Let the Data Speak: Using Rigour to Extract Vitality from Qualitative Data

Alistair J Campbell
Edith Cowan University, Perth, Australia
alistair.campbell@ecu.edu.au
DOI: 10.34190/JBRM.18.1.001

Abstract: Qualitative data can be gathered from an array of rich sources of research information. One of the popular ways to collect this data is by interviewing a range of experts on the topic, followed by transcription, resulting in a database of written documents, often supplemented by other documented data that informs the topic. Thematic or Content Analysis can then be used to explore the data and identify themes of meaning that enlighten the research topic, with the themes being gathered into nodes. The researcher now has an array of nodes, which needs to be organised into a coherent model, and more importantly, one that represents the views of the research informants. To do this with some degree of rigour, the researcher needs some way of ranking the nodes in terms of their relative importance. The node ranking can be based on experience, or on the literature, but neither of these approaches looks to the data itself. If the database contains new or unexpected knowledge, neither experience nor the literature will guide us to it, and vital new insights may easily be missed. The framework outlined in this paper aims to provide a sound first-cut analysis of the data, based on the evidence in the research interviews themselves. Clearly the literature and research experience have an important role to play in shaping the results of any research. However this paper argues that one should proceed only after the data itself has been offered “the first chance to speak”. The node classification matrix detailed here, identifies distinct node categories, each ranging in significance and with particular characteristics that reveal key aspects of the informants’ views. In this way the researcher can use the nodes to reveal the voice of the experts, and build a scientifically rigorous set of results from a qualitative database.

1. Introduction
1.1 Qualitative Approach:

Qualitative research is quite different to a quantitative or statistical research approach which typically rests on a finite range of answers to questionnaires, or hypotheses that are put to a sample of numerical data. These statistical results may indicate a correlation, and occasionally causality, among variables that are thought to be fundamental to the investigation. The quantitative approach is however limited to the confirmation of known concepts, and assumes that we know enough about the phenomenon being investigated to ask probing questions, and pose credible hypotheses. Statistical results may therefore add to extant knowledge by polling a range of known phenomena in new geographical or social environments.

However if we are dealing with a topic that is exploring new ground, then a set quantitative questions will be unlikely to uncover anything unusual. Davidsson (2005) puts it this way: "events that are infrequent, unanticipated or extraordinary... phenomena of this kind may be difficult to capture with conventional, quantitative approaches".

This is where the open-ended nature of qualitative research, and semi-structured interviews in particular, begin to show value (Alasuutari 2010). For any interview there will of course be a topic of focus, and this is what gives it structure. However a skilled researcher will allow the respondent considerable leeway to explore the topic in their own terms (Denzin and Lincoln 2013). In this way, particularly through multiple interviews, and other sources of qualitative data, a new body of knowledge can be created by probing the experiences and insights from a range of experts in the field.

Qualitative approaches are particularly useful when investigating research topics that study creativity, where the research method needs to accommodate the unexpected, or surprises, and aspects that might be novel or even unique (Davidsson 2005, Yin 2018). A quote often attributed to Einstein asserts that: 'No problem can be solved by the same kind of thinking that created it' (Bohm and Peat 2000). In this sense, and particularly where creativity or innovation is being explored, the research approach needs to be open to accept answers that are outside a pre-conceived framework of questions. Only then can new knowledge be added, novel insights be garnered, or innovative models be proposed.
Pursuing the qualitative route however involves complexity, as qualitative data comes in many forms; for instance Yin (2018, pg.114) lists six possible sources of qualitative data: documentation, archival records, interviews, direct observations, participant-observations, and physical artefacts. An important point to make is that a mix of these sources can be used to inform a single investigation, and having multiple sources naturally strengthens the rigour and depth of the database (Yin 2014, Fabregues and Molina-Azorin 2017).

Also, when seeking to uncover new knowledge using qualitative approaches, rather than confirming existing constructs using statistical methods, the so-called Meno’s Paradox comes into play. Fine (2014, pg.8) paraphrases this paradox as follows:

But how will you inquire into this, Socrates, when you don’t at all know what it is? For what sort of thing, from among those you don’t know, will you put forward as the thing you’re inquiring into? And even if you really encounter it, how will you know that this is the thing you didn’t know?

Clearly in many cases we know enough to at least begin a research enquiry. However the central issue remains – if our knowledge is incomplete, how will we know what questions to ask, and will the answers expand our knowledge? Mir (2018) argues that this boils down to little better than an educated guess (which is undoubtedly preferable to an uneducated one), but the questions are still largely guesswork. It is therefore in the focussed (but unfettered) expert knowledge that the potential value lies, and it is up to the researcher to unearth and present this value.

Qualitative data is therefore rich (Yin 2018, Yang 2013, Neergaard and Ulhoi 2007), and extracting just the relevant themes from the database can be a challenge, requiring focus, dedication, skill, and repeatability from the researcher. Nowell et al (2017) point out that Thematic Analysis is often used as a basis to guide this exploration, given that it is a relatively easy concept to understand, especially for early career researchers or those that are new to qualitative approaches.

Nowell et all (2017) goes on to set out a framework to improve trustworthiness in Thematic Analysis, while noting that it is one of few contributions in this field. This paper will therefore seek to add to this general aspect of rigour as well. However the qualitative researcher still faces another important hurdle.

Once a collection of relevant themes has been amassed, it is vital to understand the hierarchy among them – are some themes more relevant than others? There are a number of ways this hierarchy could be achieved – using the researcher’s experience could be a way, but the question of impartiality would be the first issue to arise. Another option would be to base the hierarchy on extant models or constructs in the literature – in this case the question is whether the research is then just destined to discover what is already known? Crucially, neither of these options leverages the unique power of qualitative research approaches to uncover the new, the unexpected, or the surprising.

This paper will therefore seek to develop a coherent and repeatable framework for determining a hierarchy in the identified themes in the data, by using attributes that are found in the database itself, and therefore reflect the value of the expert knowledge that has been gathered into that database.

1.2 Aim and Objectives

To summarise the above introductory discussion, the Aim and Objectives are stated separately here:

Aim:
This research will seek to develop a rigorous framework for researchers using qualitative methods, that supports the extraction of coherent, novel and validated results from the gathered data.

Objectives:
- Leverage off the ubiquitous and extant nature of thematic and content analysis, using software search tools and concepts from phenomenology to augment thoroughness, meaning and rigour when identifying themes.
- Develop a framework that analyses the relative value of themes identified in the database, to reflect the value inherent in the expert knowledge in the data.
• Illustrate the developed framework using real-world examples.

1.3 Working with Qualitative Data

This paper is therefore focused on qualitative research, and as many experienced researchers can attest, in some fields of exploration this still runs against the current (Yang 2013; Mir 2018; DeGregorio, Tagliafico and Verde 2018). Mir (2018) has some strong views on this, asserting that "splitting empirical research in the social sciences into "qualitative" and "quantitative" camps is an act of discursive and epistemic violence". Likewise Yang (2013) describes her experiences as "Surviving as a qualitative researcher in a quantitative world". While quotes like these may not typify the majority of qualitative researchers, it is certainly enough evidence to be mindful that the status quo may include critics of qualitative approaches.

It is therefore important that the qualitative methods of analysis are rigorous, and can support results and conclusions that are both valuable and robust (Neergaard and Ulhoi 2007, Fabregues and Molina-Azorin 2017).

This paper explores a framework that can be applied to a wide range of qualitative data, with a particular focus on documented data from multiple sources, and these could be emails, interview transcripts, promotional material, or other written documents.

1.4 Thematic and Content Analysis — a search for Meanings

Some research makes a distinction between Thematic and Content Analysis (Crowe, Inder and Porter 2015; DeGregorio, Tagliafico and Verde 2018) although the differences can be quite subtle at times, with a large degree of overlap and similarity (Vaismoradi et al 2016). The framework outlined here is a qualitative analysis of the documents, and will consider the terms Thematic Analysis and Content Analysis interchangeable, thus siding with those authors that refer to Thematic Content Analysis as a single concept (DeGregorio, Tagliafico and Verde 2018; Braun and Clark 2006).

Returning to the focus of this paper, in essence we are searching for meanings that the interviewee (or document author) is intending to communicate, and in this context it is useful to borrow the concept of meaning-units from the field of phenomenology (Giorgi 1985, 2009). This helps emphasise that it is the meaning that is being sought (not the words), and so "meanings have words" rather than the other way around. Since an array of possible words might be used to convey the same meaning, it is up to the researcher to determine which words in the database might point to a particular meaning being communicated. This underlines the strong argument for including "a human in the loop" when employing software analysis tools (Michaud 2017, Basit 2003). The use of words as "search-clues" to identify key meanings, is elaborated on later under the heading "meanings not words".

With this human intelligence in focus, we then move on to identify the meanings in the qualitative database. Here there are two possibilities.

• First, the researcher could be looking for a particular attribute or construct that is identified in extant theory, and the aim is to ascertain whether that theoretical construct is found in the context under scrutiny.
• The second possibility is that the research project is exploratory, and here the aim is to search the data to see what possible meanings and themes emerge.

An example of the first instance is exploring the 5-dimensional construct of Entrepreneurial Orientation (EO) among managers. EO has dense definitions for each of the 5 dimensions (Lumpkin and Dess 1996, 2001; Dess and Lumpkin 2005) and these can be used to search the interview data in a rigorous and well-defined manner.

The second instance is exemplified by interviews with a range of successful entrepreneurs, to explore and describe the paradigms which guide their thinking, their decision-making, and new venture creation.

In both instances we are searching for meanings, and to do that we need words; words that typify the meaning that we suspect might be in the database. The second, exploratory case is a bit more difficult, but finding potential search-words can be aided by reading and re-reading the transcripts (Corbin and Strauss 2015, Strauss and Corbin 1998).
1.5 Coding the Themes, Meanings

Coding can be described as the process of ‘mining’ the database (Corbin and Strauss 2015) to extract value (meanings), and then placing them in distinct piles (nodes), that all share a common meaning. Nodes can also be regarded as ‘containers’ for the separate themes that are identified from the data.

Sometimes the researcher is the interviewer as well, and in such cases the data coding can be informed by the data gathering, but typically the sheer volume and richness of the data will mean that this analysis is largely at an intuitive level. For example, the research for which this methodology was originally developed had 35 individual transcripts and documents as its database, comprising around 200,000 words.

Manual coding of nodes is still practiced by some experienced researchers. Basit (2003) for instance, describes each node of meaning in the database being assigned a different colour pen, and marked up accordingly. This however limits the number of identifiable themes to the range of available colours (Bengtsson 2016). Clearly this is not ideal, and only suited to smaller qualitative databases.

In contrast to intuitive or manual analyses, this paper agrees with those like Basit (2003) and Bengtsson (2016) that advocate the use of software analysis tools such as NVivo® and many others. First, this helps guide a systematic process that allows the full extent of the data to be analysed comprehensively (Alasuutari 2010), and thus bolsters the aspect of rigour which is especially important in qualitative research. Second, this also ensures a repeatable and verifiable procedure, which again reinforces rigour, while using a well-known and trusted method such as thematic analysis. And thirdly, using software to search the database, sometimes totalling hundreds of pages, reduces fatigue in the researcher, and therefore increases the likelihood that the research will accesses more of the informants' vitality that is recorded in the qualitative database (Michaud 2017).

2. Let the data speak

This is an especially helpful phrase in qualitative research data when seeking to reveal the nuances of new and novel insights. Having gone to the trouble of gathering a range of expert views, it makes sense to explore that value as best we can. Essentially this aim can be achieved by focussing on three main aspects:

- the research informants (respondents)
- meanings not words
- classifying nodes

The word "informants" is one borrowed from Yin (2018) and is a useful synonym for the respondents in a research project, in that it emphasises that these people are there to inform the research, by adding perspectives and insights that are unknown to the researcher at the beginning of the project.

Qualitative databases typically include interviews recorded as transcript documents, and when searching through the database for coherent threads of meaning (nodes), it is also useful to emphasise that the search is for "meanings not words". This will be explained in the next section.

When classifying nodes it is essential to allow the most significant nodes to "emerge from the data" rather than placing an emphasis on nodes that is based on factors external to the data (Vaismoradi et al 2016). This requires a framework which analyses the database itself, and identifies how well each of the nodes are represented in that database. Once this hierarchy of nodes is understood, one can proceed to relate the nodes to one another, and then also to extant theory and practice, to form a coherent model.

2.1 Transcripts to Nodes

2.1.1 Meanings not Words

As mentioned before, when searching for nodes the concept of "meanings not words" is useful especially when a variety of informants or sources are used. The concept of “meaning units” (Giorgi 1985, 2009) was introduced previously. By focussing on the meaning during the analysis, rather than the words, this helps to minimise the drift of meaning across a diversity of contexts in the research. In this way the qualitative researcher can bolster the all-important aspect of academic rigour, when using the ubiquitous method of
thematic analysis. Meaning units therefore become important in extracting coherent meaning from a range of inputs, by helping to identify the different words that interviewees may use to describe one particular meaning (meaning unit).

Take for instance the word "integrity" which typically describes an ethical approach. However in a research project one informer used the same word "integrity" to convey a completely different meaning – to denote "integrity" (cohesion) among multiple aspects of new venture creation. To complicate matters, ethics was also one of the aspects of the research question being investigated. Simply accepting the word "integrity" as a valid contribution to a node (theme) would have resulted in confusion, and nonsense, despite the power of sophisticated computer-aided tools. To maintain academic rigour, these sorts of verbal acrobatics require careful attention from the researcher (Michaud 2017), and "meaning-units" is therefore a useful concept which makes it possible to triangulate qualitative evidence rigorously, despite a diversity of sources in the database.

Using this approach of "meanings-not-words" consistently will therefore allow the researcher to identify a number of significant and coherent meanings within the database, and to code these as a number of individual nodes.

While a simple word-search is often used, backed up by the researcher’s scrutiny for the correct meaning, software analysis tools typically include other more complex search tools as well. One example is a proximity-search for two words that form a single meaning, such as "strategic marketing". The search will then return those passages where the two words are in close proximity, for instance in the same paragraph, or a certain number of words apart. So the meaning could be expressed in the database as "marketing but with a strategic approach" which would be a valid inclusion in that node of meaning, but would not be picked up by a simple word-search for "strategic marketing".

However one of the most important features of the search-tools in analysis software, is their ability to sort through a large volume of diverse documents, with different file types, in one search action. Once the search is complete, it also records a score for that search, indicating the number of passages identified, how many documents contained this theme, etc. As observed before, these tools also do not suffer the human frailty of researcher fatigue, and will faithfully search through hundreds of pages of documents, accurately identifying every occurrence of the search-item. This again helps to bolster the rigour of qualitative research by reducing researcher fatigue and human error, and by increasing the scope and repeatability of the search process.

By using a combination of software searches and attention to the meanings of each theme, the researcher therefore can maintain rigour while assembling a collection of nodes from the database, using the familiar and dependable basis of thematic or content analysis.

Once the researcher has a collection of identified nodes, the next question becomes:

- Among the many nodes of meaning identified, which are the most significant?

This is where the Matrix Sorter comes into play.

2.2 The 2x2 Matrix

The 2x2 matrix was first used by the Boston Consulting Group (BCG) to categorise a portfolio of products within a market context (Henderson 1970). This matrix divided the products into 4 groups: cash cows, stars, pets, and question-marks, which then allowed firms to follow a different strategy for each product type, and thus maximise their profits.

This 2x2 matrix tool has since been used across a range of other contexts (Lowy and Hood 2011), as an aid to strategic, organisational and personal decision-making. Some examples are the 4-quadrant "I'm ok, you’re ok" analysis, the prisoner’s dilemma which presents 4 outcomes of co-operating or not, the sorting of teams into 4 distinct types, and using the virtue matrix to guide corporate responsibility.

The essence of this conceptually simple 2x2 matrix is that two key elements are combined, to form four distinct quadrants, each with a particular character. The power of the matrix derives from the choice of the two elements – it is essential that these two are central to the issue being analysed (Lowy and Hood 2011). The
The 2x2 matrix then examines a combination of scenarios where the two key elements are alternately assumed to be high or low. For instance the BCG focused on market-growth and market-share, so high market-growth and high market-share was clearly the quadrant to be in, and those products were identified as "stars".

By using these two key elements, the 2x2 matrix thus considers 4 distinct scenarios, and giving each one a name helps to guide how each scenario is considered. The complexity of a 4-way decision-making process is thus sequenced, as each of the four scenarios is considered individually, without being confused or clouded by the other 3 alternatives.

The relevance of the matrix to a range of applications can be attributed on the one hand to its largely uncomplicated structure, and on the other to its strong persistence in the business world. Madsen (2017) puts it this way: "The BCG Matrix is simple and powerful, in much the same way as other strategic tools and frameworks such as SWOT analysis and Porter’s Five Forces framework". This certainly locates the matrix in familiar territory, and within a business context. Madsen (2017) also analysed critics of the matrix and noted: "the review of criticism levelled at the BCG Matrix reveals a certain level of scholarly disdain". The article goes on to point out that this academic criticism occurs "in spite of [the matrix] being relevant and widely used by practitioners... [and] still commonly referred to in marketing and strategy textbooks, as well as in MBA education". (Madsen 2017).

Given the above, and especially the underpinning analysis provided by Lowy and Hood (2011), the 2x2 matrix is considered a sound and appropriate basis for this paper that is located in business research methods.

2.3 Applying the 2x2 matrix to node significance

The next step is to apply this 2x2 matrix to analyse the relative significance of the nodes that have been identified in our qualitative data. It is important to emphasise that this approach explores what our experts considered to be more significant, or less so. Organising the nodes into a hierarchy dictated by a particular theory is perhaps convenient, and may be well-received in that it fits within a prevailing paradigm. In some instances this may be appropriate, depending on the research design.

However the power of a qualitative research approach is that it can reveal unexpected aspects of the topic under investigation – facets that were not necessarily predicted or included in extant theory (Davidsson 2005). In this way the existing body of knowledge in that field is fleshed out, added to, and gradually better aligned with real-world observations.

To examine the relative significance of nodes we need two key parameters (Lowy and Hood 2011) that identify the ranking that our experts placed them in. There are any number of parameters that qualitative analysis software can display regarding each of the nodes or themes that have been identified in the database. These include the number of characters, words, paragraphs, passages, documents, references or files that are stored at each node, and each software package uses different terms for these parameters. To narrow down the discussion, three main parameters will be examined, using the labels found in NVivo®.

For each node (theme), the software typically records the following parameters:
- References
- Files
- Coverage

First, the number of references refers to the individual passages that were found to match the theme that has been coded at that particular node. The number of references is irrespective of the source of the passage, so it could be from interviews, promotional material, annual reports, etc. Each time any passage in the database refers to the theme of that node, the number of references for that node increases.

In terms of node significance, the more references that match the theme of that node, the more it was mentioned in the database of documents, and therefore the more significant the node becomes. This is clearly a key parameter in determining the relative significance of the nodes in the data.
Second is the parameter called files. In most databases, a separate file is created for each data source (informant), and if not, it is relatively easy to organise the collected data into an individual file for each data source. Data files could be interview transcripts, company documents, field notes, and so on. The number of files in a particular node will indicate how many different data sources mentioned the particular theme that is contained at that node. If only a few files are found at a particular node, then only a few of the data sources have mentioned the theme at that particular node. In contrast, if many of the data sources have mentioned that theme, then many files will be counted at that node. Again, in terms of significance, having many files at a particular node will indicate that many of the data sources (e.g. interviewees) made mention of that particular theme. This would identify that particular node as more significant than nodes with a low number of files, and that the theme (node) has a diversity of support among many of the research sources. Files is therefore also a key parameter in deciding the hierarchy among our nodes.

Third, the parameter of coverage is expressed as a percentage, and refers to the proportion of a particular document that is devoted to the theme at that node. In a short document, the coverage for a particular node (theme) would typically be quite high, since most of the document consists of just that node. In contrast, a lengthy interview might mention a particular theme many times, but it will be buried in amongst the introductory parts of the interview, the interviewer’s questions, talking about the weather, and other bits of chit-chat. The theme might be very eloquently expressed, many times over, and therefore coded at that node each time it occurs, but the coverage will be quite small, since it makes up only a small percentage of the overall interview.

In the case of the longer interview, the theme is significant, but the coverage percentage is low. In contrast, in a short document the coverage is high, as the passage containing the theme is a larger percentage of the document, but without indicating that the theme (node) itself is any more significant than in the lengthy interview. It is therefore evident that coverage is not a parameter that has a straightforward link to the relative significance among a collection of nodes.

2.3.1 Two key factors for Node Significance

Following the discussion above, the two major factors to emerge which determine node significance are therefore:

- the number of References at that node
- the number of Files at that node

Before proceeding, it is useful to reiterate that in this article the NVivo® descriptors are used as the default. The following summarises the two key parameters and provides examples of different words that are used to denote the same concepts.

- **References** – an instance where the particular meaning is referred to. Alternate descriptions for References are:
  1. Passages
  2. Citations

- **Files** – a record in the database from a specific person or source – with each independent source being recoded in a separate file. Alternate descriptions for Files are:
  1. Sources
  2. Documents
  3. Interviewees

These two key parameters can now be used to evaluate the relative significance among the nodes (themes) in the data. Typically the results for each node will be listed in a table, for instance:
Table 1

<table>
<thead>
<tr>
<th>Node</th>
<th>References</th>
<th>Files</th>
<th>notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>First node</td>
<td>20</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>Second node</td>
<td>80</td>
<td>18</td>
<td>most significant</td>
</tr>
<tr>
<td>Third node</td>
<td>51</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>Fourth node</td>
<td>15</td>
<td>5</td>
<td>least significant</td>
</tr>
<tr>
<td>...etc...</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The number of references can also be expressed as the number of passages coded to that node, the number of “mentions” that a theme received, or the number of times that the theme at that node was cited in the research data.

