The Death of Mixed Methods: Research Labels and their Casualties

Jenny E. Symonds a  and Stephen Gorard b

a Faculty of Education, University of Cambridge, 184 Hills Road, Cambridge, CB2 2PQ
b School of Education, University of Birmingham, Edgbaston, Birmingham, B15 2TT

Email: jes81@cam.ac.uk
Email: s.gorard@bham.ac.uk

Labels can be helpful in structuring our understanding of phenomena. But they are also restrictive when enforcing categorical differences. The classification of all numerical research as quantitative and all other research techniques as qualitative has necessitated the construction of a third category – that of 'mixed methods', to describe studies which use both types of techniques. This paper reviews the historical development of mixed methods then analyses its construct and content validity, and propensity for bias. Underlying problems with the concept's descriptive accuracy and exhaustiveness are revealed. The findings suggest that, as a label grown out of two stereotypes, mixed methods is in danger of acting against its own aims by inhibiting new growth in research and by enforcing separation amongst methodologies and their practitioners. A case is, therefore, made for removing paradigms altogether and for a rebirth of research from the ashes of mixed methods.

KEY WORDS: mixed methods, paradigm, qualitative, quantitative, research philosophy

Introduction

'Mixed methods' is a term increasingly used in social science to describe “the class of research where the researcher mixes or combines quantitative and qualitative research techniques, methods, approaches, concepts or language into a single study” (Jonson and Onwuegbuzie 2004:17). At time of writing, mixed methods is evolving into a dominant design structure for educational research. However, our closer inspection here suggests that the concept of mixed methods has logical underpinnings rooted more in philosophy than in empirical reality. This bodes ill for current and new research projects whose design is corralled into one of the (now) three research paradigms – quantitative involving only numbers, qualitative involving anything else, and mixed methods involving both numbers and anything else - without consideration of why and how the fences between them are there, and of what benefit there may be in breaching these restrictions.

In this paper, we consider in turn the construct and content validity of the notion of 'mixed methods' [hereafter treated as a singular noun], and any tendency towards bias. This reveals underlying problems with the descriptive accuracy and
exhaustiveness of mixed methods. As a label grown out of two existing stereotypes, mixed methods is in danger of acting against its own aims by prescriptively inhibiting new growth in research. A case is, therefore, made for dropping the paradigms altogether and for a rebirth of real-life research from the ashes of mixed methods.

A Brief History of Mixed Methods

Multiple method beginnings
The beginning of mixed methods is cited by some (Creswell and Plano Clark 2007:5, Jonson et al. 2007) to Campbell and Fiske (1959) who used multiple ‘quantitative’ measures in a single study and referred to this as multitrait or multimethod research. These numerical beginnings served to demonstrate how by juxtaposing the results of multiple methods, different facets of a phenomenon can be identified – a concept later formalised by Webb et al. (1966) as ‘triangulation’. Triangulation is seen to increase validity when multiple findings either confirm or confound each other (thus reducing the chances of inappropriate generalisations). A second argument for triangulation is that “all methods have inherent biases and limitations, so use of only one method to assess a given phenomenon will inevitably yield biased and limited results” (Greene et al. 1989:256). In accord, triangulation is often cited as having methodological superiority over single methods.

Before the advent of mixed methods, many studies used multiple methods to achieve the benefits of triangulation (see for example Galton and Wilcocks 1983) without restricting themselves to any paradigmatic membership or methodological category (Tashakkori and Teddlie 2003). However, there was still awareness of the difficulties present in producing results based on multiple data types.

“The real problems associated with rapprochements come when the analysis is proceeding... Here the real problems of between method triangulation 'rise in green robes, roaring from the green hells of the sea, where fallen skies, and evil hues, and eyeless creatures be' (Chesterton, 1927)” (Galton & Delamont 1986:171).

The qualitative/quantitative debate
Perhaps in response to these issues or stemming from personal epistemologies, some single method theorists appeared unconvinced that different data types could benefit each other or that methods other than their own could provide a coherent version of reality. A growing division between camps involved different perspectives of objectivity and subjectivity in research. Survey-based researchers spent time arguing that “measurement enables us to transcend our subjectivity” (Bradley and Schaefer 1998:108) in a way that open-ended data and analyses do not. Others state that "qualitative methods are more faithful to the social world than quantitative ones” (Gergen and Gergen 2000:1027) as they allow for data to emerge more freely from context. Maxwell and Loomis (2003:342) define this as “two fundamentally different ways of thinking about explanation”. The terms quantitative and qualitative are taken to describe these perspectives and thus are
fixed onto research paradigms, and continue to be taught as social science facts to new researchers. Allegiance to one of these is still often exhibited by university faculties, journals and funding bodies, thus ‘institutionalising’ the divide (Green and Preston 2005).

The birth of mixed methods
Of particular importance during the 1970s and 1980s was the argument that the epistemological differences between the qualitative and quantitative paradigms made them fundamentally incompatible. This ‘incommensurability thesis’ created a dilemma for researchers who used methods of both qualitative and quantitative orientation in their studies. In order to justify use of multiple methods, the notion of triangulation as a superior approach was brought to the fore. Here, mixed methodologists had an ‘edge’ in being able to capitalise on the strengths of each paradigm whilst offsetting their weaknesses. Thus, a central premise of mixed methods became that “the use of quantitative and qualitative approaches in combination provides a better understanding of research problems than either approach alone” (Creswell and Plano Clark 2007:5). This is based on the logic that “one cannot separate methods from the larger process of research of which it is a part” (Tashakkori and Creswell 2007:304).

