177/1558689815570092>
- Archibald, M.M. 2016. Investigator Triangulation: A Collaborative Strategy with Potential for Mixed Methods Research. *Journal of Mixed Methods Research*. 10(3) 228–250.
- Babbie, E. & Mouton, J. (2001). *The practice of social research* (South African ed.). Cape Town: Oxford University Press, Southern Africa.
- Bazeley, P. (2009). Integrating data analyses in mixed methods research. *Journal of Mixed Methods Research*, 3(3):203–207.
- Bergman, M. (2011). Editorial: the good, the bad, and the ugly in mixed methods research and design. *Journal of Mixed Methods Research*, 5(4):271–275.
- BLINDED, XXXX.
- BLINDED, YYYY.
- Braun, M., Meitinger, K., & Behr, D. (2020). Combining Quantitative Experimental Data with Web Probing: The Case of Individual Solutions for the Division of Labor Between Both Genders. *Methods, data, analyses*, pp. 1-25. <https://doi.org/10.12758/MDA.2020.03>

- Brockmeier, J. (2010). After the archive: remapping memory. *Culture & Psychology*, 16(1):5–35.
- Bryman, A. (2012). *Social research methods* (4th ed.). Oxford: Oxford University Press.
- Caillaud, S., Doumergue, M., Préau, M., Haas, V., & Kalampalikis, N. (2019) The past and present of triangulation and social representations theory: A crossed history, *Qualitative Research in Psychology*, 16:3, 375-391, DOI:10.1080/14780887.2019.1605272
- Caillaud, S., Doumergue, M., Préau, M., Haas, V., & Kalampalikis, N. (2019). The past and present of triangulation and social representations theory: A crossed history. *Qualitative Research in Psychology*, 16(3), 375–391. <https://doi.org/10.1080/14780887.2019.1605272>
- Castro, F.G., Kellison, J.G., Boyd, S.J. & Kopak, A. (2010). A methodology for conducting integrative mixed methods research and data analyses. *Journal of Mixed Methods Research*, 4(4):342–360.
- Cohen, L., Manion, L. & Morrison, K. (2000). *Research methods in education* (5th ed.). London: RoutledgeFalmer.
- Creswell, J.W. & Plano Clark, V.L. (2011). *Designing and conducting mixed methods research*. Thousand Oaks, CA: Sage.
- Creswell, J.W. & Tashakkori, A. (2007). Editorial: Differing perspectives on mixed methods research. *Journal of Mixed Methods Research*. 1(4): 303-308.
- Doepken, D., Lawsky, E. & Padwa, L. *Modified Fennema–Sherman attitude scales* [Online]. Available: <http://www.woodrow.org/teachers/math/gender/08scale.html> [24 June 2011].
- Feldman, J.M. & Lynch, J.G. (1988). Self-generated validity and other effects of measurement on belief, attitude, intention, and behavior. *Journal of Applied Psychology*, 73(3):421–435.
- Fetters, M.D. & Freshwater, D. (2015). Editorial: the 1+1=3 integration challenge. *Journal of Mixed Methods Research*, 9(2):115–117.
- Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). London: Sage.
- Flick, U. (2017). Mantras and Myths: The Disenchantment of Mixed-Methods Research and Revisiting Triangulation as a Perspective. *Qualitative Inquiry*, 23(1): 46–57.
- Gerring, J. (2007). *Case study research: principles and practices*. Cambridge: Cambridge University Press.
- Gibson, C. B. (2017). Elaboration, Generalization, Triangulation, and Interpretation: On Enhancing the Value of Mixed Method Research. *Organizational Research Methods*, 20(2), 193–223. <https://doi.org/10.1177/1094428116639133>
- Heesen, R., Bright, L. K., & Zucker, A. (2019). Vindicating methodological triangulation. *Synthese*, 196(8), 3067–3081. <https://doi.org/10.1007/s11229-016-1294-7>
- Helle-Valle, J., & Borchgrevink, A. (2018). Household Histories and Methodological Triangulation, *Forum for Development Studies*, 45:2, 191-215, DOI: 10.1080/08039410.2016.1260049.
- Hesse-Biber, S. (2010). Qualitative approaches to mixed methods practice. *Qualitative Inquiry*, 16, 455-468.