The number of files can also be stated as the number of different respondents, or sources of information, in which the theme for that node is identified. For instance if most of the interviewees mention a particular theme, then the theme (node) can be regarded as widespread among the experts interviewed for the research.

2.4 The Matrix Sorter – a Hierarchy of Nodes

These two key indicators of relative node significance, are now placed into the 2x2 matrix format in Fig.1. The number of references is located at the top axis of the matrix, and the possible outcomes are that the node will have either few or many references.

Likewise for the number of files identified at each node, there might be many files at each node, or only a few, and this is indicated on the left axis of the matrix.

```
<table>
<thead>
<tr>
<th>Many Files</th>
<th>Few References</th>
<th>Many References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Widely Held View</td>
<td>Well Supported</td>
<td></td>
</tr>
<tr>
<td>Weak Support</td>
<td>The Committed Few</td>
<td></td>
</tr>
</tbody>
</table>
```

Figure 1: Node Significance

Combining these possible outcomes using the matrix therefore gives us 4 distinct scenarios. Each one is given a description that refers to its significance.

The two extreme scenarios are more readily understood, beginning with the scenario in the bottom left quadrant.

In this case the node has few references and also few files, and it is clear that this theme has only weak support from the data that we collected. Not only was this meaning very seldom mentioned, it was also quoted by very few of our sources, and the description Weak Support is therefore apt.

Conversely, the other extreme in the top-right quadrant, identifies those nodes that are referenced many times, and that also have many files which indicates that many different sources in our database identified this theme. Clearly this quadrant contains the most significant nodes in our database – hence the description Well Supported. So these first two node classifications are literally at opposite ends of the scale, and are vital indicators in determining node significance. It is important to emphasise at this point that this difference in significance arises, not from theory, not from the literature, not from existing models, but as signalled by the expertise represented in the database.

The two remaining quadrants are mixed scenarios. Starting at the bottom-right, these nodes have many references, but relatively few files, and this requires a bit of thought. In this outcome only a few sources in the data mentioned the theme, but those that did mention it, were quite enthusiastic about it, and mentioned it many times over. This scenario is characterised as "The Committed Few".
As an example, in a particular research project the issue of ethics was found to be important to certain respondents. However less than half of the interviewees mentioned this aspect, but those that did mention it, emphasised it by repeating it many times over. This type of node classification therefore identifies meanings that are somewhat out of the mainstream, but nonetheless very important to a key subset of the informants. These nodes are quite different to the other classifications, and can often indicate potential sources of new research directions or sub-divisions.

Lastly, the top-left quadrant gathers those nodes that have a large number of different files, but with relatively few references. Again this requires a bit of thought, as the theme was clearly identified by many different sources in the database. However these sources did not refer to the theme (node) very often. An example of this type of node is where "everybody knows". For instance in new venture creation, innovation is a given, and there is no need to repeat it over and over, although nearly all the respondents mention it at least once or twice. This type of node is therefore characterised as a "Widely Held Position". Again, this classification is quite distinct from the others, and offers the researcher yet another option when considering the relative significance of the nodes that have been identified.

This now gives us four distinct categories for our nodes:

- Well Supported
- The Committed Few
- Widely Held View
- Weak Support

Each of these scenarios identifies nodes that are have quite different levels of significance. It is also not a simple winners-losers ranking either – it is a recognition that all of these four categories have a different role in the way they convey meaning to our research project.

For instance one might be tempted to discard the last option of Weak Support as trivial, since it identifies only low endorsement from the informants. However particularly given the rich and complex nature of qualitative data, it can be equally valuable to know what is unimportant, as this can eliminate distraction by fringe-aspects. In this way our qualitative research can remain focused and rigorous, despite the sometimes distracting richness of the data.

2.5 Examples: using the Meanings and Matrix Sorter in practice

2.5.1 Searching for Meanings

First, during the word-search phase there is a need for words that typify the meaning we are looking for. One example is the concept of Risk-taking – does the respondent show instances where risk is acknowledged and managed? The obvious choice is the word "risk" itself, and by using this short four-letter version in the search, the software will also identify related extensions such as "risky" and "riskiness". As outlined before, it is then up to the researcher to determine which of the passages containing that word, also fit with the meaning being sought. This can be thought of as the first level of word-search.

The search can be extended more comprehensively, by exploring other ways that our respondents may have described a risky situation they faced. For instance they might describe a situation that had "a real prospect of failure" – in other words the manager identified a risk-taking event. So by using the search-word "fail" this will identify all the passages containing that root-word and its related extensions (fails, failure, etc.). Repeating this with other possible words related to the meaning will lead to a more exhaustive and richer range of nodes, thus increasing the quality and robustness of the research (Fabregues and Molina-Azorin 2017, Michaud 2017).

A caveat in extending the word-search in this way, is that when moving away from the root-word (risk) the researcher has to increase the level of vigilance in examining the passages that the software identifies, to ensure that they still fit the core meaning of that node (Basit 2003).

2.5.2 Searching for Known Constructs

Then to an example where the Matrix Sorter was used to good effect. A research project mentioned before, explored the Entrepreneurial Orientation (EO) among a range of arts managers using interviews. EO has 5 dimensions, and the expectation is that in most contexts all 5 will be found (Lumpkin and Dess 1996, 2001;
Dess and Lumpkin 2005). Our question was whether this sample of managers would show an entrepreneurial slant (EO), and if so, to what extent?

In this project each EO Dimension was assigned as a Node. The number of Passages (=References above) at each node was counted, and it also tallied how many of the Interviewees (= Files above) mentioned that particular node. This resulted in a table with the two variables for each node (e.g. Table 1), and the table can then be used to create an x-y scatter chart which has the basic format of the 2x2 matrix sorter.

The x-y scatter chart is helpful in that it shows the results graphically, and this can be produced using standard spreadsheet applications. In cases where the node parameters (references, files) have a wide variance (10:1 or more), it can be helpful to use logarithmic axes in the x-y scatter-plot, for a more even spread of nodes.

In Chart-1 below the results are graphed, and the matrix sorter then identifies the relative significance among the five dimensions of EO within this sample of interviewees. As in Figure 1 before, node significance is low in the bottom-left quadrant, and increases towards high significance in the top-right quadrant of the matrix. So in this case Innovativeness (Innov) is clearly the most significant meaning to emerge from these informants, while all the other dimensions show a strong presence, except for one.

The node Compet is firmly in the quadrant of Weak Support (Fig.1) and therefore very much less significant than the rest. This node represents the EO dimension of Competitive Aggressiveness, and therefore showed very clearly that among our respondents, this dimension was "lacking". So while the original intention was just to confirm if the arts managers were entrepreneurial (or not), the answer was much more interesting than imagined. By using the Matrix Sorter we had discovered a small but intriguing anomaly. Without this framework and its reliance on two key significance-factors, it is quite likely that this anomaly would have been missed.

The data therefore confirmed that the respondents were certainly entrepreneurial, but in this case they were not aggressively competitive (with other arts companies), which is not what the EO construct suggests. This key insight later initiated an extension of the research to identify an important counter-dimension of collaboration (rather than competition) among the respondents, all drawn from what might be described as the Arts Community. So although these arts managers showed all the other EO traits, and were running profitable businesses, a key difference was that they achieved this without the need for "competitive aggressiveness".

This is a clear case of the Matrix Sorter rigorously guiding the research, and importantly also in uncovering "the unexpected" – something novel, new, surprising – and thus allowing a vital strength of qualitative research to emerge (Davidsson 2005). This insight also spawned a new line of research that explores potential contextual modifiers for EO. The evidence was always there in the qualitative database, but without the methodical approach described here, it is quite possible that this knowledge would have remained undiscovered.
2.5.3 Searching for the Unknown

In another case example, this time in an exploratory research context, the Matrix Sorter was used to construct a model from the nodes of meaning uncovered in the database. In this case the coding had identified 20 individual themes, gathered into nodes, and the question was then how to organise them into a coherent model. Importantly, that model had to represent the emphasis that the respondents had placed on each meaning. In this case the key factor named "references" in this article was termed "passages", and the factor that this article calls "files" was termed "documents". The raw results of the thematic content analysis are shown in the table below. Only the two right-hand columns are relevant here.

Table 6.2: Extent of Support for each Meaning Unit (Node)

<table>
<thead>
<tr>
<th>Node</th>
<th>Characters</th>
<th>Paragraphs</th>
<th>Passages</th>
<th>Documents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ambiguity, Risk Tolerance</td>
<td>7609</td>
<td>19</td>
<td>18</td>
<td>8</td>
</tr>
<tr>
<td>Cool with Chaos</td>
<td>3678</td>
<td>13</td>
<td>13</td>
<td>8</td>
</tr>
<tr>
<td>Don't Tell yet</td>
<td>2773</td>
<td>15</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Flexibility</td>
<td>5926</td>
<td>33</td>
<td>24</td>
<td>11</td>
</tr>
<tr>
<td>Holism</td>
<td>1536</td>
<td>19</td>
<td>15</td>
<td>10</td>
</tr>
<tr>
<td>Look Ahead</td>
<td>5856</td>
<td>54</td>
<td>51</td>
<td>15</td>
</tr>
<tr>
<td>Market Feedback</td>
<td>9649</td>
<td>88</td>
<td>85</td>
<td>23</td>
</tr>
<tr>
<td>Money Matters</td>
<td>14232</td>
<td>127</td>
<td>139</td>
<td>26</td>
</tr>
<tr>
<td>Multiple Experts</td>
<td>12034</td>
<td>67</td>
<td>66</td>
<td>16</td>
</tr>
<tr>
<td>NL Approach</td>
<td>10488</td>
<td>73</td>
<td>72</td>
<td>18</td>
</tr>
<tr>
<td>NL Environment</td>
<td>8956</td>
<td>78</td>
<td>79</td>
<td>24</td>
</tr>
<tr>
<td>Non-Linear</td>
<td>44159</td>
<td>144</td>
<td>97</td>
<td>27</td>
</tr>
<tr>
<td>Pattern Recognition</td>
<td>5798</td>
<td>31</td>
<td>28</td>
<td>11</td>
</tr>
<tr>
<td>Ratchet in Assumptions</td>
<td>1866</td>
<td>11</td>
<td>12</td>
<td>8</td>
</tr>
<tr>
<td>Relational Balance - Fit</td>
<td>28040</td>
<td>117</td>
<td>83</td>
<td>20</td>
</tr>
<tr>
<td>Relationships in NVs</td>
<td>3526</td>
<td>29</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td>Strategic Minimalism</td>
<td>7038</td>
<td>40</td>
<td>31</td>
<td>13</td>
</tr>
<tr>
<td>Sustainability &amp; Ethics</td>
<td>9597</td>
<td>30</td>
<td>26</td>
<td>11</td>
</tr>
<tr>
<td>The Entrepreneur</td>
<td>16751</td>
<td>75</td>
<td>72</td>
<td>19</td>
</tr>
<tr>
<td>What's Changed</td>
<td>5744</td>
<td>58</td>
<td>56</td>
<td>18</td>
</tr>
</tbody>
</table>

NVivo® allows these "free" nodes to be organised into main (parent) nodes with sub-nodes (siblings), and if required, further levels of sub-nodes. While every node has a distinct meaning, carefully defined in the research, there are also some relationships among the nodes that emerged. In this context the hierarchy of the nodes is paramount, so that significant nodes are brought to the fore. To this end the raw data above is again entered into an x-y scatter-graph, as shown in Chart 2 below.
The chart uses cryptic abbreviations to identify the nodes, but the important observation here is that the random assortment of free nodes (Table 6.2) has now been arranged in a clear hierarchy. The low-ranking nodes are in the bottom-left corner, with significance increasing to its maximum in the top-right corner. It is also interesting to note that in this case there are no nodes of meaning in either of the outlier (high-low) categories of "The Committed Few" or "Widely-held View".

Having this hierarchy revealed now means the nodes can now be organised into a clear and cogent model, with each meaning taking a place that depicts the significance accorded it by the research informants. Without the framework of this matrix, it is hard to know how the researcher would assign the relative importance of nodes in accordance with the priorities of the research informants.

To illustrate this example further, the resulting model is shown below, with some nodes clearly more significant than others, and yet inter-related as well. It is not necessary to understand the model itself, only its structure, and the logic behind that structure. Importantly, the structure is guided by the Matrix Sorter, which in turn relies on the relative importance of each node that emerged from the database.

To explain the structure and the inter-relationships a bit further, the node and sub-node arrangements are helpful. Within each node-square, the numbers below the description indicate the significance of the node itself, while the numbers above add the significance contributed by its sibling nodes as well. For instance to be "Cool with Chaos" requires "Ambiguity and Risk Tolerance" and an attitude of "Don't Tell Yet" (because things are fluid and unconfirmed). The parent node has a score of 13 itself, but since it is supported by the two sub-nodes (score: 18 and 5), these two siblings inform the parent node, adding to its meaning and significance. The node "Cool with Chaos" thus ends up with a score of 36 (=13+18+5).

Similarly the parent node of "Strategic Minimalism" (score 31) is informed by the sub-nodes "What's Changed" (56), "Look Ahead" (51), and "Ratchet in Assumptions (12)". Its significance-score therefore becomes 150 (=31+56+51+12). This linked significance therefore places "Strategic Minimalism" above the sub-nodes below.

Finally, to explain the relatively high positioning of "Holism" and "Ethics", despite their low scores, this relies on strong support from the literature. To acknowledge their lesser support from the database respondents, they have been placed in an "auxiliary" or help-function, similar to what is found in some organisational charts. In this way the contribution of the literature and the research informants are combined into a coherent picture.
2.6 Integrating the Node Hierarchy

As illustrated in the examples above, the focus on thematic meaning and the matrix sorter produce a hierarchy of nodes, and by making this the initial analysis, the data has been given the opportunity to "speak first". These data-based results can then be integrated coherently into the overall context of the research project. This integration would typically include comparison with existing theory frameworks, matching with industry experience, and questions around how this knowledge might be applied in practice. The final illustration above (Holism and Ethics nodes) is an example of where the database significance is integrated with the knowledge in the extant literature.

3. Implications:

Qualitative researchers have many challenges to deal with (Michaud 2017, Yang 2013, Neergaard and Ulhøi 2007) but for the persistent the rewards are many (Mir 2018, Davidson 2005). To that end, the method and matrix sorter outlined here, is offered as a pathway forward, enabling the qualitative researcher to follow a rigorous and systematic process. This starts with a pile of documents and transcripts, and arrives at a coherent hierarchy of meanings (nodes). All the while this helps the researcher to reconnoitre the valuable content itself – from the research informants – and to reveal potential insights, confirmations and discoveries that are somewhere in the qualitative database (Alasuutari 2010).

The paper has therefore presented two ways in which to increase rigour while using the familiar method of thematic content analysis, which is especially important in qualitative research.

- Utilising the concept of meaning-units to maintain thematic focus during the search for themes in the qualitative database
- Using software analysis tools during the search-phase to reduce researcher fatigue, and to increase thoroughness and consistency when identifying key meanings in the qualitative database

Also, the paper has developed a simple but powerful matrix which sorts the identified themes into four distinct categories, with associated characteristics that enable the researcher to arrange the themes in a way that reflects the data.

Finally, this framework has been illustrated using actual research examples, to flesh out both the utility and capability of the framework.
Overall, the paper outlines how themes can be coherently identified and then categorised in a clear hierarchy, offering some confidence that the data itself has been "allowed to speak". The power of the 2x2 matrix sorter has been used to explore the unique value in the qualitative database itself, independent of the expectations of theories or experience. Importantly, this has retained "the human in the loop" to ensure that the results still rest on the expertise of the researcher(s), rather than on a simple machine output (Vaismoradi et al 2016). The use of computer-aided searching and counting has been leveraged to enhance the researcher's skills of perception and understanding (Basit 2003). The expert views that have so carefully been gathered in the database, are now analysed and organised in a way that is both scientifically repeatable and academically rigorous (Fabregues and Molina-Azorin 2017).

4. Conclusion

Many researchers, including those referred to in this article, identify the typical obstacles facing those that choose the qualitative route. These can range from "it's a quant world", to the sometimes overwhelming richness of the data, and even identifying that qualitative approaches are often driven by a particular worldview that questions and challenges conventional group-think.

The method and matrix-sorter is therefore presented as one way to address these obstacles. There are some quantitative aspects to the technique which will hopefully appeal to those with a penchant for that approach, while importantly also being useful in analysing qualitative databases. The attention to meanings-not-words and software tools can assist in maintaining focus through the richness of the data, and the matrix-sorter presents a relatively straightforward yet powerful categorisation of nodes. Finally, for those that have chosen the qualitative approach and worldview, it is hoped that the framework outlined here will help to open the path ahead somewhat, and support the continued exploration of new knowledge and valuable insights.

References

Bengtsson, M., 2016. How to plan and perform a qualitative study using content analysis, International Journal of Nursing Practice, Research and Education, 2, pp.8-14
Braun, V., and Clarke, V., 2006. Using thematic analysis in psychology, Qualitative Research in Psychology, 3(2), pp.77-101
Formative Measurement Scale Development: An Example Using Generalized Structured Component Analysis

Dawn Hall¹ and Joshua D. Shackman²
¹Tetraverta, Fayetteville, NC, USA
²California State University Maritime Academy, Vallejo, CA, USA
dmhall.wk@gmail.com
Tetraverta@gmail.com
JShackman@csum.edu
DOI: 10.34130/JBRM.18.1.002

Abstract: Unlike reflective measurement scales, the steps for development of formative measurement scales tend to be highly subjective and involve mostly the judgment of the researcher. Formative scales have been criticized for this reason. This paper extends Christoffersen and Konradt’s (2008) method of joint development of a formative and reflective scale to assess mutual validity of each scale. We utilize a second order method to reduce measurement error in the formative scale as suggested by Edwards (2011), and test the efficacy of Generalized Structured Component Analysis (GeSCA) for this purpose. For illustrative purposes, we utilize a sample of formative and reflective job satisfaction survey data both to test our joint formative/reflective scale development technique and to assess which formative aspects of job satisfaction align with commonly used reflective job satisfaction scales.

Keywords: Formative measurement scales, job satisfaction, Generalized Structured Component Analysis

1. Introduction

There has been a great deal of controversy in recent literature regarding the development and utilization of formative measurement indices, culminating in special issues of Psychological Methods in 2007, the Journal of Business Research in 2008, and MIS Quarterly in 2011, and the Academy of Marketing Science Review in 2013. In spite of the extensive commentary and conceptual discussion about the appropriateness of the use of formative versus reflective scales, there has been little practical guidance and research available for researchers wishing to develop formative scales.

This paper presents a simple empirical example of a practical method to develop and validate a formative scale that addresses some of the criticisms and limitations of formative measurement. Edwards (2011) points out that whereas the development of reflective scales involves standardized steps and statistical tests for validation such as factor analysis for assessing construct validity or Cronbach’s Alpha for internal consistency, these steps do not apply for formative scales. Selection of initial items for both reflective and formative scales involve similar steps such as a literature review and the judgment of experts, but steps to evaluate the appropriateness of a set of formative items are not nearly as clear or well established. Furthermore, Edwards (2011) points out that the weighting of items in a formative scale depends on the choice of other variables in the model which lead to problems of interpretation and lack of consistency between studies. The method outlined in this paper is intended to address the issue of empirically assessing the validity of formative items using objective criteria.

Our paper makes three main contributions to the literature on formative scale development. The first contribution is that we extend the work of Diamantopoulos and Siguaw (2006) and Christoffersen and Konradt (2008) regarding the development of standardized steps for developing and validating a formative scale. Like these authors, we illustrate a method of jointly developing both a formative and reflective scale for the same construct as a way of jointly validating both scales. However, we extend the approach of these authors by utilizing a second-order measurement model for both the reflective and formative scales. More specifically, we measure the formative scale through second order formative dimensions each measured by first order reflective items. This second order approach has been proposed conceptually by Mackenzie, Podsakoff, and Jarvis (2005) and Mackenzie, Podsakoff, and Podsakoff, (2011). Similarly, Edwards (2011) has proposed a similar second order approach as a way of reducing measurement error for formative constructs. While such an approach has been proposed conceptually, little or no research has been done to illustrate such a method empirically.
Second, this paper contributes to research on formative scale validation by using a relatively new path modeling technique called Generalized Structured Component Analysis (GeSCA) (Hwang and Takane, 2004). This method was designed to incorporate advantages of both traditional covariance based structural equation modeling (CBSEM) and partial least squares path modeling (PLS), and has some unique features that are ideal for formative scale development. More specifically, we employ the ability of GeSCA to estimate mixed-indicator measurement models with second order formative dimensions and first order reflective items. While studies have been done testing GeSCA’s capabilities both Monte Carlo simulations (Hwang, et al., 2010; Henseler, 2012) as well as empirical data (Tenenhaus, 2008), little or no research has been done utilizing GeSCA’s second order measurement model capabilities.