Mixed methods: the third paradigm
Mixed methods was thereafter hailed by some as the end to the incompatibility thesis as it had proven that paradigms could be combined in a single study. It thus became known as “the third methodological movement” (Tashakkori and Teddlie 2003:ix) alongside quantitative and qualitative research. As the traditional paradigms had well developed epistemologies, mixed methods’ theorists felt it necessary to find a separate philosophical foundation to support and differentiate the concept (Tashakkori and Teddlie 2003). Mixed methods is often thought of as emancipatory as it is found to be “welcoming all legitimate methodological traditions” (Greene 2005:207). Capitalising on this, many theorists, including Jonson and Onwuegbuzie (2004) and Teddlie and Tashakkori (2003) have proposed that pragmatism, or the ‘philosophy of free choice’, is the most appropriate epistemology for mixed methods. With the additional strength given by the adoption of pragmatism, mixed methods was recently claimed to be a ‘third paradigm’ (Jonson and Onwuegbuzie 2004, Jonson et al. 2007) in a trinity of otherwise incommensurable approaches.

Current definitions of mixed methods
The development of mixed methods as a concept propelled a recent study by Jonson, Onwuegbuzie and Turner (2007:118), who sought to formalise a definition by synthesising the perspectives from 31 ‘leaders’ in the field. They concluded that:

“mixed methods research is the type of research in which a researcher or team of researchers combines elements of qualitative and quantitative research approaches (e.g., use of qualitative and quantitative viewpoints, data collection, analysis, inference techniques) for the broad purpose of breadth and depth of understanding and corroboration"
This definition is almost identical to that given in the Handbook of Mixed Methods Research (Creswell and Plano Clark 2007:5):

“Mixed methods is a research design with philosophical assumptions as well as methods of inquiry. As a methodology, it involves philosophical assumptions that guide the direction of the collection and analysis of data and the mixture of qualitative and quantitative approaches in many phases in the research process. As a method, it focuses on collecting, analyzing, and mixing both quantitative and qualitative data in a single study or series of studies”.

The central element of each definition is that mixed methods is the use of both quantitative and qualitative approaches on one or more of the levels of epistemology, methodology and methods. This rests on the logic that methods, methodologies and paradigms are strongly linked.

The Case for Mixed Methods

By following the historical overview above, we can trace how mixed methods, as a concept, arose from the use of multiple methods in the mid 20th Century, then developed its identity (and became restricted to the ‘mixing’ of quantitative and qualitative components) in response to the incommensurability thesis of the paradigm wars (Gorard 2004). The possible case for mixed methods is summarised below, from its epistemological arguments to the more empirical levels where research is actually carried out.

Epistemological rationale for mixed methods:

1) That all singular methods (i.e. interview, survey) and data types (numerical, audio, visual, word based) can be classified under one of two succinct paradigms (quantitative and qualitative)
2) That elements from each of these two paradigms can coexist in a single study
3) That a third category is needed to refer to studies which use elements of both paradigms
4) That pragmatism is the philosophical basis for this third category
5) That this third category should be in itself a separate paradigm

The link between philosophy and practice:

6) That direct, normative links exist between paradigms and methodologies/methods and types of data

Empirical rationale for mixed methods:

7) That there should be a special focus on ‘mixing’ when using elements from both quantitative and qualitative paradigms in the same study, as these are inherently different in form
8) That the triangulation of epistemologies/methodologies/methods provides better quality data than a single approach
9) That by this, mixed methods is a very effective method of research
We now apply several standard validity checks – those of construct and content validity, and bias - to the construct of mixed methods as envisaged here. Our findings are that mixed methods has low construct validity and is more simply a perspective on how research can be done, and not on what research actually is. Furthermore, mixed methods can lead to bias against other real-life forms of research, and is likely to restrict the research community’s potential for creating new and more effective methods.

A Possible Case Against Current Definitions of Mixed Methods

**Construct validity: Does the construct measure what it was designed to measure?**

Mixed methods is operationalised by its description, which is contingent on the trajectory of assumptions outlined above. The foundation assumption is that all singular elements of research can be categorised into quantitative and qualitative paradigms, and this is connected to practice by the assertion that strong links exist between the data types, methods and methodologies in each category. The section below examines the construct validity of mixed methods by evaluating these two key assumptions. It asks whether singular elements of research have, by their nature, paradigmatic membership, and thereby queries the links between methods and paradigms. In doing so it exposes whether there is a basis for claims that the combination of elements from both paradigms in a single study should result in a third paradigm.