- Jentoft, N., & Olsen, T. S. (2019). Against the flow in data collection: How data triangulation combined with a 'slow' interview technique enriches data. *Qualitative Social Work*, 18(2), 179–193. <https://doi.org/10.1177/1473325017712581>
- Jick, T. D. (1979). Mixing Qualitative and Quantitative Methods: Triangulation in Action. *Administrative Science Quarterly*, 24(4), 602. <https://doi.org/10.2307/2392366>
- Jick, T.D. (1979). Mixing qualitative and quantitative methods: Triangulation in action. *Administration Science Quarterly*, 24:602-611.
- Johnson, R.B., Onwuegbuzie, A.J. & Turner, L.A. (2007). Toward a definition of mixed methods research. *Journal of Mixed Methods Research*, 1(2):112–133.
- Kelle, U. (2006). Combining qualitative and quantitative methods in research practice: purposes and advantages, *Qualitative Research in Psychology*, 3:4, 293-311.
- Kern, F. G. (2018). The Trials and Tribulations of Applied Triangulation: Weighing Different Data Sources. *Journal of Mixed Methods Research*, 12(2), 166–181. <https://doi.org/10.1177/1558689816651032>
- Moffatt, S., White, M., Mackintosh, J., & Howel, D. (2006). Using quantitative and qualitative data in health services research – what happens when mixed method findings conflict? [ISRCTN61522618]. *BMC Health Services Research*, 6(1), 28. <https://doi.org/10.1186/1472-6963-6-28>
- Moon, M. (2019). Triangulation: A Method to Increase Validity, Reliability, and Legitimation in Clinical Research. *Journal of Emergency Nursing*. 45(1): 103–105. <https://doi.org/10.1016/J.JEN.2018.11.004>
- Morgan, D. L. (2019). Commentary—After Triangulation, What Next? *Journal of Mixed Methods Research*, 13(1), 6–11. <https://doi.org/10.1177/1558689818780596>
- Moseholm, E. & Fetters, M.D. (2017). Conceptual models to guide integration during analysis in convergent mixed methods studies. *Methodological Innovations*. 10(2):1-11.
- Mouton, J. (1996). *Understanding social research*. Pretoria: J.L. Van Schaik.
- Napper, L.E., Fisher, D.G., Reynolds, G.L. & Johnson, M.E. (2010). HIV risk behavior self-report reliability at different recall periods. *AIDS Behavior*, 14(1):152–161.
- O'Connor, C., & Joffe, H. (2020). Intercoder Reliability in Qualitative Research: Debates and Practical Guidelines. *International Journal of Qualitative Methods*, 19, 160940691989922. <https://doi.org/10.1177/1609406919899220>
- Onwuegbuzie, A.J., Bustamante, R.M. & Nelson, J.A. (2010). Mixed research as a tool for developing quantitative instruments. *Journal of Mixed Methods Research*, 4(1):56–78
- Plano Clark, V.L., Anderson, N., Wertz, J.A, Zhou, Y., Schumacher, K., & Miaskowski, C. (2015). Conceptualizing longitudinal mixed methods designs: a methodological review of health sciences research. *Journal of Mixed Methods Research*, 9(4):297–319.
- 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 Research Approaches*, 3(1):58-72.
- Pluye P., Bengoechea E.G., Granikov, V., Kaur, N., and Li Tang, D. (2018). A World of Possibilities in Mixed Methods: Review of the Combinations of Strategies Used to Integrate Qualitative and Quantitative Phases, Results and Data. *International Journal of Multiple Research Approaches*. VOL. 10 (1): 41–56.