A final contribution of this study is specific to the measurement of job satisfaction. This paper uses job satisfaction as an example for this method due to the common use of both formative and reflective items in measuring this construct, and because job satisfaction is a widely studied construct in organizational research. While the primary purpose of our study is to illustrate a generalizable method of formative scale development rather than to specifically investigate job satisfaction measurement, this study also provides useful data and results that shed light as to which formative aspects of job satisfaction line up the closest with commonly used reflective job satisfaction scales.

2. Literature Review

Coltman, et al. (2008) point out that while reflective measurement scales are dominant in psychology and management, the use of formative measurement scales are common in economics and sociology. A major conceptual difference between formative and reflective scales include the notion that in a reflective scale the direction of causality runs from the construct to the scale items, whereas in a formative scale the indicators cause the construct. All items in a reflective scale are expected to measure the full domain of the construct, and the meaning of the construct should not change if an item is removed. By comparison, in a formative scale each item is measuring a specific aspect of the construct and the domain of the construct will change if an item is removed.

A key empirical implication of a reflective scale is that all items should be intercorrelated since each item should be measuring the full domain of the construct. Traditionally, standard methods for evaluating measurement scales such as structural equation modeling, Cronbach’s alpha, item response theory, or factor analysis have all assumed the use of reflective indicators (Bollen and Diamantopoulos, 2017). But given that formative scales do not need to be intercorrelated, these traditional methods will not be appropriate for validating a formative scale.

Diamantopoulos and Siguaw (2006) outline a four-step method for development of a formative scale. They illustrate these steps by identifying 30 items intended to measure export coordination. In order to divide these items into formative or reflective items, they use steps such as factor analysis or tests for multicollinearity. Items found to have a low degree of multicollinearity with each other are identified as formative, whereas items that load heavily on a single factor are identified as reflective. Finally, items chosen as formative were incorporated in a Multiple Indicators Multiple Causes (MIMIC) model to assess the correlation between these formative items and a latent variable measured by two reflective indicators of export coordination. Formative items were retained or removed based on optimization of the model fit.

Christophersen and Konradt (2008) use a similar approach as Diamantopoulos and Siguaw (2006) to jointly develop a formative and reflective scale to measure online store usability. However, Christophersen and Konradt’s (2008) approach has two important differences with Diamantopoulos and Siguaw (2006). First of all, they establish a priori which items are expected to be formative and which ones are expected to be reflective rather than start with a common pool of items. This has the advantage of starting the process with an attempt to identify conceptually which formative aspects should be included in the scale. Second, instead of using a MIMIC approach they estimate separate reflective and formative scales. Validity of individual items are assessed in the formative scale by their statistical significance, and the validity of the formative scale is assessed by its path coefficient to the reflective scale of the same construct as well as the path coefficient to an outcome variable (purchase intention). A criticism of the MIMIC approach is that the interpretation can be difficult and it is not clear if the formative items “measure” the latent variable or are separate independent variables that predict a reflectively measured variable (Diamantopoulos, 2011). By estimating separate reflective and formative latent variables, the interpretation is clearer and the precise level of correlation between the two
scales can be measured. In addition, the ability of the formative and reflective scales to predict a criterion variable can be directly compared.

Edwards (2011) criticizes the use of formative measurement scales due to their inability, unlike reflective scales, to attenuate for measurement error. He proposes an alternative method whereby each formative item is measured by a multi-item reflective scale in a second order measurement model. This approach not only allows for formative items to be incorporated into a path model, but also allows for a combined reflective/formative approach that combines the advantages of both. Similarly, Mackenzie, Podsakoff, and Jarvis (2005) and Mackenzie, Podsakoff, and Podsakoff, (2011) propose the use of mixed-indicator second order measurement models with multiple formative dimensions each measured with a reflective scale. Christophersen and Konradt’s (2008) approach does involve the use of multiple items to measure each formative aspect. However, instead of using a second order approach they first reduce the number of items in their formative scale using factor analysis.

Christophersen and Konradt (2008) use Smart-PLS for their model estimation, which does not allow for second order measurement models and thus the Edwards (2011) approach cannot be applied. While most other PLS software packages have built in formative latent variable features, they typically do not have second order measurement model capabilities. Generalized Structured Component Analysis (GeSCA) software (Hwang and Park, 2009) allows for second order latent variable computation, and thus it can be used to implement Edwards’ (2011) method.

Diamantopoulos (2011) points out that it is a common misperception that only PLS can estimate formative latent variables and that this cannot be done using CBSEM. He outlines an approach to estimating formative latent variables using CBSEM and points out several advantages of using CBSEM over PLS such as the ability for CBSEM to compute fit statistics and thus better refine the formative measurement model. However, he also points out several challenges of formative measurement with CBSEM such as identification issues and the need to use programming language to implement some formative estimation techniques. Unlike PLS, GeSCA does compute a fit statistic but retains advantages of PLS such as built-in formative measurement capabilities, a lack of identification issues, and low sample size requirements (Hwang and Takane, 2004). Thus GeSCA is an attractive option for development of formative measurement scales. The next section of this paper will outline our methodological approach for evaluating a hybrid reflective/formative measurement model using GeSCA.

3. Methodology

3.1 Study Population

This study was conducted utilizing U. S. military, NATO and Department of Defense civilian personnel throughout the Naples, Italy area during May and June of 2012. While this is a narrow and specialized sample, it does have its advantages for collecting job satisfaction data in that the military is one of the only large organization that tends to have required or obligated service. This means that unlike most civilians, they cannot leave a job that they dislike until their fixed term of employment ends. So unlike civilian populations, a military oriented population may have less of a sample selection bias issue resulting from unhappy employees leaving their jobs. Harvey, Billings and Nilan (1985) used National Guard members in a similar fashion when evaluating the Job Diagnostic Survey.

3.2 Formative Indicator Selection

The sources of reflective and formative job satisfaction items are adopted from five commonly utilized job satisfaction measurement instruments which include, the Minnesota Satisfaction Questionnaire (MSQ), the Job Descriptive Index (JDI), the Job Diagnostic Survey (JDS), the Michigan Organizational Assessment Questionnaire-Job Satisfaction Subscale (MOAQ-JSS), as well as the Job Satisfaction Survey (JSS) (see Van Saane, et al. (2003) for a review of these instruments).

Table 1 below gives an overview of the facets covered in these scales. The most commonly used items involve pay, coworkers, growth, supervision, and general reflective items. The MSQ covers 20 different formative aspects, but the other instruments are far more parsimonious. The JSS includes nine formative aspects, and the remaining three include five facets each.
Table 1: Job Satisfaction Survey Indicators

<table>
<thead>
<tr>
<th>MSQ</th>
<th>JDI (JIG)</th>
<th>JDS</th>
<th>JSS</th>
<th>MOAQ-JSS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ability Utilization</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Achievement</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Activity</td>
<td>Work</td>
<td>Nature of Work</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Advancement</td>
<td>Promotion</td>
<td>Promotion</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Authority</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Company policies</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Compensation</td>
<td>Pay</td>
<td>Pay</td>
<td>Pay</td>
<td>Extrinsic Rewards</td>
</tr>
<tr>
<td>Coworkers</td>
<td>Coworkers</td>
<td>Social</td>
<td>Coworkers</td>
<td>Interpersonal Relations</td>
</tr>
<tr>
<td>Creativity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Independence</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moral Values</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recognition</td>
<td></td>
<td>Contingent Rewards</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Responsibility</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Security</td>
<td></td>
<td>Security</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social service</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social status</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Supervision (Human relations)</td>
<td>Supervision</td>
<td>Supervisor</td>
<td>Supervisor</td>
<td></td>
</tr>
<tr>
<td>Supervision (technical)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variety</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Working Conditions</td>
<td></td>
<td>Operating Conditions</td>
<td>Work Environment</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Growth</td>
<td>Intrinsic Rewards</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Communication</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fringe Benefits</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Miscellaneous</td>
</tr>
<tr>
<td></td>
<td>JIG-General</td>
<td>General</td>
<td>Global</td>
<td>JSS-General</td>
</tr>
</tbody>
</table>

Three items for each formative aspect were selected and/or adapted with permission, from an item pool of current scales to represent each of the selected formative indicators. The formative indicators were measured using a seven-point Likert scale ranging from Very Dissatisfied to Very Satisfied. Each statement is prefaced with “At this job, this is how I feel about...”.

3.2.1 Pay

All five instruments listed in Table 1 include pay as a formative facet of job satisfaction. Thus it is a relatively straightforward decision to include pay as one of our formative aspects. We include two items directly related to pay and one item related to benefits. Appendix I includes a list of our formative items and the source of the items – most of which come from the MSQ, JDI, and JSS.

3.2.2 Social

Besides pay, the only other formative facet shared by each of the five instruments in Table 1 relates to coworkers or the “social” aspect of the workplace. We include Social as a formative aspect that encompasses items related to coworkers, community, and working relationships.

3.2.3 Supervision

Four of the five instruments in Table 1 measure some form of satisfaction with supervision. Thus this facet is also a natural choice for inclusion. We include items related to perceived support from the supervisor, praise for good work, and competence of the supervisor.
3.3 Work and Growth

While there is broad convergence among the instruments we reviewed in covering coworkers, pay, and supervision there is considerable divergence regarding other facets of job satisfaction to include. Facets related to work, the working environment, and promotion each measured in three of the instruments in Table 1. Two other facets were measured by two instruments – these facets are related to growth and to job security. We elected to select three items directly related to everyday tasks regarding the work itself based on items found in work or working environment scales and label this facet as Work. And finally, we elected to take items related to promotion, growth, and job security as the Growth facet of job satisfaction. Promotion and growth have been linked as “motivator” aspects of job satisfaction. Furthermore, job security and growth have been linked together in the psychological contract literature as “relational motivators” as opposed to “transactional motivators” that are monetary in nature such as pay and benefits. (Robinson, Kraatz and Rousseau, 1994; Shore and Tetrick, 1994). Taken together, the items related to job security, growth, and promotion can also be considered part of the long-term potential of the employee at the organization.

3.4 Reflective Indicator Selection

Reflective indicators vary depending on the measurement model or the structural model being applied. The first type of reflective indicator is the single item answer, “All in all I am satisfied with my job.” The single item response will be used to test various assumptions and models. The second type of indicator is the adapted JIG, which was amended from the Y/N format to a 7 point Likert scale premised with “At this job, my work is usually...” This phraseology follows the original design, with only the response format differing. The third type of indicator will be the trio of responses from the MOAQ-JSS also on a 7-point Likert scale. Both the adapted JIG, and the MOAQ-JSS will act as reflective indicators for the latent variable as illustrated in the composite model in Figure 2.

3.5 Structural Model Dependent Variable

A three-item scale was used to measure Intent to Leave: “I often think of leaving the organization,” “It is very possible that I will look for a new job next year, or as soon as possible after my current commitment” and “If I may choose again, I will choose to work for the current government organization.” These items were taken were adopted from a scale originally developed by Cammann, et al. (1979) which was also shown to have validity/reliability in additional studies (Chen, Hui and Sego, 1998; Valentine, et al., 2011). Minor modifications were made to the second and third questions to reflect the nature of our military sample population.

3.6 Statistical Analysis

As previously explained, GeSCA is one of the few path modeling packages that allows for second order measurement models and has built-in formative measurement model capabilities. GeSCA is the primary statistical method used to estimate our model. Figure 1 illustrates the basic reflective/formative modeling approach used by Christophersen and Konradt (2008). Figure 2 illustrates the approach we use to estimate our model. This includes five formative aspects at the first order level, each measured by multiple reflective indicators. For the purely reflective scale, we also employ a second order approach with the JIG and MOAQ-JSS scales used at the first order level and each measured by multiple reflective indicators at the second order level. The three item Intent to Leave scale serves as the outcome variable in this model.

![Figure 1: Christophersen and Konradt Reflective/Formative Measurement Model Approach](image-url)
3.7 Halo Effect

Minimizing a potential halo effect was a concern due to development of our formative scales, as a strong halo effect might lead to artificially high correlations between formative aspects and thus multicollinearity. Multicollinearity can bias formative scale estimation (Diamantopoulos and Winkhofer, 2001). Items were then grouped according to each indicator, and not mixed throughout the survey, allowing the individual to concentrate on each conceptualization or subscale before moving to the next idea (Garbarino and Johnson, 1999). Finally, all items were modified to align with a 7-point Likert scale allowing for a larger range of response with greater nuance as recommended by Bownas and Bernardin (1991).

4. Results

4.1 Descriptive Statistics

Of over 320 completed surveys, 307 were considered valid and usable for the study. The population was divided evenly between junior and senior personnel in which enlisted ranks and civilians below the GS-7 pay grade are considered the Junior population, GS-9 civilians or higher were grouped with all officer ranks and designated as Senior in Table 2. In general, most rated their job satisfaction at a level slightly above the neutral of 4 with Senior individuals more satisfied that Junior personnel and less likely to leave. However, the standard deviations in several areas are high, illustrating that those who weren’t ambivalent tended to either love or hate their jobs.

Table 2: Positional

<table>
<thead>
<tr>
<th></th>
<th>JIG</th>
<th>MOAQ-JSS</th>
<th>ITLeave</th>
<th>Work</th>
<th>Pay</th>
<th>Social</th>
<th>Superv</th>
<th>Growth</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>153</td>
<td>153</td>
<td>153</td>
<td>153</td>
<td>153</td>
<td>153</td>
<td>153</td>
<td>153</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>1.35961</td>
<td>1.78121</td>
<td>1.87216</td>
<td>1.44882</td>
<td>1.38055</td>
<td>1.5109</td>
<td>1.86117</td>
<td>1.54655</td>
</tr>
<tr>
<td>Senior</td>
<td>Mean</td>
<td>5.0592</td>
<td>5.0217</td>
<td>3.658</td>
<td>4.8745</td>
<td>5.1061</td>
<td>4.9091</td>
<td>4.7619</td>
</tr>
<tr>
<td>N</td>
<td>154</td>
<td>154</td>
<td>154</td>
<td>154</td>
<td>154</td>
<td>154</td>
<td>154</td>
<td>154</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>1.3383</td>
<td>1.70424</td>
<td>1.84156</td>
<td>1.42142</td>
<td>1.4665</td>
<td>1.39783</td>
<td>1.6342</td>
<td>1.4867</td>
</tr>
<tr>
<td>N</td>
<td>307</td>
<td>307</td>
<td>307</td>
<td>307</td>
<td>307</td>
<td>307</td>
<td>307</td>
<td>307</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>1.35634</td>
<td>1.75694</td>
<td>1.86224</td>
<td>1.43995</td>
<td>1.46058</td>
<td>1.46956</td>
<td>1.76034</td>
<td>1.56498</td>
</tr>
</tbody>
</table>
There was little variation between the military and civilian counterparts as seen in Table 3, although civilians were slightly more satisfied in all areas except Growth. This may be

because there tend to be fewer opportunities for advancement for civilians who would prefer to remain in one area or in one field of expertise. The commands with the highest levels of satisfaction were JFC/NATO at an average of 5.21 and Other Naples Areas at an average of 5.17, who consisted of individuals based in Naples at a different commands or units aside from the larger groups. The one ship included in the study had the lowest overall levels of satisfaction averaging 3.11.

Table 3: Pay Structure

<table>
<thead>
<tr>
<th></th>
<th>JIG</th>
<th>MOAQ-JSS</th>
<th>ITLeave</th>
<th>Work</th>
<th>Pay</th>
<th>Social</th>
<th>Superv</th>
<th>Growth</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>213</td>
<td>213</td>
<td>213</td>
<td>213</td>
<td>213</td>
<td>213</td>
<td>213</td>
<td>213</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>1.4088</td>
<td>1.4420</td>
<td>1.78376</td>
<td>1.86623</td>
<td>1.41637</td>
<td>1.45759</td>
<td>1.7049</td>
<td>1.54865</td>
</tr>
<tr>
<td>Civilian Mean</td>
<td>5.2973</td>
<td>5.1312</td>
<td>3.5709</td>
<td>4.9574</td>
<td>4.7624</td>
<td>4.7695</td>
<td>4.8014</td>
<td>4.2376</td>
</tr>
<tr>
<td>N</td>
<td>94</td>
<td>94</td>
<td>94</td>
<td>94</td>
<td>94</td>
<td>94</td>
<td>94</td>
<td>94</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>1.1390</td>
<td>1.4169</td>
<td>1.65062</td>
<td>1.83594</td>
<td>1.56395</td>
<td>1.50112</td>
<td>1.86608</td>
<td>1.60764</td>
</tr>
<tr>
<td>N</td>
<td>307</td>
<td>307</td>
<td>307</td>
<td>307</td>
<td>307</td>
<td>307</td>
<td>307</td>
<td>307</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>1.3563</td>
<td>1.4399</td>
<td>1.75694</td>
<td>1.86224</td>
<td>1.46058</td>
<td>1.46956</td>
<td>1.76034</td>
<td>1.56498</td>
</tr>
</tbody>
</table>

The highest levels of satisfaction by ranked groups are the Civilian Senior Supervisors (5.74), Junior Civilians (5.35), and Civilian Senior Managers (5.29). The lowest levels of satisfaction were Civilian Supervisors (4.20), and Junior Enlisted personnel (4.39). Strangely, while Civilian Supervisors were the least satisfied, they were also the least likely to leave. Junior Enlisted personnel were the most likely to leave for another job. In general, women were on average more satisfied (ranging 5.08-5.14) then men (ranging 4.71-4.84) and slightly less interested in leaving, 3.41 compared to 3.94.

Table 4 shows the overall averages for the formative job satisfaction aspects with satisfaction being strongest with Pay and lowest in Growth, for the entire sample. Overall intentions to leave were slightly lower than neutral at 3.83 indicating a general confidence in remaining with government service. However, it should be noted that with a standard deviation of 1.86 there appears to be little ambivalence toward the decision.

Table 4: Formative Indicator Results

<table>
<thead>
<tr>
<th></th>
<th>JIG</th>
<th>MOAQ-JSS</th>
<th>Work</th>
<th>Pay</th>
<th>Social</th>
<th>Superv</th>
<th>Growth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Mean</td>
<td>4.899</td>
<td>4.7807</td>
<td>4.7318</td>
<td>4.7742</td>
<td>4.6895</td>
<td>4.5559</td>
<td>4.3062</td>
</tr>
<tr>
<td>N</td>
<td>307</td>
<td>307</td>
<td>307</td>
<td>307</td>
<td>307</td>
<td>307</td>
<td>307</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>1.3563</td>
<td>1.75694</td>
<td>1.43995</td>
<td>1.46058</td>
<td>1.46956</td>
<td>1.76034</td>
<td>1.56498</td>
</tr>
</tbody>
</table>

4.2 Model Testing

Cross loading analysis was conducted to assess the discriminant validity of the scales and can be seen in Table 5. Overall the loading patterns indicate strong convergent and discriminant validity of the scales with most items loading above .7. The item JIG3 was the only item that illustrated a low loading on the JIG construct, but it still had a higher loading on the JIG construct than for any of the other scales indicating that it still related primarily to only the JIG. It also had poor path coefficients throughout all of the models, however because the JIG is being utilized in its entirety as a previously validated scale, the item has not been removed in this study in order to maintain consistency. Both the MOAQ-JSS and the JIG are utilized throughout in their entirety in order to maintain integrity and reliability. The JIG asks the individual to consider their current job in general and “all in all, what is it like most of the time?” Item number 3 is the statement “Ideal”, which considering the military community, may be identifying independently from the actual levels of satisfaction. In other sample populations this term may have less resistance. The JIG4, “Waste of Time,” and JIG14, “At this job, all in all my work is usually...Inadequate,” were the next lowest loading of all the items in the study. This could possibly correspond to a sense of duty overriding feelings of job satisfaction in general.
### Table 5: Crossloading Matrix