<table>
<thead>
<tr>
<th>Qualities of the Data</th>
<th>QUANTITATIVE PARADIGM</th>
<th>QUALITATIVE PARADIGM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Objective</td>
<td>Close-ended data</td>
<td>Subjective</td>
</tr>
<tr>
<td>Close-ended data</td>
<td>Open-ended data</td>
<td></td>
</tr>
</tbody>
</table>

“There is a misconception among social scientists that statistical analysis is somehow a technical, essentially objective, process of decision-making, whereas other forms of data analysis are judgement based, subjective and far from technical” (Gorard 2006:67). These assumptions underlie the categorisation of methods into separate paradigms by others. However they are not an accurate description of what is actually happening in research. Firstly, both open- and close-ended data if describing human perception are representing open-ended phenomena. The subjectivity inherent in a participant’s response to an interview question is also there when choosing a number on a Likert scale. Therefore most close-ended data is a subjective measure of subjective phenomena! Secondly, as argued in Gorard (2006), personal judgement is at the heart of all decisions that we make as researchers – in our choice of research questions, samples, questions to participants and methods of analysis – no matter whether the research is numerically based or is made up of word, audio or visual data. The choice of what to include, as described above, permeates every level of the research process and thus any type of data gathering exercise can either reduce or increase subjectivity and objectivity as it wishes.
The idea that the quantitative and qualitative paradigms are representative of objectivity and subjectivity sui generis is therefore simply a social representation based on a misunderstanding of how research is actually conducted. As Hammersley reminds us, “paradigmatic categories such as postpositivism and interpretivism... do not pick out clearly defined and differentiated positions. Indeed, some of these terms are applied almost exclusively to other researchers, rather than being self-applied” (2005:142). If all methods and evidence can be equally subjective (or objective) then a mixed methods paradigm offers us nothing further in this regard.

**Data Collection Tools**

<table>
<thead>
<tr>
<th>QUANTITATIVE PARADIGM</th>
<th>QUALITATIVE PARADIGM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Closed-Ended/Structured:</td>
<td>Open-Ended/Semi-Structured:</td>
</tr>
<tr>
<td>Questionnaire</td>
<td>Questionnaire</td>
</tr>
<tr>
<td>Interview</td>
<td>Interview</td>
</tr>
<tr>
<td>Systematic Observation</td>
<td>Observation</td>
</tr>
<tr>
<td>Document Analysis</td>
<td>Document Analysis</td>
</tr>
<tr>
<td>Official Statistics</td>
<td>Image Analysis (any image type)</td>
</tr>
<tr>
<td></td>
<td>Video Recording</td>
</tr>
</tbody>
</table>

The traditional categorisation of the many different ‘tools’, ‘techniques’ or ‘methods’ for collecting data seems to be based on whether they are closed or open in their style of data gathering and treatment of data. Clearly methods of data collection can be close- or open-ended, but their placement in a quantitative or qualitative paradigm is based only on assumptions about how the data is used. Data from close-ended methods are most often quantified, whereas data from open-ended measures are usually grouped inductively into themes or codes. However, the coding of data from open-ended methods is still quantifiable. Inversely, close-ended data does not need to be quantified for it to be used in a study. The answers to a questionnaire for a single case can be examined to create a portrait of an individual – either in one wave or across time – without the reporting of any numbers. A further supposition is that close-ended data do not give detailed, inductive versions of reality. However, detailed ‘qualitative’ data can be gathered by using close ended measures. An example of this could be in using a computer program to generate a bank of almost unlimited potential answers that the participant navigates to give an unrestricted ‘real life’ response. Here, if the survey gives options greater than the potential response of the participant to a particular question, then it is ‘open-ended’. The current assignment of close- and open-ended data gathering methods into separate paradigms is therefore based on their *most common* use, and not on their potential, or in some cases their actual, uses. Thus, mixed methods as a paradigm is not needed to combine these tools in one study.

**Sampling**

<table>
<thead>
<tr>
<th>QUANTITATIVE PARADIGM</th>
<th>QUALITATIVE PARADIGM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Large Sample Size</td>
<td>Small Sample Size</td>
</tr>
<tr>
<td>Multiple Cases</td>
<td>Single or Small Scale Case Study</td>
</tr>
<tr>
<td>Representative Generalisations</td>
<td>Non-Representative Generalisations</td>
</tr>
</tbody>
</table>
Creswell and Plano Clark (2007:111) state that qualitative data collection involves a “small number of participants and sites” whereas the sample of any quantitative study should be of an “adequate size to reduce sampling error and provide sufficient power”. This perspective suggests that qualitative research is unable to cover as many cases as quantitative research. However this is inaccurate. Incidences in research such as interviewing a hundred children in a school of five hundred are not unheard of. Nor are the ‘experimental’ quantitative researchers who might use a survey ten times on a sample of 20 children to produce the same amount of data. Both Gorard and Rees (2002) and Selwyn et al. (2006) interviewed 1,100 adults in their own homes, for example, and treated the data gathered as individual life histories and as akin to a detailed survey. On the other hand, Smith and Gorard (2005) conducted a field trial in one school with only 26 students in the treatment group, yielding both attainment scores and contextual data. Therefore, sample size is not a descriptor of any one research paradigm, nor is it attributable to method.