- Sapsford, R. & Jupp, V. (2006). *Data collection and analysis* (2nd ed.). London: Sage.
- Seawright, J. & Gerring, J. (2008). Case-selection techniques in case study research: a menu of qualitative and quantitative options. *Political Research Quarterly*, 61(2):294–308.
- Sheldrake, R. 1998. Experimenter effects in scientific research: how widely are they neglected? *Journal of Scientific Exploration*, 12(1):73–78.
- Thoe, N.K., Thah, S.S. & Fook, F.S. (2010). Development of a questionnaire to evaluate students' perceived motivation towards science learning incorporating ICT tool. *Malaysian Journal of Educational Technology*, 10(1):39–55.
- Tuan, H., Chin, C. & Shieh, S. (2005). The development of a questionnaire to measure students' motivation towards science learning. *International Journal of Science Education*, 27(6):639–654.
- Uprichard, E., & Dawney, L. (2019). Data Diffraction: Challenging Data Integration in Mixed Methods Research. *Journal of Mixed Methods Research*, 13(1) 19–32.
- Wallace, D., Cesar, G., & Hedberg, E. C. (2018). The Effect of Survey Mode on Socially Undesirable Responses to Open-ended Questions: A Mixed Methods Approach. *Field Methods*, 30(2), 105–123. <https://doi.org/10.1177/1525822X18766284>.

Appendix 1. Table of Statistical Results

Variable	Sequence of data instrument	χ^2	df	T1 (%)	T2 (%)	P-value
Science as a school favourite subject	QN (T1) vs. QL (T2)	6.53	1	68.7	47.8	<.05
	QL (T1) vs. QN (T2)	10.80	1	52.7	77.0	<.001
Mixed-sex school liked most (vs single-sex school)	QN (T1) vs. QL (T2)	15.11	1	70.1	35.8	<.001
	QL (T1) vs. QN (T2)	21.78	1	31.1	68.9	<.001
Structural factors informing school subject choice	QN (T1) vs. QL (T2)	20.63	1	62.7	20.9	<.001
	QL (T1) vs. QN (T2)	28.49	1	23.0	70.3	<.001
Personal factors informing choice of school subjects	QN (T1) vs. QL (T2)	27.13	1	97.0	53.7	<.001
	QL (T1) vs. QN (T2)	31.41	1	48.6	95.9	<.001
Science as the anticipated field for tertiary studies	QN (T1) vs. QL (T2)	33.11	1	38.8	91.0	<.001
	QL (T1) vs. QN (T2)	34.38	1	91.9	40.5	<.001
Positive views and attitudes towards the role of women in society	QN (T1) vs. QL (T2)	2.61	1	25.4	38.8	>.05
	QL (T1) vs. QN (T2)	5.76	1	6.8	21.6	<.05
Science subjects reported as school subjects for both boys and girls	QN (T1) vs. QL (T2)	23.15	1	97.0	59.7	<.001
	QL (T1) vs. QN (T2)	25.14	1	59.5	95.9	<.001
Non-science subjects reported as school subjects for both boys and girls	QN (T1) vs. QL (T2)	41.00	1	100	61.0	N/A
	QL (T1) vs. QN (T2)	29.12	1	52.7	94.6	<.001
Non-science subjects as school subjects for boys	QN (T1) vs. QL (T2)	18.62	1	37.3	4.5	<.001
	QL (T1) vs. QN (T2)	0.04	1	20.3	21.6	>.05
Both girls and boys can perform equally well in science subjects	QN (T1) vs. QL (T2)	35.10	1	88.1	32.8	<.001
	QL (T1) vs. QN (T2)	41.09	1	31.1	89.2	<.001
Both girls and boys can perform equally well in non-science subjects	QN (T1) vs. QL (T2)	34.38	1	86.6	29.9	<.001
	QL (T1) vs. QN (T2)	40.33	1	28.4	87.8	<.001