<table>
<thead>
<tr>
<th></th>
<th>MOAQ</th>
<th>JIG</th>
<th>Pay</th>
<th>Social</th>
<th>Superviso n</th>
<th>Work</th>
<th>Growth</th>
<th>Intent Leave</th>
</tr>
</thead>
<tbody>
<tr>
<td>MOAQ1</td>
<td>0.9416</td>
<td>0.7903</td>
<td>0.5230</td>
<td>0.7436</td>
<td>0.7029</td>
<td>0.7860</td>
<td>0.5921</td>
<td>-0.7398</td>
</tr>
<tr>
<td>MOAQ2</td>
<td>0.8938</td>
<td>0.7062</td>
<td>0.4547</td>
<td>0.6171</td>
<td>0.6117</td>
<td>0.6610</td>
<td>0.5000</td>
<td>-0.7155</td>
</tr>
<tr>
<td>MOAQ3</td>
<td>0.9022</td>
<td>0.7311</td>
<td>0.5203</td>
<td>0.6565</td>
<td>0.6055</td>
<td>0.7364</td>
<td>0.5251</td>
<td>-0.6942</td>
</tr>
<tr>
<td>JIG1</td>
<td>0.7003</td>
<td>0.8297</td>
<td>0.3656</td>
<td>0.5926</td>
<td>0.5658</td>
<td>0.6907</td>
<td>0.4493</td>
<td>-0.5698</td>
</tr>
<tr>
<td>JIG2</td>
<td>0.6580</td>
<td>0.7969</td>
<td>0.3334</td>
<td>0.5328</td>
<td>0.5704</td>
<td>0.5945</td>
<td>0.4202</td>
<td>-0.5435</td>
</tr>
<tr>
<td>JIG3</td>
<td>0.3260</td>
<td>0.3821</td>
<td>0.1903</td>
<td>0.2696</td>
<td>0.2393</td>
<td>0.3051</td>
<td>0.2512</td>
<td>-0.2443</td>
</tr>
<tr>
<td>JIG4</td>
<td>0.5190</td>
<td>0.6785</td>
<td>0.2508</td>
<td>0.3671</td>
<td>0.4109</td>
<td>0.5236</td>
<td>0.3193</td>
<td>-0.4318</td>
</tr>
<tr>
<td>JIG5</td>
<td>0.6495</td>
<td>0.8194</td>
<td>0.3379</td>
<td>0.5413</td>
<td>0.5395</td>
<td>0.6328</td>
<td>0.4101</td>
<td>-0.5479</td>
</tr>
<tr>
<td>JIG6</td>
<td>0.6022</td>
<td>0.8046</td>
<td>0.2577</td>
<td>0.4869</td>
<td>0.5366</td>
<td>0.5991</td>
<td>0.3729</td>
<td>-0.4731</td>
</tr>
<tr>
<td>JIG7</td>
<td>0.6432</td>
<td>0.7745</td>
<td>0.3838</td>
<td>0.5127</td>
<td>0.5066</td>
<td>0.6050</td>
<td>0.4076</td>
<td>-0.5441</td>
</tr>
<tr>
<td>JIG8</td>
<td>0.6731</td>
<td>0.8292</td>
<td>0.3046</td>
<td>0.5111</td>
<td>0.5170</td>
<td>0.6324</td>
<td>0.3934</td>
<td>-0.5276</td>
</tr>
<tr>
<td>JIG9</td>
<td>0.5316</td>
<td>0.6802</td>
<td>0.2239</td>
<td>0.4601</td>
<td>0.4590</td>
<td>0.4779</td>
<td>0.3550</td>
<td>-0.4197</td>
</tr>
<tr>
<td>JIG10</td>
<td>0.6616</td>
<td>0.8415</td>
<td>0.3189</td>
<td>0.5519</td>
<td>0.5130</td>
<td>0.6471</td>
<td>0.4478</td>
<td>-0.5712</td>
</tr>
<tr>
<td>JIG11</td>
<td>0.6699</td>
<td>0.8320</td>
<td>0.2682</td>
<td>0.5428</td>
<td>0.5430</td>
<td>0.6045</td>
<td>0.3850</td>
<td>-0.5424</td>
</tr>
<tr>
<td>JIG12</td>
<td>0.7512</td>
<td>0.8635</td>
<td>0.4209</td>
<td>0.6301</td>
<td>0.5727</td>
<td>0.7418</td>
<td>0.4801</td>
<td>-0.6243</td>
</tr>
<tr>
<td>JIG13</td>
<td>0.7310</td>
<td>0.8629</td>
<td>0.3428</td>
<td>0.5897</td>
<td>0.5811</td>
<td>0.6753</td>
<td>0.4454</td>
<td>-0.6079</td>
</tr>
<tr>
<td>PAY1</td>
<td>0.4692</td>
<td>0.3551</td>
<td>0.8745</td>
<td>0.3772</td>
<td>0.3216</td>
<td>0.4193</td>
<td>0.4648</td>
<td>-0.3730</td>
</tr>
<tr>
<td>PAY2</td>
<td>0.4418</td>
<td>0.3164</td>
<td>0.7757</td>
<td>0.3906</td>
<td>0.3432</td>
<td>0.3863</td>
<td>0.4910</td>
<td>-0.4505</td>
</tr>
<tr>
<td>PAY3</td>
<td>0.4697</td>
<td>0.3441</td>
<td>0.8704</td>
<td>0.3879</td>
<td>0.3559</td>
<td>0.3923</td>
<td>0.4667</td>
<td>-0.4045</td>
</tr>
<tr>
<td>SOCIAL1</td>
<td>0.5678</td>
<td>0.5080</td>
<td>0.3315</td>
<td>0.8254</td>
<td>0.5763</td>
<td>0.6103</td>
<td>0.4604</td>
<td>-0.4735</td>
</tr>
<tr>
<td>SOCIAL2</td>
<td>0.6802</td>
<td>0.6113</td>
<td>0.4143</td>
<td>0.9161</td>
<td>0.7304</td>
<td>0.6724</td>
<td>0.5417</td>
<td>-0.5795</td>
</tr>
<tr>
<td>SOCIAL3</td>
<td>0.6413</td>
<td>0.5797</td>
<td>0.4247</td>
<td>0.8240</td>
<td>0.6249</td>
<td>0.6342</td>
<td>0.5852</td>
<td>-0.5758</td>
</tr>
<tr>
<td>SUPERV1</td>
<td>0.6487</td>
<td>0.6145</td>
<td>0.3655</td>
<td>0.6882</td>
<td>0.9198</td>
<td>0.6481</td>
<td>0.4886</td>
<td>-0.5772</td>
</tr>
<tr>
<td>SUPERV2</td>
<td>0.6519</td>
<td>0.5795</td>
<td>0.3999</td>
<td>0.7026</td>
<td>0.9375</td>
<td>0.6336</td>
<td>0.5348</td>
<td>-0.6003</td>
</tr>
<tr>
<td>SUPERV3</td>
<td>0.6514</td>
<td>0.6213</td>
<td>0.3601</td>
<td>0.7094</td>
<td>0.9205</td>
<td>0.5917</td>
<td>0.5339</td>
<td>-0.5716</td>
</tr>
<tr>
<td>WORK1</td>
<td>0.6697</td>
<td>0.6462</td>
<td>0.4475</td>
<td>0.6459</td>
<td>0.5901</td>
<td>0.8524</td>
<td>0.4578</td>
<td>-0.5409</td>
</tr>
<tr>
<td>WORK2</td>
<td>0.7752</td>
<td>0.7500</td>
<td>0.3917</td>
<td>0.6896</td>
<td>0.6142</td>
<td>0.8862</td>
<td>0.4924</td>
<td>-0.6690</td>
</tr>
<tr>
<td>WORK3</td>
<td>0.5718</td>
<td>0.5566</td>
<td>0.3778</td>
<td>0.5604</td>
<td>0.5084</td>
<td>0.8140</td>
<td>0.3865</td>
<td>-0.4662</td>
</tr>
<tr>
<td>GROWTH1</td>
<td>0.5283</td>
<td>0.4609</td>
<td>0.4983</td>
<td>0.5464</td>
<td>0.5007</td>
<td>0.4850</td>
<td>0.9251</td>
<td>-0.5117</td>
</tr>
<tr>
<td>GROWTH2</td>
<td>0.5342</td>
<td>0.4623</td>
<td>0.5076</td>
<td>0.5487</td>
<td>0.5306</td>
<td>0.4761</td>
<td>0.9000</td>
<td>-0.5332</td>
</tr>
<tr>
<td>GROWTH3</td>
<td>0.4684</td>
<td>0.4042</td>
<td>0.4543</td>
<td>0.5136</td>
<td>0.4141</td>
<td>0.4014</td>
<td>0.7561</td>
<td>-0.4463</td>
</tr>
<tr>
<td>LEAVE1</td>
<td>-0.6593</td>
<td>-0.5764</td>
<td>-0.3861</td>
<td>-0.5610</td>
<td>-0.5603</td>
<td>-0.5849</td>
<td>-0.4985</td>
<td>0.8886</td>
</tr>
<tr>
<td>LEAVE2</td>
<td>-0.5547</td>
<td>-0.4470</td>
<td>-0.3541</td>
<td>-0.4546</td>
<td>-0.4473</td>
<td>-0.4674</td>
<td>-0.4236</td>
<td>0.8362</td>
</tr>
</tbody>
</table>
| LEAVE3 | -0.7559| -0.6317| -0.4828| -0.5882 | -0.5764     | -0.6216 | -0.5321 | 0.8227       

Overall the discriminant validity appears to be strong in which each of the designed scales as well as those of the previously published JIG and MOAQ-JSS. Due to discriminant validity between the JIG and MOAQ-JSS scales we modelled reflective job satisfaction as a second order construct rather than merge the two scales as one. Figure 3 below presents the model results. For brevity sake only the second order weights and loadings are shown for the formative Specific Satisfaction scale and the reflective General Satisfaction scale. First order loadings (not shown in the figure) were all over .7 except for three items in the JIG scale, two of which were over .6. As expected, the coefficient between the formative and reflective job satisfaction scales is very high at .867. The coefficients between the two job satisfaction latent variables and Intent to Leave are both negative as
expected and significant at the 1% level. However, the reflective latent variable has a larger coefficient of 0.536 compared to -0.274 for the formative scale.

**Figure 3: Model Results**

Some interesting results were found concerning the weights of the five formative indicators. All weights were significant at the 1% level except for Social which was positive and similar in magnitude to Supervision and Pay but failed to reach significance. An unusual result is that the weighting for Work is negative. However, this appears to be an anomaly within GeSCA as the factor loadings for the reflective items used to measure Work were also negative so the negative signs cancelled out. Work appears to be the strongest formative contributor with a weighting of -0.546 which should be interpreted as positive due to the GeSCA anomaly.

While job satisfaction was chosen primarily for illustrative purposes, an interesting result was found that the Social aspect was not found to be significant. Social includes commonly used items in job satisfaction such as relationships with coworkers, so this result is surprising. Given that the sample was a specific one involving military employees, limited generalizability should be given to this result. However, this result is still interesting in that it shows that even with constructs as widely used as job satisfaction it is still not necessarily clear what formative aspects align with the reflective aspects. Furthermore, this result is consistent with at least some prior analysis of formative job satisfaction scales. Kinicki, et al. (2002) and Buckley et al. (1992) used meta-analysis and multitrait-multimethod matrix analysis to assess the comparative validity of the five facets of the JDI, and found coworkers to have the lowest validity of these five facets. Future research using different sample populations can help better refine which formative aspects of job satisfaction should be included both in measurement scales and in the conceptual definition of job satisfaction.

Another interesting result is that the reflective job satisfaction scale had a much larger coefficient when predicting intention to leave than the formative job satisfaction scale. This may be due the fact that the reflective scale involves attenuation for measurement error at both the second and first order level, whereas the formative scale only attenuated for measurement error at the first order level. The result suggests that for pure accuracy of measurement and predictive power, reflective scales may be preferable. However, formative scales also provide more practical information to researchers or practitioners. For example, the formative job satisfaction scale could be used by management to see what formative aspects of job satisfaction such as satisfaction with work or satisfaction with supervision can be improved in order to improve retention. The benefits of the accuracy of reflective scales and the detailed information provided in formative scales is a further argument in favor of joint development of formative and reflective scales along the lines of the method presented in this paper.
5. Conclusion

Methodologically this paper has contributed to the literature by providing an empirical example of the hybrid approach of measuring a scale formatively at the first order level and reflectively at the second order level as proposed by Edwards (2011). The results indicate that this technique is relatively easy to implement and generally provided results that were expected. GeSCA was overall effective other than one anomaly that was easy to interpret and didn't impact the overall results. While this study used job satisfaction for illustrative purposes, future researchers designing a new scale can use this technique to validate a set of chosen formative indicators. Researchers choosing to use existing formative scales may also wish to utilize GeSCA to measure formative latent variables using formative items at the second order level and reflective items at the first order level.

The method in this study has potential to be an improvement over other scale development methods. First of all, Cronbach's alpha, average variance explained and other validity/reliability statistics for reflective scales do not apply to formative scales. In place of these statistics, this method has shown how the statistical significance of each formative aspect in the GeSCA model can be used to assess the validity of each aspect. For example, the lack of significance of the Social formative aspect shows that even widely used formative job satisfaction aspects may not have validity when included in a full model with both formative and reflective scales. Second, this method allows for the joint development of both a reflective and formative scale simultaneously. The result that both the formative and reflective job satisfaction scales correlate high with each other as well as with intent to leave gives a high level of confidence for both scales. For these reasons, new scales can be developed and existing formative scales can be re-evaluated using the method in this paper.

References


**Appendix I: Formative Scale Items**
All items prefaced by “At this job, this is how I feel about...”.

<table>
<thead>
<tr>
<th>Item</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pay</td>
<td>Adopted from MSQ</td>
</tr>
<tr>
<td>“The amount of pay for the work I do”</td>
<td>MSQ</td>
</tr>
<tr>
<td>“How my pay compares with that for similar jobs in other companies”</td>
<td>JSS</td>
</tr>
<tr>
<td>“The benefits I receive (medical, leave, retirement, etc.)”</td>
<td>MSQ</td>
</tr>
<tr>
<td>Coworkers</td>
<td>Adopted from MSQ and Teacher Job Satisfaction Questionnaire</td>
</tr>
<tr>
<td>“The way my co-workers get along with each other”</td>
<td>MSQ</td>
</tr>
<tr>
<td>“The amount of helpfulness provided by the people with whom I work”</td>
<td>JDI</td>
</tr>
<tr>
<td>“The chance to be ‘somebody’ in the community”</td>
<td>JDI</td>
</tr>
<tr>
<td>Supervision</td>
<td>JDI</td>
</tr>
<tr>
<td>“The way my boss backs up his/her employees (with top management)”</td>
<td>MSQ</td>
</tr>
<tr>
<td>“The way my supervisor praises good work.”</td>
<td>MSQ</td>
</tr>
<tr>
<td>“The level of competence my supervisor has in doing his/her job.”</td>
<td>JSS</td>
</tr>
<tr>
<td>Work</td>
<td>JDI</td>
</tr>
<tr>
<td>“The working conditions”</td>
<td>JDI</td>
</tr>
<tr>
<td>“The sense of accomplishment work gives me”</td>
<td>JDI</td>
</tr>
<tr>
<td>“The amount of work expected of me”</td>
<td>JDI</td>
</tr>
<tr>
<td>Growth</td>
<td>JDI</td>
</tr>
<tr>
<td>“The opportunities for promotion.”</td>
<td>JDI</td>
</tr>
<tr>
<td>“When people do well on the job they have a fair chance of being promoted.”</td>
<td>JDI</td>
</tr>
<tr>
<td>“My job security.”</td>
<td>MSQ</td>
</tr>
</tbody>
</table>
Deconstruction of a Multi-Embedded Supply Chain Coordination Problem Using Mixed Methods

Oluka Pross Nagitta ¹ and Marcia Mkansi²
¹Department of Economics and Managerial Sciences, Uganda Management Institute, Uganda
²Department of Operations Management, University of South Africa, South Africa

poluka@umi.ac.ug
nagittaoluka@gmail.com
mkansm@unisa.ac.za

DOI: : 10.34190/IBRM.18.1.006

Abstract: It is without doubt that there are many overviews of mixed methods research in supply chain management. However, there is relatively little research and representation on the application of robust methodological approaches and techniques that take into account the dynamic nature of a multi-embedded and specialised medicine supply chain coordination (SCC) problem. In Uganda, the distribution of artemisinin-based combination therapies (ACTs) involves a multi-embedded supply chain that runs across the macro, market, and micro levels of stakeholders. The multi-embedded levels have created a coordination challenge tied to stock-outs and unavailability of ACTs to the detriment of the patients. This study aimed to: 1) demonstrate how a mixed methods approach facilitated a better understanding of a multi-embedded and specialised supply chain coordination problem, and 2) reveal the major factors for coordinating a multi-embedded supply chain that can improve the availability of ACTs in the general hospitals of Uganda. An exploratory sequential mixed method approach was employed to disentangle the problem. In the first phase, focus group discussions were predominantly used to collect qualitative data, the findings of which acted as the foundation for the quantitative survey questionnaire. A Confirmatory Factor Analysis was used to determine and validate the market and macro supply chain dimensions that emerged from the survey results. The results showed that the most influential supply chain market environment dimensions that affect the availability of ACTs include information sharing with external stakeholders and supply chain interdependence. The results from the macro-environment showed that the socio-cultural, economic, technological, and legal dimensions influence the availability of ACTs. The contribution of this paper advances the use of mixed methods in deconstructing a complex embedded supply chain problem with implications for supply chain academics and practitioners and government bodies.

Keywords: mixed methods, coordination, supply chain management, malaria therapies.

1. Introduction

The availability of medicine in health facilities is a substantial contributory factor to quality health care. However, achieving quality care presents a significant challenge, especially in low-income countries (Khulza and Heide, 2017). Malaria has remained one of the oldest and deadliest diseases of humankind (Shretta, et al., 2017), especially among children under five and expectant mothers (Damien, et al., 2018). Currently, international efforts focus on malaria eradication by rolling out new technologies, and through substantial growth in the political and financial commitment by countries, regions, and their global partners by 2050 (Feachem, et al., 2019), especially in Africa (Alonso, 2016). However, before this ambitious plan can be attained, the supply and distribution of Artemisinin-based Combination Therapies (ACTs) must be efficient. ACTs are the most affordable effective option to treat uncomplicated malaria in many sub-Saharan countries, where widespread malarial parasite resistance to medicines is rife (Khera and Mukherjee, 2019).

As it is difficult to ascertain the number of deaths attributed to the unavailability of ACTs, this may have far-reaching consequences on patients’ lives. It is also regrettable that empirical interventions to improve ACT stocks in Uganda that might reduce eventual mortality rates have also not yet succeeded, as the applied supply chains are perceived to be notably inept (Uganda Bureau of Statistics, 2015). Similarly, the National Planning Authority Uganda (2015) reports that in spite of the significant investments in malaria control, the malaria burden remains high with rampant stock-outs of ACTs.

The growing importance of curbing the occurrence of malaria in the developing world is evidenced by the current research seeking possible solutions to making ACTs available for the treatment of malaria cases in the developing world (Spisak, et al., 2016; World Health Organisation, Malaria Report, November 2017). A common issue emerging from the empirical evidence related to the supply and distribution of ACTs shows a highly embedded
SCC problem that cuts across both macro and market levels (Nagitta and Mkansi, 2015). The embedded complex coordination problem is tied to the unavailability, understock, and or overstock of ACTs in some hospitals (Jahre, et al., 2012). Although previous studies have provided platforms for understanding the factors affecting the levels of medicine stock, unfortunately, there is recognition in the discipline that SCM studies are skewed towards either the positivist paradigm (Nyari, et al., 2010; Kumar, Singh and Shankar, 2013) or normative (literature reviews and theoretical models) and quantitative (modelling and surveys) (Flint, et al., 2012).

At the same time, the environment in which the supply chain phenomena are located is becoming increasingly complex, especially in the public health sector (Yadav, 2015), necessitating balancing the different conflicting goals and interests of politicians, health workers, patients, and citizens (Singh, Kumar, and Kumar, 2016). Hospitals often face the dilemma of designing new appropriate decision-making models that try to find a balance between the different supply chain players (Saltman and Duran 2015) without compromising quality standards through policy shifts, if one research approach is used. In order to accurately describe, truly understand and begin to explain these complex phenomena (such as market and macro supply chain coordination dimensions (SCCD), research streams should include more studies using multiple methods.

Researchers who exclusively choose one approach seriously delimit the scope of their inquiry, and thereby, their ability to contribute to the body of knowledge. Henceforth, there is a need for a more balanced approach to research using inductive research methods (qualitative approach), in addition to deductive methods (quantitative) in supply chain management to reduce incidences of stock-outs, which this study intended to assess. Some scholars have ascertained that literature on the medicine Supply Chain Management (SCM) in the public sector is still very fragmented and under-researched, especially in the context of developing countries (Jain, Dangayach, Agarwal, 2010; Msimangira and Tesha, 2014). Therefore, mixing research approaches may be seen as a candid initiative, where critical thinking, interest, reasoning, experiences and expertise are combined with the purpose of discovering the truth, so as to find solutions to the problems confronting the health supply chain through investigation and analyses.

The mixed approach complements rather than competes: “the limitations of one method can be offset by the strengths of the other method, and the combination of quantitative and qualitative data provide a more complete understanding of the research problem than either approach by itself” (Creswell and Clark, 2011, p.45). The essential purpose of mixing methods in this study was to get the best measurement of SCCD, using the strengths of each of the methods proposed. Babbie and Mouton (2010) claim that the mixed design is appropriate when the problem is persistent and can be used for developing more focused questions. The mixed methods approach was convenient, since neither a qualitative nor a quantitative approach was adequate to fully answer the research questions (Tashakkori and Teddlie, 2010). However, the two methods are irreconcilable because each has its own unique procedures of gathering and analysing data, despite the fact that they have different strengths and logic (Saunders, et al., 2016).

Therefore, this paper aimed at answering the following questions: i) How does the mixed methods approach enable the understanding of a multi-embedded supply chain coordination problem from the general hospital perspective? and ii) What are the major dimensions that hospitals should consider when coordinating a multi-embedded supply chain at the market and macro environments?

This investigation contributes to the SCM literature related to overcoming stock-outs in the hospital contexts in three aspects: (i) the application of a mixed methodology for making a diagnosis, and following from this (ii) the identification of critical dimensions for state interventions, and (iii) the use of the study findings for policy uptake at the local level.