Furthermore, sample size is often incorrectly thought to correlate with generalisation. Large, representative samples supposedly allow for the spread of inferences to a population. Small samples are ‘non-representative’ resulting in generalisations that can only be used subjectively by the reader. However, as social scientists we study people whose group membership extends in waves from homes and families, to local communities (including institutions) and to countries and so on… thus the notion of representing a fixed population is entirely dependent on the researcher’s choice of topic. A representative sample could be 30 out of 120 children in a primary school – who are studied with only ‘qualitative’ techniques. Therefore, the quality of generalisations is not restricted to any particular sample size, nor to a specific research paradigm. Hence mixed methods research cannot claim that by nature it is ‘mixing’ different types of sample size and generalisations in accordance with one method or another.

<table>
<thead>
<tr>
<th>Type of Data Produced</th>
</tr>
</thead>
<tbody>
<tr>
<td>QUANTITATIVE PARADIGM</td>
</tr>
<tr>
<td>Numerical</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

Perhaps the most obvious difference between the two paradigms is that of the form of data associated with it or produced by it. One paradigm is concerned wholly with numbers, whilst all other types of data are lumped into the qualitative paradigm, whether they be word based, visual, auditory or indeed any other kind of sensory data. In many cases, the numerical data began as word, visual or audio data yet in the separation of paradigms this connection is lost. Surely this is not supportive of validity. Also an injustice resides in the current paradigmatic model where numbers gain a higher status than just words, sounds or images alone. The initial argument for issuing numbers with separate paradigm was that they somehow give a more objective view of reality. However, this stance is more historical than practical in origin and as such has long been challenged by qualitative theorists. Furthermore, there is currently no proof that the differences
between two particular data types are of lesser or greater value in comparison, such as that between words and images versus numbers and words. Considering the above points, there is nothing to stop theorists from claiming separate paradigms for the treatment of any one or two types of data, whether these are word, numerical, visual or audio. The mixed methodologists’ definition that using numbers and any other form of data in the same study is mixed methods is easily challenged as there is no justification for numbers to have a separate paradigm in the first place.

### Validity and Bias

<table>
<thead>
<tr>
<th>QUANTITATIVE PARADIGM</th>
<th>QUALITATIVE PARADIGM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internal</td>
<td>Descriptive/Credibility</td>
</tr>
<tr>
<td>External</td>
<td>Transferability</td>
</tr>
<tr>
<td>Ecological</td>
<td>Interpretative (social meanings)</td>
</tr>
<tr>
<td>Construct</td>
<td>Dependability</td>
</tr>
<tr>
<td>Content</td>
<td>Evaluative (judgemental)</td>
</tr>
<tr>
<td>Face</td>
<td><em>(From Lincoln and Guba 1985 and Maxwell 1992)</em></td>
</tr>
<tr>
<td>Predictive</td>
<td></td>
</tr>
<tr>
<td>Bias</td>
<td></td>
</tr>
</tbody>
</table>

Validity checks in quantitative research are usually performed by statistical analysis (e.g. construct, content and predictive validity) or relate to questionnaire design (e.g. face validity and bias). However, the underlying concepts of each validity check can be equally applied to qualitative research. For example, Dellinger and Leech (2007) argue that qualitative constructs are often negotiated between the researcher and participants, whilst in quantitative research constructs are decided on at the onset of the study. Yet it is unlikely that any construct used in research is not, to some degree, part-researcher, part-participant constructed – therefore the clarity and robustness of that construct and the resulting generalisations are at stake in any type of study. The face validity of a question or its propensity for bias relates to any research that questions the participant. Ecological or ‘real world’ validity permeates all research findings. Thus, the terms used in qualitative research are either carbon copies of those in quantitative research (and vice versa), or raise attention to areas that are consistently unchecked in quantitative methods (even though they should be). For example, ‘interpretive’ validity or subjectivity in research is a considerable issue for survey-based researchers whose personal judgement permeates the entire process, from the selection of questionnaire items to the type of statistical process chosen for use in analysis (Gorard 2006). Findings from Bryman et al.’s (2007:8) survey of 251 UK social policy researchers about validity suggest that paradigmatic division “allows both camps to perpetuate poor design without challenge from experts in the other camp” and that “inventing an alternative taxonomy” (in relation to mixed methods) “is to run the risk of making research even less accessibly to and owned by users than it is now”. Perhaps the core elements of each type of validity check are simpler and more applicable than we currently imagine. Therefore it is redundant to argue that mixed methods should have a distinct set of validity descriptors, as these will refer simply to ‘problem’ areas that are fundamentally common within all types of research.
None of the analytical techniques listed above are restricted by data gathering methods, input data, or by sample size. Interview data can be counted, and are anyway traditionally presented in terms of (disguised) numeric patterns such as ‘most’, ‘many’, ‘few’, ‘none’ and so on, while surveys routinely generate rich comments (Gorard 1997). The defining factor for statistical analysis is often the amount of data collected, since the necessity for randomisation is now largely evaded in statistical treatments based on sampling theory. Thus, statistical analysis can be performed on the types of words used in a single interview transcript, or on the geometric properties of items in a single image. Conversely, numerical data and its results can be analysed by inductive coding - of the types of responses given across measures by individual cases - just as it can for interview responses. Numerical results can also be displayed in matrixes and conceptual maps. Qualitative evidence can be usefully modelled. Even narrative analysis applies, as it is common practice to uncover new insights through writing and talking about one’s survey findings and thus numbers are analysed through word and language based mechanisms. By this we can see that no method of analysis is fixed to any one paradigm, thus the separation of these is artificial and does not support mixed methods.