Girls do better in science subjects	QN (T1) vs. QL (T2)	9.14	1	41.8	17.9	<.01
	QL (T1) vs. QN (T2)	8.91	1	10.8	29.7	<.01
Girls do better in non-science subjects	QN (T1) vs. QL (T2)	25.00	1	59.7	14.9	<.001
	QL (T1) vs. QN (T2)	11.76	1	14.9	41.9	<.001
Factors contributing to success in science: being hard-working	QN (T1) vs. QL (T2)	46.00	1	88.1	19.4	<.001
	QL (T1) vs. QN (T2)	27.92	1	41.9	86.5	<.001
Factors contributing to success in science: having ambition	QN (T1) vs. QL (T2)	44.31	1	82.1	10.4	<.001
	QL (T1) vs. QN (T2)	44.31	1	21.6	86.5	<.001
Factors contributing to success in science: concentrating on science subjects	QN (T1) vs. QL (T2)	20.83	1	88.1	47.8	<.001
	QL (T1) vs. QN (T2)	26.47	1	47.3	90.5	<.001
Factors contributing to success in science: being talented or naturally intelligent in science	QN (T1) vs. QL (T2)	50.00	1	77.6	3.0	<.001
	QL (T1) vs. QN (T2)	41.33	1	5.4	66.2	<.001
Interests after completion of secondary education: full-time university study	QN (T1) vs. QL (T2)	46.08	1	95.5	23.9	<.001
	QL (T1) vs. QN (T2)	39.34	1	37.8	95.9	<.001
Interests after completion of secondary education: seeking employment	QN (T1) vs. QL (T2)	46.30	1	85.1	10.4	<.001
	QL (T1) vs. QN (T2)	53.00	1	9.5	81.1	<.001
Interests after completion of secondary education: starting own business	QN (T1) vs. QL (T2)	28.13	1	49.3	4.5	<.001
	QL (T1) vs. QN (T2)	23.68	1	9.5	50.0	<.001
Interests after completion of secondary education: part-time work and study	QN (T1) vs. QL (T2)	-	-	26.9	-	-
	QL (T1) vs. QN (T2)	22.54	1	4.1	39.2	<.001
Field of interest for career: public services	QN (T1) vs. QL (T2)	30.12	1	50.7	3.0	<.001
	QL (T1) vs. QN (T2)	31.00	1	4.1	45.9	<.001

Field of interest for career: medical services	QN (T1) vs. QL (T2)	11.56	1	59.7	34.3	<.001
	QL (T1) vs. QN (T2)	13.50	1	40.5	64.9	<.001
Field of interest for career: arts and humanities	QN (T1) vs. QL (T2)	-	-	26.9	-	-
	QL (T1) vs. QN (T2)	25.14	1	6.8	43.2	<.001
Field of interest for career: business and commerce	QN (T1) vs. QL (T2)	20.84	1	61.2	20.9	<.001
	QL (T1) vs. QN (T2)	16.90	1	16.2	51.4	<.001
Field of interest for career: science and technology	QN (T1) vs. QL (T2)	38.00	1	85.1	28.4	<.001
	QL (T1) vs. QN (T2)	47.00	1	20.3	83.8	<.001
Social expectations towards married people: having children	QN (T1) vs. QL (T2)	3.57	1	41.8	26.9	>.05
	QL (T1) vs. QN (T2)	12.30	1	27.0	58.1	<.001
Social expectations towards married people: equality between spouses	QN (T1) vs. QL (T2)	33.80	1	80.6	22.4	<.001
	QL (T1) vs. QN (T2)	29.82	1	23.0	75.7	<.001
Views and attitudes towards gender equity	QN (T1) vs. QL (T2)	19.20	1	86.6	50.7	<.001
	QL (T1) vs. QN (T2)	41.33	1	25.7	86.5	<.001
Life goals oriented to family and professional life	QN (T1) vs. QL (T2)	24.14	1	91.0	52.2	<.001
	QL (T1) vs. QN (T2)	41.09	1	39.2	97.3	<.001
Prospects for a career in science and technology	QN (T1) vs. QL (T2)	14.30	1	73.1	38.8	<.001
	QL (T1) vs. QN (T2)	12.60	1	40.5	68.9	<.001
Running one's own business as a career goal	QN (T1) vs. QL (T2)	19.88	1	61.2	22.4	<.001
	QL (T1) vs. QN (T2)	25.97	1	14.9	56.8	<.001
Career goal: interest in a political career	QN (T1) vs. QL (T2)	13.24	1	23.9	1.5	<.001
	QL (T1) vs. QN (T2)	16.20	1	2.7	27.0	<.001