In section two, we present a brief review of the theoretical and literature on SCC. Thereafter, we provide a contextual background to the medicine supply and distribution problem. In section three, we provide an introduction to mixed methods to contextualise this discussion with reference to SCM issues, while providing the steps to deconstruct the medicine chain, the attendant results, conclusion, recommendations and finally the limitations of the study.
2. Theoretical and literature review

A multi-embedded SCC problem is recognised within the existing literature and remains unexplored in developing countries (Singh, Vaish and Singh, 2014; Yadav, 2015). While the complexity of embedded supply chains has been greatly acknowledged in the health care sector, complementary scientific methodological approaches for addressing such a challenge remain unexplored. This study is guided by the coordination theory advanced by Malone and Crowston (1994). The theory posits that every organisation should continually identify tasks with their interdependencies (referred to as coordination dimensions) among the different actors. It presupposes that coordinating the supply chain is a critical element that ensures that all activities are systematically glued together for the achievement of joint supply chain performance. The insights and conclusions drawn from these articles have gone a long way in terms of mapping ‘what we know and don’t know’ about SCC, but have fallen short in terms of a more nuanced analysis of how SCC exhibits itself in developing markets, and moreover, in the public sector using a mixed methods approach.

Supply-chain coordination is defined variously. What is common to most of the definitions is that it is the process of managing interdependencies, the flow of products, resources, and information among channel partners to end users (Huo, Zhang and Zhao, 2015). Despite the burgeoning interest in SCM by scholars in the private and manufacturing sectors, its application in the supply and distribution of medicines in Africa remains sparse, offering no empirical validation (Msimangira and Tesha, 2014), and yet hospitals are one of the major weak points in the healthcare SCM (Dobrzykowski, et al., 2014).

Typically, the hospital supply chain includes the internal chain (patient care units, hospital warehouse, patients, etc.), while the external chain consists of distributors, vendors and manufacturers. Hospitals are bolstered by a range of health care SCM activities, such as forecasting, quantification, procurement, storage and dispensing (USAID | Task Order 1, 2011). Investigating how hospitals coordinate with their external partners in the market and macro environments through the application of a mixed methods approach may improve the efficiency of these activities, and can provide opportunities for health systems to increase the quality of care and improve the availability of medicine (Office of the Auditor General, 2015).

In Uganda, the medicine supply chain starts with individual hospitals that forecast, quantify, plan, procure and place orders with National Medical Stores (NMS), which is responsible for delivering all medicines and supplies to all public health facilities on a bimonthly distribution schedule under the pull system. The importance of exploring the context interdependence of SCC has been accentuated in recent years with calls for closer attention to the peculiar institutional arrangements, or national business system configurations of developing countries, which may ultimately lead to a different understanding of SCCD (Nagitta and Mkanis, 2015).

In a multi-purpose study, Gao, Li, and Kang (2018) attempted to classify SCCD (e.g. price discount, information sharing, collaboration through activities, such as vendor-managed inventory and consignment stock, and supply chain integration. Other scholars considered collaborative working and joint planning, the mutual exchange of information and integrated information systems, and cross-coordination at several levels in the firm’s supply chain (Ramanathan and Gunasekaran, 2014; Soosay and Hyland, 2015). Although these examples demonstrate some of the processes of supply chain coordination, the scholars do not explicitly describe all the dimensions of coordination mechanisms from a hospital perspective. From these studies, it appears that gaining insight into the other dimensions requires the use of robust methodological approaches and techniques that take into account the dynamic nature of health supply chains.

To deconstruct a multi-embedded SCC, an exploratory sequential mixed research design was adopted to gain a profound understanding of the phenomena in the SCCD of the medicine supply chain through the balance of the rich data that was gained from the qualitative approach and tested through the quantitative approach. For this to occur, we adopted the exploratory sequential mixed methods within the research stream. The mixed methods paradigm fits together with the insights provided by the quantitative and qualitative research into a “workable solution” to answer the multifaceted research questions (Tashakkori and Teddlie, 2010) and offers a practical solution in the ‘real world’ (Feilzer, 2010). Most notably, there seems to be no tool to measure supply chain coordination (Kaur, Kanda and Deshmukh, 2015).

Based on the research questions stated above, we first adopted the qualitative techniques. Similarly, to validate the developed instrument, we used a post-positivist philosophical paradigm, which called for the use of
quantitative methods (Creswell and Clark, 2011). The remainder of the paper discusses the procedures involved in collecting and analysing data using mixed methods.

3. Research approach: Sequential mixed application to the embedded problem

This study used Lee’s (1991) model because it details twelve (12) clear steps that were followed in carrying out research using the mixed methods approach, unlike the one proposed by Creswell and Clark (2007), which seemed more summarised.

3.1 Stage 1: Defining of the strategy context

Nagitta and Mkansi (2015) affirm that it is crucial to have a clear understanding of the environment in which a supply chain operates, as well as the external forces influencing its level of service delivery. Within a country, the medicine supply chain may operate at different levels, namely the central, regional/district, and service delivery points. Besides, various vertical supply chains may be working within a country, with many aspects of intersection and overlap, and a diverse set of stakeholders. Therefore, the inclusion of factors external to the hospital, may complement public sector health-based treatment channels.

In support, Fahey and Narayanan (1986), perceive the market and macro-environments as being instrumental and part of the significant management environment. From these environments, businesses are influenced by several variables emanating from one environment, or sometimes a combination of them all. From the market environment, there are stakeholders (employees, shareholders) and role players (customers, suppliers, competitors, new entrants, opportunities, and threats).

Lastly, the macro-environment factors that were considered include: political, economic, social-cultural, and technology and legal. These factors were examined in the current study. From the results, we note that the coordination frameworks for the supply and distribution of ACTs demonstrate the interplay of all the business management links that describe and analyse the supply and distribution of ACTs.

3.2 Stage 2: Review of literature

To address the SCC gap, the paper firstly, systematically reviews the literature on the SCC determinants; secondly, it argues for the use of mixed methods research to address the questions related to medicine coordination. However, the market and macro settings are ignored in many coordination frameworks. Therefore, it was presumed that if correctly done, SCC may facilitate the integration of chain activities, resulting in better performance (Shukla, et al., 2013).

Several authors, such as Singh (2011) and Kaur, Kanda and Deshmukh (2015), tried to examine the role of coordination in SCM from the private sector perspective. The first author attempted to answer the question, “what are the major factors for coordinating the supply chain?” To do this, the author first explored the literature, then eventually, a coordination index was developed and evaluated using the interpretive structural modelling (ISM) approach. Ultimately, the fundamental relationship among the different factors of coordination and responsiveness in the supply chain was highlighted, which aids strategic decisions. The framework was further tested using a case study. Hence, the inductive method. From the literature, a total of 32 enablers for coordination in a supply chain were identified. These were further grouped into six categories, such as top management commitment, organisational factors, mutual understanding, the flow of information, relationship and decision-making, and responsiveness. All six elements had a strong mutual linkage, although top management commitment was the primary driver for improving the coordination among these factors. Arshinder, Kanda and Deshmukh (2011) conducted a systematic literature review on the importance of SC coordination. The resultant conceptual model helped to highlight the various surrogate measures of SCC.

However, these studies fell short on specifying the market-and macro-factors to be considered in managing the interdependencies in the context of a multi-embedded supply chain. From the above discussion, it can be concluded that there was a need to broaden the theoretical domain of SCC within the medicine field. The general hospital focus with a holistic orientation, including the macro and market environment, creates a superior supply and distribution of ACTs in developing countries, given the dynamics in the broader economy (rooted in national, regional, and international contexts).
A content analysis was conducted by exploring journals from 1960 to 2017. Within the specified period, this study analysed coordination frameworks originating from organisational scholars such as Thompson (1967) and Malone and Crowston (1994). The rationale was to demonstrate how the dependencies have been coordinated in SCM, and also to find out whether they aptly addressed linkages between the market and macro environments. Leading journals were consulted, as these are believed to impart current trends and are held in high esteem (Üsdiken, 2014). The focus of this study is Africa, as it falls within the developing countries’ classification (Nielsen, 2013). Therefore, from a critical realist stance, this study examined previous studies that have applied coordination frameworks within the African context, with implications linked to the market and macro environments. The purpose of this qualitative approach was to identify only those studies that were concerned with the ACT supply and distribution in Africa. Therefore, the focus of this study was to use the coordination theory as a basis for making generative models.

Also reviewed were published reports and scientific publications from Malaria Journal. The journal provides research on several drug-efficacy studies that have been conducted, with examples from Uganda, Tanzania, Kenya, Malawi, Sudan, Ghana, and Nigeria, among others, in Africa. For experiences related to policy implementation, we reviewed several reports and publications (e.g., United States Agency for International Development (USAID). The findings are presented in Table 1. These indicate that the hospitals’ coordination frameworks for supply and distribution of ACTs should be inclusive of all forms of joint work from the market and macro environments.

3.2.1 Market environment dimensions

Due to the many actors in the market environment, there are many turning points hospitals go through to communicate and share information. Studies by Asamoah, Abor, and Opare (2011) and Watsierah and Ouma (2014) take us one step closer to understanding the actors involved in the market environment. These include suppliers, manufacturers, wholesalers, and retailers who need to proactively coordinate with each other.

In a related study undertaken in the health care industry in South Korea, Shou (2013) supports efficient coordination of the market environment if a steady supply of medicines were to be guaranteed. This study also indicated that there was a positive relationship between supply chain performance and customer satisfaction. Nonetheless, in high total buyer supplier relationships, it can be challenging when supply chain partners engage in opportunistic behavior and negative tactics or coercion (Eyaa, 2017). This practice is likely to disenfranchise the availability of drugs at health facilities. From the extant literature, the following dimensions were identified and published in the qualitative paper published by Nagitta and Mkansi (2019).

Table 1: Summary of market management environment dimensions

<table>
<thead>
<tr>
<th>Dimensions</th>
<th>Characteristics/ brief description</th>
<th>Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collaborative partnerships</td>
<td>Formation of strategic alliances with suppliers, agencies, donors.</td>
<td>Seiter (2010); Asamoah, et al. (2011); Watsierah and Ouma (2014).</td>
</tr>
<tr>
<td>Information sharing with suppliers, Ministry of Health (MoH), Donors</td>
<td>Sharing supply and demand information between the general hospitals, MoH and donors, etc.</td>
<td>Ongolo-Zogo and Bonono, 2010; Williams, Roh, Tokar, and Swink (2013).</td>
</tr>
<tr>
<td>Relationship among lower health units</td>
<td>Hospital’s relationship with lower health units through redistribution.</td>
<td>Ministry of Health (2016)</td>
</tr>
</tbody>
</table>

Source: Nagitta and Mkansi (2019).

3.2.2 Macro-environment dimensions

Successful macro dimensions involve understanding the political, economic, social, technological and legal environment. There are many variations in analysing the macro-environmental factors in the realm of management (Sammut-Bonnici and Galea, 2015). However, most studies on macro environment analysis are limited to determining and categorising the factors (Sammut-Bonnici, 2015). Regrettably, the available literature revealed that scant efforts have been invested into investigating the excesses of macro-environmental dynamics.
in the effective management of drug supply in the country. Partly, this prompted the consideration of an empirical or rational path in the proposed research. Table 2 provides a summary of the macro dimensions that can be used by hospitals to tackle ACTs problem.

**Table 2: Summary of macro dimensions affecting ACTs**

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Characteristics</th>
<th>Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Political dimensions</strong></td>
<td>Power, politics, and interest relationships</td>
<td>Jan de Vries and Huissman, (2011); Park, et al. (2017).</td>
</tr>
<tr>
<td><strong>Economic dimensions</strong></td>
<td>Donors’ support and government subsidies</td>
<td>Tumwine, et al. (2010); Morris, et al. (2014).</td>
</tr>
<tr>
<td><strong>Social-cultural dimensions</strong></td>
<td>Social structures and social network, beliefs, and attitude.</td>
<td>Granovetter (2005); Cohen, et al. (2015).</td>
</tr>
<tr>
<td><strong>Technological dimensions</strong></td>
<td>Use of information technology tools such as phones; Rapid Diagnostic Tests (RDTs); m-Health; Use of Internet</td>
<td>Bhakoo and Chan (2011); Hossain (2016).</td>
</tr>
</tbody>
</table>

Nagitta and Mkansi (2019).

The findings were then schematised in a new framework as shown in figure 1, with the hope that it would address the interplay of interdependence among the two management links (market and macro environments). Figure 1 illustrates the stakeholders involved in the supply and distribution of ACTs and the attendant coordination dimensions, as used in this study and derived from the literature.

![Figure 1: The market and macro dimensions of ACTs’ market coordination](www.ejbrm.com)
Source: Nagitta and Mkansi (2015)
The conceptual framework led to the following two hypotheses:

H1: The market dimensions used by general hospitals correlate positively with ACTs’ availability.
H2: The macro dimensions used by general hospitals account for more variance in ACTs’ availability.

3.3 Stage 3: Developing a case study protocol

Case studies require one to plan for the research (Brereton, et al., 2007). Among these was the need to plan a series of case studies. For consistency, we decided to develop a case study protocol template with the aim of providing a common structure for our case study protocols and guidance on how to construct them. We reviewed some of the well-known case study papers and text books but did not find any existing template on either hospital supply chain nor the medicine supply and distribution, despite the fact that Yin (2013) provided an example of a case study protocol. For this reason we constructed our own template based on the basic case study methodologies described by Yin (2003). At a high level, most of the approaches to case study research are quite similar, in spite of the very different philosophical approaches taken by different case study experts. However, Stake (1995), like Eisenhardt and Graebner (2007), takes an interpretive approach to case study research, but is particularly concerned with “programme evaluation” which involves the evaluation of social or education policies. Overall, we adopted the approach described by Yin (2003), who takes a slightly more positivist approach to case studies. We started by defining the research questions and a priori. The data helped answer the preliminary research question, helped to develop the model and two hypotheses surrounding the phenomenon, and created items for the subsequent survey.

3.4 Stage 4: Conducting multiple case study research

Under the qualitative phase, multiple cases were used with the aim of exploring phenomena within its contexts using a variety of data sources. In order to first understand the meaning of SCC as it is experienced by Drug Therapeutic Management Committee DTMCs (in a new context), we conducted focus group discussions with 32 DTMCs from four general hospitals in endemic regions (northern and eastern regions) (Uganda Bureau of Statistics [UBOS], 2016). All the participants who took part in the study were selected purposively and were asked to describe their experiences in coordinating the supply and distribution of ACTs. This was appropriate, since we wanted to answer the ‘how’ and ‘why’ questions (Yin, 2003) and to explore the relevant factors applicable in a particular context describing the world of experience (Myers, 1997). However, the qualitative research process is largely subjective, even for the most seasoned of researchers (Myers, 2000). Using the pattern-matching technique, data was displayed in matrices to facilitate the comparison of patterns predicted in the hypotheses (Yin, 2013). Thereafter, verification was made before identifying common trends of agreement or disagreement, and drawing logical connections to inform conclusions.

Previous scholars, such as Giménez (2004) studied the relationship between internal and external coordination and their relationships with logistics performance. She first explored the integration practices in the context of a grocery supply chain using a multiple case study method. The results of this qualitative research helped in the design of a questionnaire. The instrument was then used to collect data in a survey. Subsequently, structural equation modelling was applied to answer the research question on the effect of integration on performance. She attributed her good results to using these two different, but complementary, methods.

Lee, Song and Cheong (2018) used a comparative study following a balanced approach, starting with grounded theory to gain an understanding of the coordination mechanisms involved in payback contracts. They then described a real-world case study in which coordination affects the supply chain from both the suppliers and retailers’ perspective. Finally, they validated the vendor managed inventory (VMI) system with a base-stock policy as a means to coordinate the supply chain. For their part, Balcázar Camacho, López Bello and Adarme Jaimes (2016) proposed methodological guidelines for the coordinated planning of the supply chain in the context of the health care services, given the global benefits to its members by using primary information collected at several health service institutions. The study demonstrated a quantitative model to show how organisational coordination contributes to a reduction in the overall costs throughout the chain. However, the SCC phenomena described in the cases above did not fully document the medicine supply chain problem in general hospitals. Furthermore, the literature also did not identify contextual relevant variables to help understand the SCC problem in the market and macro environments.

Therefore, to further understand the SCCDs, four focus group discussions (FGDs) were held at four public general hospitals from four malaria endemic regions of Uganda to probe for detailed explanations up to the point of

www.ejbrm.com 60 ©ACPIL
theoretical saturation (Guest, Bunce and Johnson, 2006). Semi-structured questions were designed in such a way as to help the participants think about how the market and macro dimensions affect ACT availability. Permission to tape-record the sessions was also sought and granted. Participation was voluntary, and the participants understood that they had the right to decline to answer any questions and terminate the discussion at any time. Participants were informed that the purpose of the research was to explore how SCCD affects the availability of ACTs. Participants had assurance that pseudonyms would replace their actual names. The FGDs lasted from 90 minutes to two hours as supported by Cooper and Schindler (2014).

Data analysis from multiple cases: Qualitative data analysis took a three-phase approach, which included data reduction, data display, conclusions, and verification (Miles, Huberman and Saldaña, 2014). A computer-aided qualitative data analysis software program known as Nvivo was used to analyse the qualitative data. While there are many computer-aided software programs to analyse in-depth inquiry for case studies (Quinlan, 2011), we found Nvivo more user-friendly in analysing the emerging dimensions. The results from one hospital were replicated to the other general hospitals in Uganda.

The purpose of this inquiry was to develop an instrument that was subsequently validated. The strategy was inductive in nature, and specifically answered the first research question. The study explored how the management environment dimensions affect ACTs in general hospitals in Uganda. In addition, the approach helped to detect similarities and differences in SCC in the various general hospitals. Previous scholars, such as Singh (2013) posed the question, “What are the factors affecting a coordinated supply chain?” He then prioritised the different factors of a coordinated supply chain using the Analytical Hierarchy Process. However, the scholar focused on the micro level only. In this paper, we first prioritised the collection and analysis of qualitative data to assess the supply chain coordination of ACTs in order to gain a deeper understanding of drug availability or stock-outs. The results from this phase were used to develop a survey tool which was used in the quantitative analysis. The design permitted triangulation and clarification of the qualitative results with the qualitative data (Greene, Caracelli and Graham, 1989). The qualitative findings were published in a journal article by Nagitta and Mkansi (2019). Some of the results are displayed in tables 3 and 4.

Table 3: Emerging market dimensions

| SC interdependence with suppliers | Regular communication with supplier (NMS); regular meetings with the supplier; use of e-mails with/to the supplier; use of personal phone calls with NMS; routine regional monitoring with stakeholders; evaluation meetings with supplier; use of M-Track with other external stakeholders; use of the Rx tool; sharing of schedules with the supplier; regular correspondence with the supplier’s representative. |
| Collaborative partnerships for training | Joint training with NMS; collaborative training with Drug Monitoring Unit; refresher training with Ministry of Health and collaborative training with NGOs. |
| Information sharing with suppliers, MoH, Donors | Online sharing of information; regular exchanges information, using hard copies of reports; use of IT tools (M-Track); sharing of weekly or monthly reports; holding of quarterly meetings with external stakeholders. |
| Relationship between lower health units | Hospital’s relationship with lower health units; support for redistribution enhances ACT availability. |

Table 4: Emerging macro dimensions

| Political dimensions | Politicisation of ACTs; political publicity; politicians’ interference; political support whenever need arises; awareness by politicians; verification of ACTs; advocacy by politicians, surveillance or monitoring. |
| Economic dimensions | Poverty within the communities; cost sharing; availability of donor funds. |
| Social-cultural dimensions | Cliques within the community; culture of self-medication; belief of keeping ACTs by households; public attitude towards the lower health facilities; compliance to dosage. |
| Technological dimensions | Use of personal phones; use of toll free lines; M-Track system; use of RDT and use of Internet. |
| Legal dimensions | Testing and dispensing policy; clinical guidelines; regulating consumption; pull policy change. |
Source: Nagitta and Mkansi (2019).
We derived the findings by examining each case study independently and the key aspects of the SCC in the respective contexts, and then examining data and findings across the case studies. The assertions emerging from individual case studies were examined to ascertain their wider relevance. Secondly, the FGD data was examined for variables of interest or emerging key ideas, with relevant statements from interviews being grouped together using Nvivo. This was found to be a useful step in the analysis, as it synthesised evidence from the data around a concept of interest. We analysed the synthesised data by examining each of the research questions across the cases. Within each of the research questions, the variables of interest were examined individually, alternating between detailed analyses. Results from cross-case analysis were used to build a theory and an instrument with two primary hypotheses that were subsequently tested using a survey and Confirmatory Factor Analysis (CFA). These results were published in the *International Journal of SCM*. A conceptual model was eventually developed with the attendant hypotheses.

### 3.5 Stage 5: Pilot study and hypothesised model

The developed instrument had 135 items to represent the dimensions of supply chain coordination. Unlike Singh (2011) and Arshinder, et al. (2011), who did not pre-test their tools, this study pretested the instrument using 10 DTMC members from another hospital which was not part of the study. We revised the questionnaire on the basis of their recommendations. The Cronbach’s coefficient alpha was used to test the internal consistency because it is the most common and widely used method (Mehrab, Siyadat and Allameh, 2013). According to Hair, et al. (2010), coefficients equal to or greater than 0.70 indicate a high reliability of the measuring instrument. The data yielded a Cronbach’s alpha of: Market (.815) and Macro (.855). The success of the theoretical model was attributed to the results of the qualitative study, which helped strengthen the hypotheses and survey items.