Section Summary
Our review suggests that it is not clearly established whether the key components of ‘methods’, i.e. data gathering tools, sample sizes, generalisations, data types and analytical techniques, must be restricted into either the quantitative or qualitative paradigms. As they do not have a strict placement, there is no evidence for direct, normative links between methods, methodologies and paradigms. Subjectivity and objectivity are not separate conditions of the qualitative and quantitative research paradigms, nor are data collection tools as any data produced can be used numerically or in narrative analysis and are always subjective in form. Close ended measures are not more restrictive than open ended measures when the number of options exceeds the potential answers given by a participant. Sample size and generalisations do not determine a study’s paradigmatic membership as any study can have a large sample and even the smallest study is representative of some value to some degree. Data, whether they be numerical, word, image or audio, are equally differential in type to one another and require as much concerted effort to ‘mix’ or include in a single study as do numerical data with any other. Therefore there is no basis for the joint categorisation of audio, visual and word data thus rendering the need for just two paradigms, obsolete. Core validity issues are inherent in all research processes thus although validity checks can be ‘data specific’ they ultimately test the same areas. By querying all major steps of the arguments for dividing all types or aspects of research into one of either the quantitative or qualitative categories, this critique suggests that the act of creating a third paradigm, or category of mixed methods, is a fallacy. When examining the
construct validity of mixed methods, we find that it only holds true for those researchers who are philosophically committed to bipolar paradigms anyway, and has very little bearing on how research is and can be conducted.

**Content validity: How well does the construct covers all aspects of the phenomenon?**

It is difficult to test for content validity when the phenomenon being studied is only relevant under certain assumptions. Therefore, we use the criteria set out by mixed methodologists of what constitutes a 'good' mixed methods study to judge which reports meet these standards and which are not 'good' examples of mixed methods. The reviewed studies are not unusual in their approach and in this are deemed loosely representative of mixed methods studies in general.

**Criteria for mixed methods studies**

Tashakkori and Creswell as editors of the Journal of Mixed Methods Research describe a need “for distinguishing between studies that utilize two types of data without serious integration and those that integrate the findings of the qualitative and quantitative strands” (2007:4). A purposeful focus on mixing is also called for by Creswell and Plano Clark (2007:115) who state that “incorporating one form of data into a study with another form of data... does not mean that the researcher has combined or integrated the data in any meaningful way”. A good mixed methods study is therefore:

“The intentional collection of both quantitative and qualitative data...
The investigators know the reasons that both types of data are needed, and they state these reasons. Also, rigor is added when the authors report detailed quantitative and qualitative procedures and used mixed methods terms to describe their study” (Creswell and Plano Clark 2007:163-164)

The following examples suggest that many studies do not meet these mixed methods criteria, whilst others meet all criteria yet defy strict categorisation as ‘mixed methods’. We then face difficulties in deciding what these studies should be called – if they are not strictly or solely mixed methods, and are not ‘qualitative’ or ‘quantitative’ either, then what are they? Are they still to be considered research at all?

**Some studies that use two types of methods but are not mixed**

Klassen et al. (2008) examined secondary school teachers’ motivation beliefs in Canada and Singapore. They used qualitative data (interviews with 24 teachers) to “provide support and explanation” (:4) for the quantitative findings (a survey of 502 teachers). Klassen et al. give their reasons for doing this and clearly explain the use of each method. They report each study and its findings separately – and by comparing the results from these we can see that some of the findings are similar (hence providing support for each other). This is similar to Vuttanont et al (2006) who also performed two separate studies within their investigation of sex-education in Thailand. First they surveyed 2301 pupils and 351 parents, then
performed 20 focus group interviews with pupils and two with parents. They also interviewed ‘key stakeholders’ and analysed policy documents. These ‘qualitative’ data were analysed as a group. The results of all of the methods are reported together in the article, yet until this point no concerted ‘mixing’ of data sets or methods occurred in the study. Webb and Macdonald (2007) also reported their results of a survey of 556 teachers and head of departments, in tandem with results of their ‘qualitative’ field observations, document reviews and interviews with 17 participants – without any mixing of methods in the actual research process.

Although these studies calls themselves ‘mixed methods’, they do not fit the criteria set out by mixed methodologists of a ‘purposeful focus on mixing’. Each is simply an example of two separate approaches to studying the same phenomenon in one population and study. These approaches could be autonomously conducted by separate research teams and still be reported together in an article to the same effect. Therefore, as no real ‘mixing’ occurs in the design of these studies – they do not fit into the mixed methods category. However to disregard these studies would be to ignore the value contained in their triangulation of results. Furthermore, would the triangulation be as useful had they actually ‘mixed’ their methods earlier on in the piece, or would this bias the results? This is presented as an example of how research ‘standards’ can inadvertently work against their own aims.