Career goal: interest in public administration	QN (T1) vs. QL (T2)	16.13	1	40.3	7.5	<.001
	QL (T1) vs. QN (T2)	23.52	1	2.7	39.2	<.001
Career goal: senior executive	QN (T1) vs. QL (T2)	16.00	1	55.2	19.4	<.001
	QL (T1) vs. QN (T2)	20.45	1	20.3	60.8	<.001
Career goal: social sciences and humanities	QN (T1) vs. QL (T2)	6.55	1	26.9	9.0	<.001
	QL (T1) vs. QN (T2)	25.14	1	1.4	37.8	<.001
Obstacles to girls' success in science: sexual harassment	QN (T1) vs. QL (T2)	17.79	1	62.7	23.9	<.001
	QL (T1) vs. QN (T2)	19.56	1	31.1	70.3	<.001
Obstacles to girls' success in science: Men as the dominant players	QN (T1) vs. QL (T2)	30.12	1	53.7	6.0	<.001
	QL (T1) vs. QN (T2)	18.69	1	21.6	58.1	<.001
Gendered perception: science is easier for boys than for girls	QN (T1) vs. QL (T2)	13.56	1	73.1	38.8	<.001
	QL (T1) vs. QN (T2)	22.26	1	37.8	81.1	<.001
Gendered perception: girls are naturally indifferent to science	QN (T1) vs. QL (T2)	16.13	1	43.3	10.4	<.001
	QL (T1) vs. QN (T2)	7.26	1	16.2	36.5	<.001
Gendered perception: boys as the more welcome in science than girls	QN (T1) vs. QL (T2)	0.73	1	25.4	19.4	>.05
	QL (T1) vs. QN (T2)	6.26	1	13.5	31.1	<.05
Gendered perception: exceptional girls make it a success in science	QN (T1) vs. QL (T2)	9.85	1	34.3	10.4	<.01
	QL (T1) vs. QN (T2)	2.58	1	14.9	24.3	>.05
Gendered perception: girls can be trusted to do well in science as boys	QN (T1) vs. QL (T2)	52.00	1	82.1	4.5	<.001
	QL (T1) vs. QN (T2)	61.00	1	9.5	91.9	<.001
Factors contributing to success in science: self-confidence	QN (T1) vs. QL (T2)	N/A	N/A	100	44.8	N/A
	QL (T1) vs. QN (T2)	32.00	1	55.4	98.6	<.001

Factors contributing to success in science: science regarded as self-rewarding	QN (T1) vs. QL (T2)	11.31	1	71.6	40.3	<.001
	QL (T1) vs. QN (T2)	61.00	1	59.5	79.7	<.05
Factors contributing to success in science: instrumental value of science	QN (T1) vs. QL (T2)	2.57	1	79.1	88.1	>.05
	QL (T1) vs. QN (T2)	4.84	1	74.3	89.2	<.05
School-based negative factors affecting performance in science	QN (T1) vs. QL (T2)	17.78	1	83.6	43.3	<.001
	QL (T1) vs. QN (T2)	12.60	1	48.6	77.0	<.001