### 3.6 Stage 6: Conduct survey and sampling

The unit of analysis was the DTMC members from each hospital. A total of 320 questionnaires were issued to the DTMC members in 40 general hospitals. However, only 304 questionnaires were returned. Of these, 21 were discarded because they had incomplete data (more than 10% missing data). The remaining 283 questionnaires represented a response rate of 88.4%, which is regarded as high (Kothari and Garg, 2014).

The results indicate that 16.6% of the respondents were based in the central hospitals of Uganda, 26% in the North, 29.7% in the East, and 27.6% in the West of Uganda. This implies that all the regions were represented. The majority of the practitioners held a Diploma. Therefore, the information was collected from reasonably well-educated people who knew the medical systems in the country. At least 80% of the participants had experience of more than one year in the pharmaceutical sector. Therefore, the information was collected from well-informed people who knew the medical sector in Uganda.

We subjected the data from Phase Two to confirmatory factor analysis to establish content validity of the priori dimensions. Secondly, we used correlation analysis to test the four hypotheses concerning the relationships of the factors to ACTs’ availability. Support for these hypotheses provided evidence of validity for the instrument, in the sense that the dimensions identified showed mostly the same relationship to the dependent variables of interest. Construct validity exists, according to (Salisbury, 1989), when different indices (in this case dimensions) show the same relationship to other measures as one would expect, based on the theory in which they appear.

### 3.7 Stage 7: Factor analysis using confirmatory factor analysis (CFA) and construct validity

CFA was used to confirm or reject the measurement theory. The conceptual model was tested using AMOS V 21.0 to get a better understanding of the relationship among the latent variables. The hypothesised model was analysed using AMOS 21.0 with the maximum likelihood estimation of the covariance of the items, which enabled us to test the empirical validity of the hypothesised model. We utilised several criteria to determine the inclusion of the items and model fit. The appropriateness of the initial model was assessed based on the recommended goodness-of-fit (GOF) indices. Where the overall fitness of the model did not meet the required standards, it was refined based on the logical modification indices provided by the software.

The following steps were observed: first, items had to have a primary factor loading of .50. Second, items had to be unidimensional as demonstrated by the tests of internal consistency and parallelism (Hair, et al., 2010). Internal consistency requires that the items comprising a scale have a similar statistical relationship to the
primary factor. Parallelism requires that the items of a scale have a similar statistical relationship to the other factors. Since AMOS does not directly test for internal consistency or parallelism, we removed items from the model that the modification option of AMOS suggested had a path to another factor. Virtually, this procedure assured that an item only loaded on one factor. For the final model, internal consistency and parallelism were tested using the product rules of internal consistency and parallelism (Hair, et al., 2010). Third, the items had to have homogeneous content. Fourth, the items needed to show an acceptable level of reliability (Cronbach’s alpha). Bryman and Bell (2011) described the Cronbach-alpha test as a commonly used test to determine internal reliability.

3.8 Stage 8: Test model and hypotheses

3.2.3 Market environment

CFA was conducted to confirm that factors that were extracted converged as manifest variables of the global latent variable. The results of the CFA indicated that the two manifest variables significantly loaded on the organisation factors (OF), Information Sharing (IS) and Market management environment dimensions. The GOF indices of market dimensions after modification were: GFI=.930, IFI=.965, TLI=.957 CFI=.965, RMSEA = 057. The lower and upper boundaries at 90% confidence interval were .046 and .068, respectively. All the GOF parameters were above the recommended values. The results of the CFA indicated that all the variables significantly loaded on the global latent variable of Suppliers interdependence (SCI), Partnership training (CP), and Information sharing with donors and MoH (ISS). The findings confirmed the validity of the final model with excellent model fit statistics between the model and the observed data.

From the regression weights for the default model it can be stated that H₁ is supported. The availability of malaria treatment therapies strongly correlates with SCI (.59), followed by ISS (.28), while CP had no significant correlation (-.03).

Asamoah, Abor, and Opare (2011) proposed the need for a well-built coordination framework to tackle ACTs’ complex partnerships and subnational programmes, hence the rationale for sharing critical information between the general hospitals and external stakeholders. Organisations should be in a position to gather data, organise it, analyse and report accurate, timely and appropriate information to decision-makers to evaluate how supplies flow, account for products, reduce supply imbalances and improve efficiency. Together with information technology and business practices, coordination becomes easier (Ebrahim-Khanjari, Hopp and Iravani, 2012). However, in Uganda, the existing information system is predominantly paper-based.

In support of partnerships, Kembro, Näslund, and Olhager (2017) assert that supply chain partnerships lead to increased information flows, reduced uncertainty, and a more gainful supply chain. Unfortunately, the accuracy of the supply and demand data is most often challenged, something that has negatively impacted the quality of services offered, rate of responsiveness, and operational efficiency.

3.2.4 Macro-management environment dimensions

Upon modifying the model base on the modification indices, the allowable GOF values were achieved: IFl=.976, TLI=.972, CFI=.976, and RMSEA=.039. Accordingly, failure to reject the null is a sign of a good model fit that is a reverse testing procedure in CFA (Blunch, 2017). To test the second hypothesis, standard regression weights were generated. The results indicate that the availability of malaria treatment therapies is mainly affected by legal dimensions (LF, .58) followed by socio-cultural (SC, .39), while political dimensions (PF) have no significant impact on the availability of ACTs.

While political dimension had no significance, politically, medicine attracts a high level of interest, owing to its high economic value, the broad public and private investment, and its impact on the health and well-being of society. As Williams, et al. (2013) assert, all organisations, irrespective of formality, size, or makeup, are made up of structured political aspects that must be understood and managed, and the supply chain is not an exception. Parenthetically, the health budget in most sub-Saharan African countries is still below 15% of the total budget for health, as required under the Abuja Declaration.

However, in Uganda, medicines are the single most expensive out-of-pocket health expenditure item, and therefore, constitute an enormous burden for the ordinary person. With limited financial support from the treasury and also from donors, there will always be a shortage of ACTs. This notion was reiterated by Morris, et
al. (2015), stressing that while ACTs are subsidised, and therefore, the recommended first-line treatment for uncomplicated malaria in most endemic countries, they are prohibitively costly in the retail drug. In determining the most critical SCCD that emerged from the qualitative phase, the hypotheses for making ACTs available in general hospitals in Uganda were adopted for this study.

From the results it can be concluded that the most influential supply chain market environment dimensions that affect the availability of ACTs are Information Sharing with external stakeholders and supply chain interdependence, among others. The macro-environment shows that the socio-cultural, economic, technological and legal dimensions influence the availability of ACTs more. The political dimension scored least and, therefore, was dropped. From the study, the following model was derived.

Market and macro SCCD framework for ACTs in general hospitals in Uganda

![Diagram of Market and Macro SCCD framework for ACTs in general hospitals in Uganda]

**Figure 2:** A SCC framework for malaria treatment therapies in general hospitals in Uganda

### 4. Discussion

Following the logic of a mixed methods approach to deconstruct a multi-embedded supply chain coordination problem appears to be the right strategy. In particular, the strength of both methods in refining theory and advancing the knowledge related to supply chain coordination in practice. For example, using both the qualitative and quantitative paths was pertinent for first understanding the coordination dimensions as used by the hospitals, and then explaining this phenomenon in the medicine supply chain context. The study adopted an inductive strategy and employed a multiple case study approach. The success of the theoretical model was attributed to the results of the qualitative study, which helped strengthen the hypotheses and survey items. By following different research approaches, it was critical to answer the various research questions accurately.

This study explored key dimensions from expert focus groups to assist in the development of an instrument that could be used for future studies. The study of the supply chain coordination of malaria treatment pills (ACTs) in selected general hospitals was reflective of the wider public general hospitals in Uganda in terms of content, samples and patterns as supported by Cohen, Lawrence and Morrison (2018). This is especially so, since the research approach included hypotheses testing (Lichtman, 2012) and can, therefore, be repeated at a later date or place and still get the same results (Shank and Brown, 2007). Notwithstanding, there are still some advantages and criticisms related to the efficacy of the mixed method approach. The problems and benefits associated with mixed methods approach are highlighted below.
4.1 Advantages of using mixed methods in the study

The technical advantages of using mixed methods are plentiful. Each approach has clear objectives that differ from each other and complement rather than compete: “the limitations of one method can be offset by the strengths of the other method, and the combination of quantitative and qualitative data provide a more complete understanding of the research problem than either approach by itself” (Creswell and Plano, 2011, p.45). The essential purpose of mixing methods in this paper was to obtain the best measurement of multidimensional coordination problem for the study population using the strengths of each of the methods proposed.

In the research presented here, the use of mixed methods was important for three main reasons: (i) allowing for a robust analysis following the triangulation of data, sequencing of data collection and research methods and methodological bilingualism, (ii) identification of priority areas for policy, and (iii) promoting policy uptake at the general hospital level.

From the study, mixed methods were particularly effective in gathering information that would not have been collected in isolation by either the quantitatively or qualitative approach. The integration of the methodologies was essential to develop a complete diagnosis of the ACTs’ availability in public general hospitals in Uganda, where supply chain information is either fragmented or non-existent (Ntayi, et al., 2010).

Furthermore, the use of mixed method helped us better account for some of the criticisms of and weaknesses inherent to the qualitative and quantitative methods. For example, the issue of replicability associated with a qualitative research approach. Critics of this approach argue that the constructivist abandoned the scientific methods and procedures of enquiry and investigation (Cohen, 2012). It is alluded that users of the qualitative approach are said to write fiction because they have no means of verifying their true statements. Since the approach is characterised by feelings and personal reports, it is believed that the approach cannot give reliable and consistent data when compared to using quantifiable figures (Atkins and Wallace, 2012). For the purpose of the present study, the reliability of the qualitative data was ensured by considering the trustworthiness of procedures based on the following criteria: credibility, authenticity, dependability, and confirmation (Onwuegbuzie and Johnson, 2006).

For quantitative data, generally three methods were applied to ensure reliability: a) split-half technique, b) item analysis and c) Cronbach’s alpha (Cooper and Schindler, 2011). Although the qualitative method is regarded as subjective, misleading, inaccurate and subjective because of the ontological and epistemological paradigms (Cohen, Lawrence and Morrison, 2011), the study triangulated the qualitative findings with the quantitative findings making them credible. Denzin and Lincoln (2005) contend that the constructivists’ approach is purely exploratory and makes it difficult and impossible to simplify findings and observations. Qualitative researchers believe that the social world (phenomena and experiences) has many dimensions, hence explanations are based on the interpretations of the researcher (Leedy and Ormrod, 2015). In view of this, the study counteracted this argument by carrying out a survey to gain a proper explanation of the results without depending only on the explanations of the researcher. The implication is that the research can be repeated by another researcher at another place and still get the same results.

Finally, Denscombe (2003) describes quantitative research as one where the researcher detaches himself or herself. By using the quantitative approach in the second phase of the study, it was seen as a strength of quantitative research approach from one angle, yet from another angle it was seen as its weakness. The issue of the researcher being biased with either the data collection or data analysis was eliminated because we were not in direct contact with the participants, since data was collected using a questionnaire. There was full control for alternatives such as interpretations, explanations, and conclusions. Therefore, the objectivity of the researcher was not compromised. Secondly, this perhaps guaranteed respondents’ anonymity.

4.2 Disadvantages of using mixed methods in the study

Despite the usefulness of a mixed method research approach for conducting research in the medicine supply chain, there are still teething problems. By using a qualitative approach, we viewed the social world as being dynamic and not static (Johnson and Christenen, 2014). In view of this, we limited the findings to the particular group of people being studied (DTMC members in public general hospitals). However, this meant that certain behaviour and characteristics observable in a dynamic multi-layered supply chain were not considered. Put simply, even with its best intention, there are limits to the nature of the phenomenological constructs that mixed
methods can uncover in a dynamic social context. Notably, the process of combining the two approaches in a single study is very laborious.

This study was also limited by the choice of research design. It was a cross-sectional study ignoring the importance of a longitudinal design, which can be useful in investigating the characteristics of supply chain coordination mechanisms over time, based on the study variables. Furthermore, since this study was cross-sectional, it could neither discuss nor come up with conclusions on the causality of critical supply chain coordination dimensions, management environment dimensions, and availability of ACTs. Therefore, it was difficult to claim that critical supply chain coordination dimensions, and management environment dimensions can cause changes in the level of ACTs over a long period of time. Although the cross-section limitation can be attributed more to the design than the methods, at best, it reveals some of the boundaries of what can and cannot be achieved through mixed methods, to pave the way for future studies as discussed in the next concluding section.

5. Conclusion

The study aimed to demonstrate the validity of the use of mixed methods in studying the medicine supply chain and its additional value for informing policy processes. Through the integration of information, methodologies, and experience in quantitative and qualitative approaches, we designed a mixed methods diagnosis of ACTs’ availability in public general hospitals in Uganda. Through this methodological approach (an exploratory sequential mixed method), the current study contributes to our understanding of SCCD, explaining clearly ACTs’ availability in general hospitals in Uganda, and creates an instrument for its measurement, which was subsequently validated using the quantitative procedures.

The qualitative approach (focus group discussions) transcripts enabled an understanding of SCCD in the availability of ACTs and led to the development of an instrument unlike that of previous studies conducted by Singh (2011) and Arshinder, et al. (2011). The quantitative approach allowed the discovery of reality through the statistical analysis of relationships between the independent and dependent variables. Therefore, the research design provided appropriate procedures for an inquiry into the existing phenomenon under study. The use of a mixture of techniques enhanced and refined the results. The reliability of items and construct/content validity of the measurement items were purified based on CFA, thereby confirming their usefulness in measuring the different variables under study. Therefore, this study provides evidence on the use of these measurement items for future scholarly studies. Finally, this study provides direction for further research, based on the empirical and theoretical concepts used.

5.1 Synopsis

In this study, which deconstructed a supply chain coordination problem, the mixed methods approach used an exploratory sequential research approach and methods that represented different research strategies, and which differed in terms of their theoretical, epistemological, and ontological issue. The use of the approach depended on the preference of the method of data collection and analysis. Researching and deconstructing the coordination problem required a focus on the way data was collected and interpreted.

5.2 Theoretical implications

Using a mixed methods strategy, this study made significant findings on the usefulness of the qualitative and quantitative research approaches in deconstructing the supply chain coordination issues in the public general hospitals in Uganda. Therefore, the application of mixed methods research should be enhanced because it serves the advancement of the supply chain management discipline, especially where little or scant information is available. The approach provides a richer understanding and more robust explanation of the phenomena, especially in instances of lack of literature and instruments. The implication points towards fully embracing mixed methods if the supply chain management phenomenon is to be fully understood, other than using a single approach which may not be sufficient for this task.

5.3 Practical implications

The results highlight key policy implications for hospitals. For instance, hospitals and governments should rethink investing in innovative information technology tools that will enhance real-time information exchange with key market and macro chain players. With regards to the macro environment, hospitals must manage the
interdependencies by working closely with the local communities through the creation of awareness, and embrace cheaper technologies while working within the legal frameworks and standard operating guidelines.

5.4 Limitations

Our study is limited by a number of factors and we discuss the factors in this section and also recommend directions for studies in future. The study is limited to a multi-embedded health care supply chain for ACTs from a general hospital perspective. Specifically, the sample was composed of public general hospitals, leaving out regional, national referral hospitals, faith-based and other lower health centres in Uganda. Our findings cannot, therefore, be generalised across the spectrum of other hospitals. In the future, the instrument can be extended to other multi hospital levels in the country, because the results may vary.

The respondents were Drug Therapeutic Committee Members in charge of forecasting, quantification, procurement, storage and dispensing, and yet there are other people within the hospital, and external to the hospitals, such as suppliers, regulators, and donors, whose decisions affect the supply and distribution chain that were excluded from the study. This study, therefore, missed out on the vital information that these other groups of people would have provided. In future, multi-level studies should take into account all categories of supply chain partners to offer a comparable outcome, including investigating other medicines. In addition, in order to gain a true nature of supply chain coordination, a longitudinal study may be conducted in the future, since practices are gained over time.

References


Research Methods. Ch.5


www.ejbrm.com

©ACPIL
Reimagining Academic Writing in Academia 4.0 to De-incentivise Plagiarism\(^1\)

Anthony Stacey  
Graduate School of Business Administration, University of the Witwatersrand, Johannesburg, South Africa  
Anthony.Stacey@wits.ac.za  
DOI: 10.34190/JBRM.18.1.003

Abstract: Academic research and scientific publication are being influenced irreversibly by what is referred to as the fourth industrial revolution. The exponential growth in the number of research publications continues, information and communication technology (including artificial intelligence) is making available research data and tools with unprecedented capabilities, and online open access to publications has enabled greater and more rapid access by other researchers. Changes of research practice and the behaviour of researchers and authors as a result of these developments are evident, and are challenging the criteria, norms and standards by which the quality and integrity of research has historically been judged. The manner in which prior research is being accessed, reproduced, applied and acknowledged is an example of such changes. In academia, the presentation of the ideas or writings of another without being explicitly attributed to the original source has always been regarded as plagiarism and considered serious misconduct. Yet when such ideas and writings are freely available and in the public domain, they arguably fulfil the criteria for being considered common knowledge which don’t necessarily need to be referenced. This article presents examples of acceptable replication and reuse of the work of others, and examples of how plagiarism is manifesting differently because of information and communication technologies, including plagiarism software. It is argued that while paraphrasing previous authors result from understanding and applying their prior research, paraphrasing may simply be a grammatical or mechanistic process that does not attest understanding and application. It is provocatively suggested that current norms and standards of academic writing, including referencing, may no longer be appropriate. Relatively modest amendments to academic conventions and assumptions are proposed that could lead to a new paradigm of more efficient research and scientific publications, acknowledging that this would place greater burden of responsibility on the users, reviewers, editors and examiners of research to be familiar with extant knowledge.

Keywords: Plagiarism, Ghost writing, Common knowledge, Attribution, Normal science, Paradigm shift, Literature review

1. Introduction

The context within which academic research (research project reports, dissertations and theses for higher degrees) and scientific publication (in traditional subscription journals, open access journals and institutional repositories) are taking place is undeniably evolving. The growth of research literature continues despite Price’s anticipated knowledge saturation (Wagner and Kim, 2014), research is becoming increasingly accessible electronically, and software and artificial intelligence systems continue to become more powerful.

These and other disruptive technologies are resulting in changes of behaviour and research practice. Specifically, the manner in which academics and researchers access, manage, analyse, develop, grow, describe, publish, exchange and reference knowledge is evolving. These changes of behaviour and practice are testing the validity and legitimacy of traditional norms, standards and assessment criteria, which may become less and less appropriate over time.

The aim of this article is to alert researchers to the changes that are currently being experienced in academia associated with contemporary information and communication technologies, and to provoke discussion and debate on how the academic and publishing community might respond constructively to the emerging challenges.

The emphasis is on the representation of the words, ideas and thoughts of others as one’s own, because what has traditionally been considered plagiarism is one of the practices that highlights the emerging tension between the traditional and the future status quo. The reasons that students and researchers plagiarise include managing workload, ease of access to sources, inadequate knowledge of subject matter and simple laziness (Kayaoğlu et

\(^1\) An earlier version of this article was presented at the 18th European Conference on Research Methodology for Business and Management Studies, 20 – 21 June 2019, Wits Business School, Johannesburg under the title Reframing Plagiarism in Academic 4.0.
al., 2016). These will not be discussed in this paper because it is not anticipated that these antecedents will change fundamentally in the foreseeable future.

2. A Paradigm Shift towards Academia 4.0

Industry 4.0, also referred to as the fourth industrial revolution, is the term that has been coined to encapsulate current emerging and disruptive technologies in industry, such as cyber-physical systems, the internet of things, cloud computing and artificial intelligence. It is a progression from the first industrial revolution, characterised by mechanisation, water and steam power; the second industrial revolution, characterised by electricity and mass production; and the third industrial revolution characterised by computerisation and automation. In this paper, Academia 4.0 is therefore used as a metaphor for the currently emerging and disruptive trends in tertiary education.

2.1 Exponential growth of research literature

The development of scientometrics as a field of study was pioneered by Derek John de Solla Price. He found evidence that the body of scientific literature has been growing exponentially for the past few centuries (Price, 1963). This has led to the aphorism that a substantial portion (up to 90%) of all the scientists that ever lived are alive today (Gastfriend, 2015). Price (1965) estimated the growth rate of scientific literature to be approximately 7% per annum and, despite his logic that exponential growth is not sustainable, this appears to be continuing even at an increased rate (Wagner and Kim, 2014, Bornmann and Mutz, 2015).

The proliferation of research in any particular field presents challenges to academic authors. First, it may not be easy to identify seminal and cutting-edge works in the field. Second, an author may independently arrive at thoughts, ideas or words that are comparable to those of other authors, thus giving the false impression of having plagiarised.

The relentless pursuit of knowledge in academia is competitive and obligatory, as captured in the maxim “publish or perish”; the metaphor of dwarfs standing on the shoulders of giants is often used to describe academic authors, who themselves aspire to become “giants” one day. Wagner and Kim (2014) identify various authors and philosophers of science who expressed concern about the quality as opposed to the quantity of scientific output. Without engaging with the awkward question of how to define and then measure the quality of research output, it is conceivable that the growth of research literature has been and continues to be at the expense of the quality of publications.