*Studies that transform data instead of mix it*
Flumerfelt et al (2007) performed a document analysis of 34 student professional development plans, written by graduate students at a Midwestern university in the United States. Two examiners appraised how well the plans covered two standards criteria for a ‘whole person’ taxonomy, between groups of pre-service and practicing student administrators. The two standards consisted of six and eleven descriptors respectively. These results were then ‘forwarded…to the statistician for analysis’ (p.110). Student t-tests were performed on the data to show significant differences between groups for the standards. This study is an example of how one mode of data gathering produced data that was transformed into numbers. Exactly where this is ‘mixed’ is unclear – as the change from open-ended to close-ended is a feature of all research that restricts potential information into set categories.

*Studies that meet mixed methods criteria yet operate outside of these*
Blatchford et al. (2007) investigated the relationship between teaching assistants (TAs) and class size, pupil behaviour and achievement. Their reason for using mixed methods was to address the different elements of these relationships. The links between the presence of TAs and class size were examined by a survey. Teacher and pupil behaviour in class in relation to TAs was studied with systematic observation, whilst pupil achievement was determined by results on SATs and QCA optional tests. Their ‘qualitative’ component was in case studies of schools. Here they combined semi-structured interviews and freely given comments from teachers and fieldworkers with systematic observation. The use of systematic observation in a case study which is called ‘qualitative’ shows how numerical data can be used to build up a detailed picture without a large sample.
Therefore, Blatchford et al. are already operating outside of the boundaries of paradigmatic membership in the case of individual methods. Secondly, although the results from the survey and systematic observations are reported in their paper, there is no discussion of the interview findings. Without this, we can only assume that their conclusions are based solely on numerical data. The use of systematic observation for qualitative work, and the lack of reporting of open-ended data, shows how this study does not fit the strict definition of mixed methods (above) and does not meet quality criteria for this.

The discussion of mixed methods design of the ‘Variations in Teachers’ Work, Lives and Effectiveness’ (VITAE) in Sammons et al. (2007) illustrates how 300 teachers in 100 schools in seven LAs were surveyed and then interviewed to assess their teaching styles and efficacy, and how this was compared to pupil’s test results as a measure of teacher effectiveness. Already this study defies paradigmatic categorisation as interview data was taken from a large sample, whilst quantitative data was used alongside interviews to build case studies (like Blatchford et al, 2007). Both numerical and interview data in the case studies were coded thematically. From these an amount of teacher profiles emerged and the number of teachers in each profile were reported in percentages. This study is a good example of how any type of data can be analysed ‘qualitatively’ and can be quantified – thus supporting our argument that methods of data collection do not belong in any one paradigm.

Section Summary
Most of the studies claiming to be ‘mixed methods’ given above are those that perform two separate investigations of the same phenomenon in the same sample. Likewise, in Greene et al.’s (1989) review of 55 mixed methods studies, the majority of studies (n18) simply collected both types of data and synthesised the findings from these in discussion. In twenty years of mixed methods research this is still a dominant pattern – suggesting perhaps that it is a model that works well for some researchers, despite mixed methodologists’ attempts to devalue it as more ‘sophisticated’ models of mixing come about. A further set of studies show how individual methods often defy paradigmatic categorisation, and are not as much ‘mixed’ as they are examples of whole research processes in their own right. Flumerfelt et al. (2007) and Sammons et al. (2007) produce quantifiable data through an open-ended method, whilst Blatchford et al. (2007) use a traditionally ‘quantitative’ method (systematic observation) to provide detailed, rich description in a ‘qualitative’ case study.

In different ways, all of these studies fall outside of the standard description of ‘mixed methods’. Likewise, they cannot be categorised as qualitative or quantitative. They are evidence contra the argument that the three research paradigms are exhaustive of all research, and illustrate that the content validity of mixed methods is far from established.

Unintended Biases
Lastly we discuss the potential biases against other possible types of research stemming from the creation of the category of mixed methods.
**Mixed methods requires a special focus on mixing**

A common argument from mixed methods theorists is that the use of numerical data with other data types requires a special focus on mixing. By this they postulate that mixed methods deserves to be a separate category in research. As Greene (2008:18) says, “it is the mixing that is distinctive to a mixed methods methodology”. However, as discussed, there is no evidence that any two types of data, for example words and sounds, versus images and numbers, are more similar to each other than the rest. Therefore the suggestion that there needs to be a special focus on mixing numbers with other data types, such that this requires a new research method, is biased.

Furthermore, the assumption that only numbers deserve a focus on mixing distracts our attention from the issues inherent in triangulating or combining any two forms of data. Russek and Weinberg’s (1993) ethnographic ‘mixed methods’ study is a good example of this. They gathered qualitative data from interviews, classroom observations, analysis of school documents and open-ended questionnaires, and quantitative data from classroom observation checklists, lesson evaluation forms, workshop evaluation forms and two close-ended questionnaires of teacher perceptions. Over roughly a quarter of their article they discuss the ‘mixing’ of the grouped quantitative and qualitative data without any consideration of how the individual data types have been combined or triangulated within or across these categories. Although having a focus on mixing numbers with other data can improve the quality of some research, it can also result in skewed attention to numbers and a reduction of attention to other types of data triangulation.