2.2 Electronic access to research

Information and communication technology has been a fundamental enabler of access to published research. Open access – providing free, online, permanent and largely unrestricted access to research – has experienced worldwide growth. Pinfield et al. (2014) in reviewing this growth from 2005 to 2012 note that while the initial development of repositories took place in North America, Western Europe, and Australasia, while more recently there has been repository growth in East Asia, South America, and Eastern Europe. A more recent review (Piwowar et al., 2018) identifies and describes four emerging themes regarding open access: an increasing partiality of funding institutions towards open access to the research; value adding services based on open access sources; contestation regarding the ethics and appropriateness of traditional subscription publishing; and the increasing unaffordability of access to traditional subscription sources.

It is acknowledged and should be noted that electronic access to research is not a panacea for researchers. In addition to the benefits, Tennant et al. (2016) and Björk (2017) identify detriments of open access to research. For example, the costs for researchers are typically higher for open access journals, and lower barriers to entry have resulted in a proliferation of so-called predatory publishers and conferences, with a corresponding impact on the quality and reliability of research publications.

Easy and reliable access to research publications are now a reality, irrespective of the advantages and disadvantages. An implication that has not been comprehensively addressed is that, being freely available in the public domain, open access research publications now fulfil one of the criteria for being considered common knowledge. This has profound implications for the attribution of knowledge and referencing.
2.3 Collaboration and communication among researchers

One of the phenomena of the internet era has been the emergence of literally hundreds of social networking websites, among the most popular today being LinkedIn (founded 2003), Facebook (founded 2004), Reddit (founded 2005), Twitter (founded 2006), Instagram (founded 2010), and Pinterest (founded 2010). These sites are defined as internet-based systems on which users can create individual profiles, connect with other users, and make additional connections through other users that would otherwise not be possible (Boyd and Ellison, 2007). Ovadia (2014) noted the increasing popularity of specialized academic social networking sites, focussing specifically on ResearchGate and Academia.edu.

The potential benefits of these academic social networking sites are similar to those of open access, and include knowledge dissemination and greater research impact, connecting with scholars with similar research agendas, sharing of data, and keeping acquainted with current research trends (Williams and Woodacre, 2016). Scholars have also expressed reservations about the use of academic social networking sites, for example: fears that their work will be plagiarised, the speed and convenience of communication can be at the expense of the quality of engagement; manuscripts can be posted that are not of publishable standard in recognised scientific publications, and traditional publication formats are favoured for purposes of academic appointments and promotions (Williams and Woodacre, 2016). No matter what the opportunities or risks posed by these academic social networking sites, there can be little doubt that they will remain a component of the academic research and publishing ecosystem.

2.4 Information technology and artificial intelligence

The influence of information technology (both hardware and software) and artificial intelligence on research cannot be underestimated. The sizes of datasets no longer constitute a constraint on storage, access or analysis; cloud storage provides practically limitless storage capability, communication bandwidths allow for rapid exchange of vast datasets, algorithms previously too complex to be practicable can now be executed, and the expertise and tools for unstructured “big data” analytics are commercially obtainable.

It is well known that researchers may have their own discernible style and register of academic writing. As and when the technology becomes available to characterise, identify and replicate researchers’ style and register of writing, it may become impossible to attribute sources and contribution to knowledge reliably. This would undermine one of the fundamental tenets of academic integrity.

2.5 Literature Reviews

There is a degree of consensus about what constitutes a good literature review. Ryan (1979) notes that the literature review will both contextualise a study and justify the research question, and must be based on the best available resources. Similarly, Zaporozhetz (1987) referred to the “thoroughness” of the literature review, but added that there ought to be an overall flow of ideas and attention to detail. Indeed, in the latter part of the 20th century there are many mentions of a good literature review implying comprehensiveness and completeness as being the criterion of quality. Possibly this can be attributed to sources being relatively less readily accessible than they are today. Walker (1998) provides a useful editor’s perspective in commenting that authors evidently try to impress by listing an excessive number of references, whereas the literature review should rather articulate the purpose of the study and then review the literature the context of that purpose. Electronic access to prior research has been an enabler of literature reviews that are comprehensive and have a broad scope, but has also resulted in the emergence of literature reviews that are little more than a compilation of summaries of prior publications. Such compilations add little value and do not fulfil the intended functions of literature reviews.

More recent publications have formalised the purpose of and what constitutes a good literature review. The “inverted pyramid” conceptual model (Maier, 2013) in Figure 1 identifies four distinct steps in the writing of a good literature review. Similarly, Rewhorn (2018) concludes that a literature review must provide context, identify and describe gaps in existing scientific knowledge, and providing the rationale for the research questions. Nakano and Muniz Jr (2018) argue that any contribution to knowledge requires a good literature review to provide the theoretical framework. The authors cite key features of a good literature review, being: coverage of relevant sources, synthesis of a new perspective on the topic, well-articulated substantiated assertions (rhetoric) and evidence of practical and theoretical significance.
Figure 1: "Inverted pyramid" model of steps in writing a literature review (Maier, 2013)

Existing software and search algorithms have largely resolved the challenges of adequate coverage of relevant sources, thoroughness, providing context and identification of the domain of the research problem. New artificial intelligence applications are proliferating, and search tools are increasingly effective. Combined with natural language generation and text composition software, it is conceivable that the customary review of literature, the traditional foundation of research endeavour, will be automatically generated on the basis of appropriately specified key words and phrases. In this context, knowledge in the form of published research will in effect constitute research data for the potential generation of new knowledge, with debateable intervention and contribution from human researchers.

The review of literature as a research method in its own right, widely referred to as a Systematic or Structured Literature Review (SLR), was alluded to by Mittelstaedt and Zorn (1984) when indicating that it should examine the generalisability of particular relationships by means of a meta-analysis. Rowe (2014) provides another editorial perspective, with guidelines and recommendations for authors confirming that a literature review paper is not merely an overview of the literature. The greater accessibility to prior research publications enabling researchers to carry out comprehensive literature reviews is explicitly noted by Kraus, Breier and Dasi-Rodríguez (2020) who also provide a useful comparison between systematic and traditional literature reviews.

2.6 Plagiarism and ghost writing

There appears to be a degree of consensus in the literature that evidence and instances of plagiarism and ghost writing are increasing. Appropriately these phenomena are receiving increasing attention from the research community. It is self-evident that information and communication technologies have been enablers of both plagiarism and ghost writing, although the very same technologies may also have a greater role to play in the future in mitigating against them.

Singh and Remenyi (2016) note that academic misconduct is not new, citing cheating in examinations, data falsification and fabrication, and illicit pre-viewing of examination papers as well-established examples. However, the authors note that the opportunities for cheating have multiplied in the past few decades. Similarly, Fisher et al. (2016) illustrate that the modus operandi of ghost writers depends heavily on the anonymity of the internet and the proliferation of online academic programmes.

The critical urgency of responding to plagiarism and ghost writing is highlighted by Fusch et al. (2017). Both involve profound deception which compromises integrity. The authors go on to describe the ethical implications of plagiarism for all participants and stakeholders in the research as a “moral harm inflicted on all parties” (Fusch et al., 2017). It is therefore appropriate to look deeper into the emerging tension between the historic manifestations and understanding of plagiarism, and the future status quo.

3. Emergence of Plagiarism

3.1 Evolution of plagiarism

It is not within the scope of this paper to review the origins and history of plagiarism, and there are various authors who have provided comprehensive and informative synopses (for example Sutherland-Smith, 2010). Common themes of these perspectives is that plagiarism has existed since time immemorial, that it is referred to using criminalising language, and that there is not necessarily universal consensus on the interpretation or appropriate consequences. It is noted that most research is carried out within existing frameworks or paradigms.
and has been referred to as “normal science” (Kuhn, 1962), and therefore a certain amount repetition and replication of others’ ideas is probably unavoidable.

### 3.2 Acceptable replication and reuse

It would be inaccurate to suggest that all replication and reuse of the words or ideas of others without acknowledgment is regarded as plagiarism.

#### 3.2.1 Academic phrasebooks

There are many commonly used phrases in academic writing, to such an extent that some have become the butt of internet jokes and memes, examples of which are given in Table 1. While these particular interpretations are obviously not intended to be taken seriously, there are numerous internet resources that give examples of useful phrases and sentences for the benefit of academic authors, particularly those not writing in their mother tongue. The adoption of such phrases and sentences is so widespread that their use cannot reasonably be construed as plagiarism.

<table>
<thead>
<tr>
<th>Common phrases</th>
<th>Correct interpretations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Typical results are shown ...</td>
<td>This is the prettiest graph.</td>
</tr>
<tr>
<td>In a series of cases ...</td>
<td>Thrice.</td>
</tr>
<tr>
<td>According to statistical analysis...</td>
<td>Rumour has it.</td>
</tr>
<tr>
<td>It is believed that...</td>
<td>I think.</td>
</tr>
<tr>
<td>In my experience ...</td>
<td>Once.</td>
</tr>
<tr>
<td>It is generally believed that ...</td>
<td>A couple of other people think so, too.</td>
</tr>
<tr>
<td>In case after case ...</td>
<td>Twice.</td>
</tr>
<tr>
<td>It has long been known ...</td>
<td>I didn’t look up the original reference.</td>
</tr>
</tbody>
</table>

*Source: Various internet web sites. * †The sequence of these phrases has been randomised so as not to reproduce the original sources exactly.

#### 3.2.2 Imitation and Flattery

Oscar Wilde is credited with the adage “Imitation is the sincerest form of flattery that mediocrity can pay to greatness.” Many witticisms are attributed to this Irish poet and playwright, but this saying has particular relevance in academia. In some cultures, it continues to be the case that the use of the words or ideas of others is deemed appropriate and respectful of seniority and authority (Park, 2003). While this is not prevalent in Western cultures, it does suggest that plagiarism needs to be viewed with cultural sensitivity.

#### 3.2.3 Code reuse in software engineering

Plagiarism is particularly challenging in the software engineering discipline. Gibson (2009) notes that it is quite normal for existing code to be used in the development of new software, and discusses a code of practice specifically for use in academia to differentiate between legitimate re-use and plagiarism. Object oriented programming and example embedding (Barzilay, 2011) are examples of widely used software development techniques that involve legitimate and intentional reuse of other developers’ code, which would not be regarded as plagiarism.

### 4. Contemporary plagiarism

Verbatim plagiarism in academic texts – comprising phrases, sentences, paragraphs or more that have been copied from original sources and replicated in an academic text without appropriate attribution – is surely the most straightforward and easily understood form of plagiarism. It is equally the most easily perpetrated form of plagiarism, particularly with the availability of electronic and online academic content.

#### 4.1 Plagiarism software

There is a plethora of software available to address plagiarism. It is intriguing to consider how these various software products are described: they are labelled variously as “plagiarism detection”, “plagiarism analysis”, “plagiarism checking”, “anti-plagiarism”, or “plagiarism prevention” software. There are two problems with these labels. Firstly, an implied claim that the software itself might prevent plagiarism is farfetched; only the
researchers or authors themselves can avoid plagiarising with the software providing them some technological support. Secondly, current software packages use algorithms that identify similarities and matching text, making it particularly reliable at identifying verbatim plagiarism. However, not all forms of plagiarism are manifest as similar or matching text. The theft of the thoughts of others may have no similar or matching text whatsoever and may therefore remain undetected, yet remains plagiarism.

Turnitin appears to have attracted the most attention of academics and researchers. The website (www.turnitin.com) asserts that “academic integrity begins with Turnitin” and claims to be “the world’s most effective plagiarism detection solution”. Another popular plagiarism software is Grammarly (www.grammarly.com) which also checks grammar, spelling, tone and style. This software can be used as an add-in to Microsoft Office to check documents and emails. iThenticate claims on its website (www.ithenticate.com) to be “the most trusted plagiarism checker by the world's top researchers, publishers, and scholars.” While that claim may be difficult to substantiate, the website does have links to a few useful readings.

The problems of both false-positives and false-negatives do not negate the benefit and utility of these software. However, they do require more than a superficial and simplistic interpretation of a similarity metric. The software has potentially given academia a false sense of security, because it has changed how researchers and authors compose their works. Writing and editing to evade plagiarism software has become a specialised and valued skill.

4.2 Paraphrasing

When reviewing literature and prior research, it is common practice to paraphrase the text of the original authors. The principle behind this is that it demonstrates one’s understanding and interpretation of the original text in the context of one’s own research. The logic is that through understanding and interpretation of the original text in the unique context of one’s research, one will end up paraphrasing the original text. However, the reciprocal logic does not necessarily apply: paraphrasing the original text does NOT guarantee or imply that one has necessarily understood the original text and applied it to one’s own research.

Paraphrasing can be a mere grammatical exercise. There are websites and applications that will perform the function, albeit with varying degrees of competence, without any intellectual input from the researcher. By way of example, the following sentence was obtained from the Oxford University website providing academic guidance to students (https://www.ox.ac.uk/students/academic/guidance/skills/plagiarism?wssl=1) and some clauses also are to be found in various online resources:

> Paraphrasing the work of others by altering a few words and changing their order, or by closely following the structure of their argument, is plagiarism if you do not give due acknowledgement to the author whose work you are using.

This sentence was submitted to a number of paraphrasing applications, the results of which are shown in Table 2.
Table 2: Examples of paraphrasing applications

<table>
<thead>
<tr>
<th>Tool and Website URL</th>
<th>Paraphrased text</th>
</tr>
</thead>
<tbody>
<tr>
<td>QuillBot <a href="https://quillbot.com/">https://quillbot.com/</a></td>
<td>Paraphrasing other people’s job by changing a few words and changing their order, or closely following the structure of their argument, is plagiarism if you don’t offer the author whose work you’re using due recognition.</td>
</tr>
<tr>
<td>Spinbot <a href="https://spinbot.com/">https://spinbot.com/</a></td>
<td>Summarizing crafted by others by adjusting a couple of words and changing their request, or by firmly following the structure of their contention, is written falsification in the event that you don’t give due affirmation to the writer whose work you are utilizing.</td>
</tr>
</tbody>
</table>
| Small SEO Tools
  https://smallsesotools.com/article-rewriter/ | Paraphrasing the work of others by sterilisation some words and dynamic their order, or by closely following the structure of their argument, is plagiarism if you are doing not provide due acknowledgement to the author whose work you’re victimization. |
| Best Free Spinner
  http://bestfreespinner.com/         | Paraphrasing the work of some others by altering a very few words and changing their particular order, or by carefully following the structure of these argument, is plagiarism should you not give due acknowledgement towards the author whose work you might be using. |

It is clear that the ideas expressed in the original text have remained unchanged, although in most instances the style has become somewhat idiosyncratic. Nevertheless, paraphrasing using mechanistic tools such as these can be an effective means of evading plagiarism software, particularly as natural language processing and paraphrasing algorithms mature and increase in effectiveness. Therefore, it cannot be assumed that paraphrased text is free of plagiarism, even if it has evaded detection by so-called plagiarism software.

4.3 Anecdotal experience

Ison (2015) did not find that that internet had had a significant impact on the prevalence of plagiarism at doctoral level, and this author’s experience of postgraduate student research suggests that the reliability of software that can identify verbatim or cut-and-paste plagiarism has proved to be a deterrent. However, there have been perhaps unintended consequences: while there has been a decline in verbatim plagiarism, the relative number of instances of more subtle forms of plagiarism has undoubtedly increased.

4.3.1 Randomisation

It was indicated in a note to Table 1 that the entries in the table had been resequenced. Such reordering of itemised lists, or transposing of words or phrases is a common method used by academic authors to evade detection of plagiarism by the applicable software.

4.3.2 Word switching

The use of thesauruses together with the “find and replace” word processing function makes it straightforward to edit source text so that the words no longer match the original. In the comparison between the student’s submission and the original source given in Table 3 it is clear that the words “dominant players” or “the dominant” have been substituted for “monopolies” and “monopoly” respectively.

Table 3: Comparison of student’s submission and original source illustrating word switching

<table>
<thead>
<tr>
<th>Student’s submission</th>
<th>Original source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dominant players have existed throughout much of human history. This is because powerful forces exist both for the creation and maintenance of dominant players.</td>
<td>Monopolies have existed throughout much of human history. This is because powerful forces exist both for the creation and maintenance of monopolies.</td>
</tr>
<tr>
<td>The Dominant is a term used by economists to refer to the situation in which there is a single seller of a product (i.e., a good or service) for which there are no close substitutes. The word is derived from the Greek words monos (meaning one) and polein (meaning to sell). (Swannell, 2006)</td>
<td>Monopoly is a term used by economists to refer to the situation in which there is a single seller of a product (i.e., a good or service) for which there are no close substitutes. The word is derived from the Greek words monos (meaning one) and polein (meaning to sell).</td>
</tr>
</tbody>
</table>
Plagiarism software would generally draw attention to instances of plagiarism of this nature. The absurdity of this example is patently obvious in the second paragraph in which the etymology of the word “monopoly” is incongruously used to explain the derivation of the word “dominant”.

4.3.3 Paraphrasing as a means to avoid plagiarism software

In the following example the student’s submission returned a relatively low similarity index of 5%. However, closer inspection revealed that extensive sections of the Literature Review consisted of systematic paraphrasing of consecutive sentences from other authors’ publications. Evidence led at this student’s disciplinary hearing indicated that approximately 63% of Literature Review had been edited in this fashion.

By way of an example, Table 4 gives three sentences from just one paragraph of the student’s submission and the original source text. Sentences have been shown separately for greater clarity.

Table 4: Comparison of student’s submission and original source illustrating ideological plagiarism

<table>
<thead>
<tr>
<th>Student’s submission</th>
<th>Original source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Higher financial literacy levels have been associated with superior financial decision making.</td>
<td>Higher financial literacy scores have been linked to higher quality financial decisions.</td>
</tr>
<tr>
<td>The chief costly financial blunder that older family units make is not refinancing their mortgages in an environment where interest rates are falling.</td>
<td>A particularly costly financial mistake for older households in a falling interest rate environment is the failure to refinance a mortgage</td>
</tr>
<tr>
<td>Mimbmb-Johnson and Lewis (2009), established that characteristics related to financial sophistication are good predictors of refinancing behaviour that maximises wealth.</td>
<td>Campbell (2006) finds that characteristics associated with financial sophistication predict wealth-maximizing refinancing behavior.</td>
</tr>
</tbody>
</table>


While the original words have not always been used, it is clear that the student has faithfully reproduced the thoughts and ideas of the original authors without any acknowledgment. This is termed ideological- or style plagiarism, and may be considered to be an indication of a deliberate attempt to “outsmart” plagiarism software. It is interesting to note that the student replaced the original citation in the last sentence, which may be a further indication of intent.

4.3.4 Similarities with multiple sources.

The proliferation of easy access to electronic sources can result in uncertainty over the true origins of some material, as highlighted by the ambiguity regarding the origin of the opening internet meme above. When analysing a document using plagiarism software, it is not uncommon that specific phrases or word sequences are found to be similar or identical to multiple uncited sources. This would be neither unusual nor unexpected when multiple researchers are working within the same paradigm (Kuhn, 1962). Excluding cases of discipline specific jargon and technical terms, a number of uncertainties arise. For example, if the author did indeed use the words or ideas of another without attribution, it is uncertain from which source the words or ideas were plagiarised. In an environment in which academic misconduct is increasingly common (Singh and Remenyi, 2016) it is plausible that words or ideas have been plagiarised from previously plagiarised sources. While there is more than one way to skin a cat, it is surely impracticable for every author to describe a frequently discussed or universally accepted concept in their very own words. It is conceivable that the specific phrases or word sequences merely constitute common knowledge which, by convention, do not need to be cited.

Ferro and Martins (2016) refer to common knowledge as a “grey area” in academia, citing Neville (2010) in describing it as knowledge that is either commonplace in a specific discipline or field of studies, or that is in the public domain. It is this latter description that has become particularly problematic with the proliferation of online and open access to academic sources. Authors can and do argue quite cogently that online and open
access journals are de facto in the public domain, and as such can and perhaps should be considered to be common knowledge.

4.4 Referencing conventions

There are many referencing software tools available to assist researchers and authors that cater for different budgets and needs. Some of the most well-known are BibTeX, Mendeley, EndNote, Zotero, RefWorks, and Reference Manager. Appropriately attributing material to the original authors and adherence to referencing conventions are considerably easier for authors using these software tools. A major advantage is the ability to reference on the fly, both while compiling notes and in drafting of manuscripts, thereby militating against inadvertently failing to attribute correctly and facing allegations of plagiarism.

An unintended consequence of the use of such software is that authors may pay less attention to the completeness of the details in the list of references, incorrectly assuming that this will be taken care of by the software. These software are only tools and the underlying database is only as reliable as the data stored therein; authors remain responsible for ensuring that all the necessary details of sources are captured correctly.

5. Changing the paradigm of academic writing in Academia 4.0

A combination of passive voice and the third person is frequently used in academic writing. This can result in some ambiguity as to who has carried out the action or activity described. Where an in-text citation follows phrases such as “The data were gathered …”, “Participants were asked …”, “It was found that …”, or “No significant difference was found …”, it is clear that the sentence does not refer to the author’s own work. Conversely, when such phrases are not followed by a citation, by convention it is generally assumed that the action or activity was carried out by the author her- or himself. Therefore, by omitting a citation of the work of others, it may be alleged that the author has plagiarised. This can be unfortunate if the omission of the citation was a mere oversight or pertained to the grey area of common knowledge.