A simple test of whether mixed methods requires a special focus on mixing and is unique in this, can be made by imagining that advice given for mixed methods refers to the mixing of multiple qualitative or quantitative data types. Take for example Greene’s justification of why we should have mixed methods, altered to read for multiple methods as follows:

“A multiple **mixed** methods way of thinking is an orientation toward social inquiry that actively invites us to participate in dialogue about multiple ways of seeing and hearing, multiple ways of making sense of the social world, and multiple standpoints on what is important and to be valued and cherished...Better understanding of the multifaceted and complex character of social phenomena can be obtained from the use of multiple approaches and ways of knowing. (2008 :20)

As demonstrated, Greene’s advice holds true for multiple method studies and for mixed methods. The presumption that mixed methods is unique in mixing data types is inherently biased. As Yin (2006:42) states, “once freed from the quantitative-qualitative dichotomy, the relevance and reality of a broad variety of ‘mixes’ emerges. The broad variety recognizes the true diversity of the research methods used in education”.

13
Mixed methods research is somehow better than other types of research

Perhaps the main argument presented in support of mixed methods is that through triangulation, mixed methods offsets the weaknesses of single research methods and provides better quality data. However, this is inherently biased. First, it supposes that the triangulation of numbers with other data types is better than triangulating other types of data such as observations and interviews. Secondly it presumes that the triangulation of large and small samples and their resulting generalisability is restricted to studies involving numbers. This bias towards numbers is challenged by Giddings and Grant (2007:59) who warn that “in spurring on such effacement, mixed methods research is a Trojan Horse for positivism, reinstalling it as the most respected form of social research, while at the same time—through inclusion—neutralizing the oppositional potential of other paradigms and methodologies that more commonly use qualitative methods”. A third bias is the assumption that the triangulation of methods will address methodological weakness. This is illustrated in Creswell and Plano Clark’s argument (2007:9) that,

“...quantitative researchers are in the background, and their own personal biases and interpretations are seldom discussed. Qualitative research makes up for these weaknesses. On the other hand, qualitative research is seen as deficient because of the personal interpretations made by the researcher...the combination of both approaches can offset the weakness of either approach used by itself”.

This statement suggests that we should use one method to cover up the limitations of another, instead of addressing the weaknesses within the methods themselves. By this it identifies limitations that are present in the traditional paradigmatic perspectives of qualitative and quantitative – yet perhaps does not reflect on how research is actually carried out. Nor does it explain why and how the mixing must lead to offsetting the weakness of each rather than the equally plausible loss of the strengths of each.

A fourth bias exists when arguing for triangulating epistemologies and methods, as this devalues studies which collect and analyse only one type of data. Indeed it appears that some mixed methods theorists want to remove single method studies altogether: “Mixed seems to be where it is at, though some identities hang on, and the image of the introverted statistician (the nerd) or the hang-loose ethnographer are by no means eliminated” (Green and Preston 2005:167). This assumption relies only on the quality of the study at hand. There is nothing preventing a single methods study from being as rigorous, employing triangulation within a single data set and providing better quality results than a study of mixed methods. Accordingly, Yin (2006) advises that studies should not just mix numbers with other data types, but should also be free to mix ‘quantitative’ methods without any qualitative method present and vice versa.

Many of these biases are identified (perhaps unintentionally) in the definition of Bryman et al (2007) who distinguishes between a ‘universalistic’ discourse, where mixed methods is always considered superior, and a ‘particularistic’ perspective of mixed methods only being appropriate when it best answers the
research question. The former perspective is exemplified by Jonson, Onwuegbuzie and Turner (2007:129) who view mixed methods as the dominant research paradigm that “will provide the most informative, complete, balanced, and useful research results”. Tashakkori and Teddlie (2003:x) also propose that “mixed methods research designs will be the dominant methodological tools in the social and behavioral sciences during the 21st century”. Hammersley (2004:201) critiques this bias when reviewing Teddlie and Tashakkori’s handbook (2003) which he refers to as “a paradigm-building exercise” which assumes that choosing mixed methods “is always desirable”. Drawing on Hammersley’s perspective, we would argue that researchers should focus more on designing studies that best suit their research topic, no matter what types of methods are used, rather than making the assumption that triangulating qualitative and quantitative data as currently defined will necessarily be most effective.

So what could these potential biases look like at a practical level? They may take the form of research institutions teaching their students that mixed methods is the best methodology. They appear when researchers who have previously used solely quantitative or quantitative methods feel undermined by the presumed dominance of mixed methods, thereby try to excuse their choice of single methods and move towards developing a more mixed methods approach. Importantly, funding bodies can be swayed into thinking that mixed methods is nearly always the most effective kind of research, and thus distribute their funding accordingly. To combat this, we should remain keenly aware that “…research paradigms participate in a form of competitive modernism, each overselling itself in the academic marketplace” and that “all this makes the development of knowledge through educational inquiry minimal, at best” (Hammersley 2005:142-143).

**Conclusion: The Death of Mixed Methods**

"Mixing methods is wrong, not because methods should be kept separate but because they should not have been divided at the outset” (Gorard, 2007:1)

Far from freeing researchers from the restrictions of paradigms and the strife of paradigmatic struggle, mixed methods can actually reinforce categorical differences. Giddings (2006:195) argues that “the use of the terms ‘qualitative’ and ‘quantitative’ as normative descriptors reinforces their binary positioning, effectively marginalising the methodological diversity within them”. This diversity is trapped by the boundaries of research paradigms, thus the third category of mixed methods simply provides a further barrier to enforce categorical differences. As stated by Gorard (2007:1), “The process of mixing requires distinct method elements to mix and so, ironically, the metaphor of mixing actually works to preserve method schisms in part”. “This attempt at building a new mixed methods paradigm could obscure the growing points for what might be a more fundamental reintegration of qualitative and quantitative methods” (Hammersley 2004:201). By endorsing mixed methods as a ‘third category’, we uphold paradigmatic separatism and thus create a world of limitation. Within this the development of new and better research techniques is hampered, especially when mixed methodologists lean towards determining how other people should do
research. For example, Greene (2008:17) recommends for mixed methods to develop guidelines about “how to choose particular methods for a given inquiry purpose and mixed methods purpose and design”. By this she would strip power from the individual researcher and give it to the mixed methods theorists. Jonson et al. (2007:127) also argue for a ‘contingency theory’ for “the conduct of human research” where conditions for the selection of using qualitative, quantitative or mixed methods research should be met by all researchers. We would argue that this type of unified strategy would surely not suit all research and instead that “progress could be seen as an evolutionary process with no specific ‘life-form’ destined to emerge from its pressures” (Gorard 2004:12). Here, “individual researchers should be free to identify the most productive areas of inquiry and to determine the most effective means for investigating them” (Hammersley 2005:144). Without this freedom, we are unlikely to step beyond the paradigmatic boundaries and towards methods that better meet the purposes of our research.

This paper has considered the category of “mixed methods” by examining its validity as a construct, its exhaustiveness as a category, and its tendency to create bias and limitations. A review of the history of mixed methods reveals its beginnings as a bridge between the qualitative and quantitative research paradigms in the era of the paradigm wars. However its bid for methodological freedom rests on there being two very separate paradigms to begin with, and that there are direct, normative links between epistemology, methodologies and methods to allow for the mixing of elements between paradigms. When examining the types of research that can be and are conducted we can see that the elements are categorised with respect only to a particular use and therefore these links are not definitively there. Our brief review shows that mixed methods studies do not always fit the descriptors given for mixed methods, and that often two types of data are used without being mixed at all. As studies which do not fit the criteria for any of the research paradigms are still accepted as research, we can state that the categories of quantitative and qualitative, whether mixed or unmixed, are not exhaustive. This destroys the logic that a third category is required to contain the remainder of studies (that of mixed methods) and moreover that this should be a separate paradigm. When held as a separate category with particular qualities, mixed methods creates several types of bias. It assumes that a special focus on mixing is required for numbers with any other types of data, and that the resulting triangulation is better quality research. This devalues other data types, the use of single methods, and multiple methods that do not involve numbers or that mix solely numerical data. It may endanger how other types of research are perceived by funding bodies, the public and by research practitioners. Mixed methods also reinforces boundaries in research by basing itself on the quantitative and qualitative research paradigms and by inventing rules for ‘mixing’ data that researchers must follow. By doing so it necessarily inhibits the growth of new and innovative research techniques.

Accordingly, mixed methods has very low validity when considering its current use. Just like any category developed through coding, we could drop it entirely and return to the sounding board to search for new and better ways to describe our research. Hammersley reminds us that:
“Many of the purported divisions are artificial, involving spurious claims to novelty, and are based on cavalier use of philosophical and methodological arguments. I also think that we need to be rather more sociologically sophisticated in seeking to understand why educational research displays this character at the present time” (2005:142).

A ‘radical’ suggestion is to stop trying to promote a certain type of quality through the construction of overarching categories and researcher identities, and focus rather on the quality of our actual research techniques, the resulting data and on how that data is used, no matter whether this involves one or more sets or types. We could use the word ‘quantitative’ to refer only to the activity of quantification, and ‘qualitative’ to describe that which is examined in depth – without being linked to a research paradigm. Here, freedoms are opened up such as embracing the subjective basis of quantification (as being the counting of non-numerical data), and using numbers qualitatively. Mixed methods itself could refer to any study which purposefully integrates multiple techniques to achieve a final set of data, whilst ‘multiple methods’ can be used for studies that use more than one method and report the results of these separately. This use of terms better describes actual research practices and moves us closer to identifying the “universal underlying logic to all research” that leaves “little or no place for paradigms” (Gorard 2007:3). As stated by Jonson and Onwuegbuzie in support of mixed methods, “It is time that methodologists catch up with practicing researchers!” (2004:22). Considering the inappropriateness of the qualitative and quantitative paradigms and the category of mixed methods, we advocate the development of a research community where “all methods have a role, and a key place in the full research cycle from the generation of ideas to the rigorous testing of theories for amelioration” Gorard (2005:162). In order to achieve this we call for the death of mixed methods, and the rebirth of plain ‘research’ as a craft.
References


Giddings, L & Grant, B. (2007). A Trojan Horse for positivism? A critique of mixed methods research, Advances in Nursing Science 30(1)


Smith, E. and Gorard, S. (2005) ‘They don’t give us our marks’: the role of formative feedback in student progress, Assessment in Education, 12, 1, 21-38