5.1 Justification for a paradigm change

It is clear that there are powerful incentives to plagiarise. Hoover (2006) shows that the incentives make plagiarising a rational choice for individuals, while Necker (2014) concluded that the “publish or perish” culture in academia is a significant incentive for individuals to plagiarise. In response to authors expending undue time and energy merely avoiding plagiarism and adding no value, plus ease of access to prior research, paraphrasing and plagiarism software, and the proliferation of online resources, this author proposes that plagiarism be reframed as a worthless and pointless activity – a complete waste of time and energy. The objective of changing the paradigm of academic writing would be to remove all incentives for authors to plagiarise, but to make it essential to acknowledge prior research in order to be recognised for making a contribution to knowledge.

Fundamental to the paradigm change of academic writing are the criteria against which academic writing is evaluated by supervisors, mentors, examiners, reviewers, editors and the like. The author of this paper proposes a paradigm for academic writing in Academia 4.0 in which:

1. seminal and significant prior research and other sources in the public domain are regarded as common knowledge and need to be acknowledged by citation, but not reviewed;
2. unreferenced statements, claims, findings, etc. are assumed to be drawn from prior research, for which the author takes and is given no credit; and
3. the merit of academic work is judged solely on the basis of that which authors explicitly claim as their own and which patently enhances the extant theory or body of knowledge.

There are various justifications for de-incentivising plagiarism, particularly in literature reviews. Plagiarism consumes and wastes research resources; the time expended by researchers, supervisors, examiners, moderators, editors, and others produces no contribution to scientific knowledge whatsoever. The proliferation of research literature within any given paradigm (Kuhn, 1962) results in unnecessary and similarly unproductive duplication of references. In the foreseeable future, technology in the form of smart search algorithms and natural language generators will be at least as effective as human researchers at summarising and applying prior research. Finally, literature reviews have been found to be relatively poorly valued by the research community (e.g. Madden, 2018; Walker, 1998; Mustaine and Tewksbury, 2016).
While at first sight, such a paradigm might seem to be taking an unduly lenient view of plagiarism, it is suggested that this will be an altogether more rigorous and potentially demanding approach to academic writing. In this paradigm, unattributed findings would be disregarded as academic puffery and poor academic writing, while plagiarism would comprise authors explicitly taking credit and responsibility for statements, activities, findings, etc. that were not their own.

5.2 Direct impact of a change of paradigm of academic writing

The most obviously impact of the proposed change of paradigm of academic writing would be substantial de-emphasis of the traditional literature review. The justification for this is that the seminal and significant published works would have been acknowledged and would be being readily accessible, and therefore restatement would be redundant. In proposing to do away with the literature review in the criminological and criminal justice disciplines Maddan (2018) suggests that researchers would get to the substance of their article much quicker, more rapid article production would facilitate quicker dissemination of research findings, shorter articles would enable the publication of more articles per journal, and the entire publication process would be expedited.

However, authors would need to familiarise themselves and be thoroughly acquainted with the existing theory or body of knowledge in order to lay claim to their unique contribution. Similarly, those evaluating academic work (for example: research supervisors, mentors, examiners, reviewers, and editors) would also need to be sufficiently conversant with the discipline in order to make a meaningful evaluation of the work. Conversations and engagement among stakeholders using academic social networks are set to become essential components of the knowledge creation process. These will be expected to be visible in the public domain as they will replace some of the discourse that currently takes place through journal publications, at conferences or in personal correspondence. All participants will need to be active in these domains to maintain their credibility and influence.

The de-emphasis of the traditional literature review and the corresponding intensified focus on the authors’ explicit contribution would potentially reduce the recycling of knowledge that tends to occur in “normal science” (Kuhn, 1962). This would increase the likelihood of researchers challenging or at least questioning the essential assumptions of current frameworks, leading to paradigm shifts. Similarly, traditional publication metrics may need to be reconsidered as they are framed and understood within the current paradigm.

Finally, authors may need to adapt their writing style, as writing in the first person and active voice – as recommended in APA 6th Edition, Section 3.18 (American Psychological Association, 2010) – would necessarily become the new norm. This would also be consistent with an increasing level of informality that has been observed across various disciplines (Hyland and Jiang, 2017).

While academic misconduct in the form of plagiarism and the related and more insidious practice of ghost-writing would not be eliminated, they would require a greater investment of time and intellect; the cost versus the benefit would largely mitigate against them.

6. Conclusion and Recommendations

There is an evolution in the manner in which academic misconduct is manifesting, driven at least in part by emerging technologies. Acknowledging the transition to Academia 4.0, mirroring the so-called fourth industrial revolution, it is suggested in this paper that there may be an alternative to the traditional definition and response to plagiarism.

It is posited that the relative ease of access to prior research and information has lessened the importance of the traditional literature, and that much more can now be thought of as common knowledge than has been the case historically. The suggestion is that although attribution of sources remains fundamental, in future authors must explicitly claim their own interpretation of and contribution to knowledge, and unattributed or unsubstantiated statements should be disdained and discounted as academic puffery rather than being considered indicators of potential plagiarism.
The proposed reframing of plagiarism and placement of greater emphasis on unique, new knowledge generation may be more intellectually onerous on researchers and evaluators, and the quantity of research may decline. However, it is suggested that the quality and impact of the research will be substantially enhanced.

The practicalities of changing writing standards and effect such implementation would have on academic examination, review and publication processes has not been explored. Given that these standards and processes are well-established, it is clear that implementation will be extremely disruptive (Pinheiro, Cope and Kalantzis, 2019). It is therefore suggested that it is imperative that further research begin on the barriers and enablers of a transition to a new paradigm of academic writing and publication, from the perspective of each and every group of stakeholders.

References


www.ejbrm.com

ISSN 1477-7029
The Electronic Journal of Business Research Methods Volume 18 Issue 1 2020


Trade-Association Endorsements, Postal & Online Responses, and Other Factors Affecting Response Rates: Reflections on a Survey of Micro-Firms

Marc E. Betton, J. Robert Branston, and Philip R. Tomlinson

University of Bath, UK
M.Betton@bath.ac.uk
ORCID ID: 0000-0001-9041-1521
J.R.Branston@bath.ac.uk
ORCID ID: 0000-0002-2332-2403
P.R.Tomlinson@bath.ac.uk
ORCID ID: 0000-0001-6292-3204
DOI: 10.34190/ JBRM.18.1.004

Abstract: Researchers using survey-based methods must do everything they can to enhance their response rates in order to improve the robustness and validity of the data they collect for analysis. Extant literature has explored several techniques to do this across many contexts, including experiments/meta-analyses. This paper addresses two particular deficits by considering the difficult context of micro-firms and also the impact of trade-association endorsement (rather than sponsorship). This is done by conducting an in-depth ex-post review of a single, successful survey case conducted and utilised in a study of English tourism firms and regulation. We explore several factors: the type of trade-association support (including levels of endorsement, access to member firms, and messages sent on behalf of the research team); postal versus online questionnaires; message type/frequency; and the importance of giving respondents an opportunity for open comments. The paper ends with a comprehensive set of suggestions for future researchers.

Keywords: Survey methods, response rate, online questionnaires, trade-associations, endorsement, sponsorship.

1. Introduction

Survey response rates are an important aspect of survey-based research methods (Mellahi and Harris, 2015). Researchers aim for a high response rate to ensure that they have sufficient data (and variation within that data) to run their desired analyses (Baruch and Holtom, 2008). A high response rate is also a robust defence against the risks of non-response bias, external validity, and generalisability (Dillman, Smyth and Christian, 2014; Dillman, 2016). Furthermore, there is evidence that researchers, journal reviewers, and journal editors, generally consider response rates to be an indicator of quality and the validity of survey data (Mellahi and Harris, 2015). It is therefore in the interests of researchers using survey data that all efforts to enhance response rate are taken, especially given evidence that response rates are falling worldwide (Schoeni et al., 2012). However, much of the research-into, or information-about response rates is limited to either experimental surveys (i.e. not ‘real-world’ experience) or the brief information that authors can include in their applied results papers. We believe that researchers have much to gain by sharing their real-world survey experiences with others in more detail.

In our case, a study of owner-managers of English micro-firms, defined here as firms with 0-9 employees (Department of Trade and Industry, 1995; Department for Business, Innovation and Skills, 2016), was undertaken in order to explore regulatory issues, with the applied results having already been published in Betton, Branston and Tomlinson (2019) and Betton, Branston and Tomlinson (2020). The contribution of this paper is to provide a more detailed overview of our survey process than is generally possible in applied academic papers (such as our own), so that future researchers can better understand the successful method utilised in this case involving difficult to survey micro-firms.

The aim therefore, is to conduct an ex-post in-depth review of a single survey case which employed a combination of survey approaches, in order to gain insight into improving response rates, while also identifying useful wider lessons. The case was an independent academic survey in which we made sure to collect extra data on the distribution and response patterns in order to learn wider methodological lessons from the applied work. This paper is therefore not based on a survey designed exclusively to explore survey distributions methods, and is thus somewhat constrained by the data we were able to collect given the nature of the context. For instance,
this means we have data on postal and online questionnaires, but we did not randomly split our survey population in two in order to explore this difference in great detail.

In particular, we examine the impact on response rates of: the effect of trade-association endorsement (including different levels of support and number of endorsements); the differences between postal and online questionnaires; the number, type, and time of messages before questionnaire completion; and the use of a project website; as well as wider lessons, such as the use of an unsolicited comments box. This exploration is made through observation of the methods of communicating with potential respondents and associated mean responses received, as well as an OLS regression model. From this, we would hope future researchers looking to survey a variety of firms on a range of issues, most especially micro-firms, can not only draw upon the techniques which we found to be successful, thereby improving their own response rates, but also consider including similar small-scale tests in their own research, which can then be reported upon.

The remainder of this paper is set out as follows. The second section provides an overview of recent literature relating to response rates, while further literature is also discussed in relevant later sections. The third section explores our case, detailing how the questionnaire was developed and distributed, as well as discussing related aspects to be explored herein, such as the use of trade-association support, and methods of communication and completion. The fourth section explores how the various aspects of our study impacted on response rates, including the communication method (postal versus online), the levels of trade-association endorsement, and the use of an unsolicited comments box. The final section then briefly concludes and summarises our suggestions for future survey-based research (Table 7).

2. Literature Overview

There is already considerable research into various techniques to improve questionnaire response rates. One popular strand involves exploring the impact of the delivery mode through numerous means. For example, Hardigan, Popovici and Carvajal (2016) found postal questionnaires to have a higher response rate than online or mixed mode surveys. However, Beebe et al. (2018) found mixed-mode surveys (and multiple messages) to have a 10% higher response rate than single modes. Bucks, Couper and Fulford (2019) found that while there were differences in sequential/concurrent modes (i.e. web then mail / mail then web, or web and mail), these differences were eliminated by the end of the survey period. Other recent investigations include Lavrakas et al. (2017), who explored envelope colour and mode of address; while Cook et al. (2016) found that non-monetary incentives and a posted reminded had no effect on response rates.

Another strand of literature (e.g. Cucyota and Harrison, 2006; Baruch and Holtom, 2008; Manfreda et al., 2008) adopt a meta-analysis approach of the methods reported in applied published papers. One recent and thorough example is Mellahi and Harris (2015), who ultimately examine the methods of 1,093 papers, and which can therefore be considered to be an authoritative source. They provide an extensive review of extant literature on a number of potential strategies to improve response rates, including survey design, questionnaire distribution, and the use of incentives, which they explore across a number of managerial fields.

It is however, important to note that there are still some contradictions in the literature. An example, Dennis Jr (2003) found that no form of incentive had a meaningful impact on response rates, while Rose, Sidle and Griffith (2007) and Gendall and Healey (2008) found that the inclusion of money (or similarly valued items) with postal questionnaires would improve response rates. One reason for the continued uncertainty is that these findings may be highly reliant on their context, i.e. it may be that monetary incentives improve response rates in populations of the general public, while they have no impact for business populations. In addition, much of the existing literature is response rates (e.g. Millar and Dillman (2011) or Petroviči, Petrič and Lozar Manfreda (2016)), is based on experiments, often exploring just one particular aspect (e.g. mode of communication), with these frequently explored within a classroom setting, rather than reviewing actual fieldwork with real-world conditions. It would therefore seem prudent that researchers in all fields and contexts share as much detail about their methods as possible so that wider lessons can be drawn.

In particular there are a number of areas which remain poorly considered in the extant literature, such as the use of trade-association endorsement. This is materially different to survey sponsorship, as that implies a greater level collaboration between the sponsor and the research team throughout the research process; and has already been well explored (see Presser, Blair and Triplett, 1992; Dennis Jr, 2003; Fan and Yan, 2010;
Edwards, Dillman and Smyth, 2014). Trade-association endorsement implies a lower level of engagement and support of the survey, but could still prove helpful, yet is not as well investigated. Moreover, there is an ongoing debate around the use of online questionnaires, which remain contentious, even though the use of online services seems to have only increased over time (Manfreda et al., 2008; Fan and Yan, 2010; Schley, 2013; Dodou and de Winter, 2014; Zhang et al., 2017). Furthermore, micro-firms are somewhat under-served in academic research (Dennis Jr, 2003; Freelon and Harrison, 2006; Betton, Branston and Tomlinson, 2019) and hence the literature on appropriate research methodology targeted at such small firms is similarly lacking.

Micro-firms, the smallest category of firms (typically defined as having 0-9 employees (Department of Trade and Industry, 1995; Department for Business, Innovation and Skills, 2016)) are notoriously difficult to research, which makes methodological insights especially useful. There are a number of reasons for this difficulty, including that their own internal resource constraints leaves them with little time for participation (Edwards et al., 2002; Sauermann and Roach, 2013) and that they have a common unease for any form of outside interference (Johnson, 2002). Such difficulties result in their either being included as part of a wider ‘small firm’ segment, or excluded altogether (for example: Arrowsmith et al., 2003; Getz and Carlsen, 2005; Servon et al., 2010). It is therefore unsurprising that there are relatively few contributions that offer insights into micro-firm survey engagement, but what is available suggests low response rates are to be expected. For example, Jay and Schaper (2003) cite studies of micro-firms with response rates of 4-6%, while Duarte Alonso, Bressan and Sakellarios (2017) reported a 19.9% response rate. Despite these difficulties, it is essential that such firms are included in research, as they make up 96% of firms in the UK and account for 32% of private sector employment (Department for Business, Energy & Industrial Strategy, 2018). They are similarly important to other economies. For instance, they account for 93% of firms in the EU (Eurostat, 2018), and 79% in the US (US Census Bureau, 2018).

Given the apparent preference for high response rates to ensure robustness of results, there remains a further issue in that different contexts and industries have different expectations for what is considered acceptable (Saunders, Lewis and Thornhill, 2015). For example, Saleh and Bista (2017) report a response rate close to 80% in an educational setting, while others accept as reasonable a rate closer to 8% within an SME context (Brustbauer, 2016). Indeed, some authors suggest any response rate can be acceptable if sufficient tests are made to explore the statistical power of results and the risk of non-response bias (Hair, Black and Babin, 2010; Malhotra, Birks and Nunn, 2017). Nevertheless, it seems important to consider the geographical and industrial contexts at hand when considering response rates alongside checks for statistical robustness. For instance, since micro-firms are generally less likely to respond to surveys (Freelon and Harrison, 2006), lower response rates are to be expected. This makes survey design especially important in this context given that a small increase in response rate can therefore have a relatively large impact. In this regard it is notable that Freelon and Harrison (2006) reported response rates around 11% internationally, while UK-based research, somewhat similar to the study considered herein (although with a good number of larger firms), presents a pattern towards a response rate above 20% (Tomlinson and Fai, 2016; Tomlinson and Branston, 2018). Additionally, a trade-association active in the area of our case previously conducted two surveys of its members on a similar topic with response rates of between 19% and 28% (BH&HPA, 2011).

3. Our Case

The survey explored herein investigated regulation in English micro-firms, with regulation defined as the “imposition of rules by government, backed by the use of penalties that are intended specifically to modify the ... behaviour of individuals and firms in the private sector” (Organisation for Economic Co-operation and Development, 1993, p.73). This is an especially difficult subject and context given the known issues surrounding micro-firm data collection and the somewhat sensitive subject of regulation (since firms do not wish to risk being reported for non-compliance). In particular, we selected the accommodation industry in England because it is highly regulated and has a high proportion of micro-firms (Tourism Regulation Taskforce, 2012; Tourism Alliance, 2016). The wider UK nations were excluded due to regulatory differences between them, which were beyond the scope of the study.

We developed a database of firms from a number of publicly available sources, which was used to directly contact potential respondents. In addition, the database included details of all communications with potential respondents, including the date, method, and content. This level of detail not only aided the data collection process (i.e. allowing us to know who to contact and when), but it also provides much of the data for the analyses.
herein. The survey ran for a six-month period from October 2014. While there may be risk of self-selection bias 
(Saunders, Lewis and Thornhill, 2015), we believe this is mitigated through numerous means. Our distribution 
consisted of directed contact and wider marketing, both with and without trade-association support. While 
there is a risk that ‘only’ those with an interest in the subject were motivated to respond, regulations are a 
central issue to micro-firms and the all-encompassing role of the owner-manager means that they are highly 
likely to have views to share.

3.1 Trade-Association Support

Sponsorship has been frequently found to improve response rates (Fan and Yan, 2010) and this, in the form of 
endorsements, was factored into the study at its inception. Furthermore, there is considerable evidence that 
trade-associations play a vital role in supporting micro-firms (Bennett and Ramsden, 2007; Lawton, Rajwani and 
Doh, 2013; Pleasance and Balmer, 2013; Department for Business, Energy & Industrial Strategy, 2016). Indeed, 
in a regulatory context trade-associations have been noted to “offer some guidance that is more concise and 
reader-friendly than official government advice” (Better Regulation Executive, 2010, p.14). Moreover, in the 
accommodation context, two trade-associations were formed as a direct result of regulatory changes, thereby 
other demonstrating both their role in regulation and their potential role in supporting research (Bed and 
Breakfast Association, 2014; BedPosts, 2014).

We sought trade-association involvement from the inception of the study for several reasons, including: 
endorsement (as a means of promotion and demonstrating legitimacy); access to member firms (to promote 
the survey); and to leverage their unique knowledge of their membership and the industry to design and frame 
the survey questions and communications. Several trade-associations and marketing groups were approached 
throughout the study. In exchange for support they were offered a summary of results, which has since been 
provided. The Bed and Breakfast Association helped in the development of the survey, particularly with the 
questions relating to fire regulations, and they were the first to endorse the study (shortly after the 
questionnaire was launched). As the study and survey continued, a total of four trade-associations endorsed the 
study, while a further three promoted it without formal endorsement. The direct effects of trade-associations 
and endorsements on response rates are explored in Sections 4.2-4.5.

3.2 Questionnaire Development and Format

The questionnaire was original, developed by the authors specifically for the research project exploring 
regulation and micro-firms. Questions were informed by past questionnaires in extant literature, and were 
carefully developed in line with Dillman, Smyth and Christian (2008) and Dillman, Smyth and Christian (2014) 
to ensure a link with the underlying concepts to be explored and to avoid bias such as leading questions. 
Furthermore, scale questions were developed in line with Spector’s (1992) guidelines. The questionnaire was 
pilot tested to explore sample generation, ascertain response rates through various methods of contact, and to 
access the internal validity and reliability of the questions and method utilised. The main trade-associations were 
also approached to check the wording and understanding of questions, while also attempting to secure 
endorsements, with feedback received from the largest. Following this, several small changes were made to the 
questionnaire and data collection plan, resulting in the methods described here. Finally, the questionnaire ended 
with an unsolicited comments box, allowing respondents to make any (or no) comments on the issues raised in 
the questionnaire or of the survey itself. The use of these comments is explored in Section 4.6.

The questionnaire was designed as a small booklet for physical completion, with each postal questionnaire 
containing a unique ID number. This was then transposed into an online questionnaire using the Qualtrics 
platform, which also has a unique ID number for each response. In addition to the question order, fonts and 
formatting were also closely matched. The advantages of the Qualtrics platform allowed the online version to 
use ‘logic functions’ to hide questions which were irrelevant to certain respondents (based on their answers to 
earlier questions), whereas the postal version simply instructed respondents to skip certain questions based on 
the required conditions. Differences between the postal and online versions of the questionnaire, are explored 
in Sections 4.1 and 4.4.

3.3 Communications

The communication strategy for the study broadly followed the pattern laid out by Dillman, Smyth and Christian 
(2008) and updated by Dillman, Smyth and Christian (2014). This included the method, frequency, and content 
of communication. Contact with potential respondents was made either via post or email, and was personalised

www.ejbrm.com

©ACPI
where possible. Potential respondents without email addresses were sent postal questionnaires, as well as randomly selected firms with email addresses. A total of 1,985 postal questionnaires were distributed with freepost return envelopes, which was the maximum possible due to cost and very limited budget. All follow-up messages were sent via email (where an email address was known), while no follow-up messages were sent to potential respondents without an email address, again due to budget limitation. Therefore, our survey included both mixed-modes of communication (postal and email) as well as mixed-modes of response (postal and online) (Edith, 2018).

Messages were broadly sent in six waves throughout the data collection period. This meant there was a constant stream of invitation and follow-up messages being sent at all times, while also ensuring firms had a prolonged period to respond. This was essential in this case given that many small accommodation firms close ‘out of season’. In addition, the waves of communication allowed for constant updates to messages, such as the addition of further trade-association endorsements as they became available. Logos and names of each trade-association that endorsed the study were included on cover letters. The effects of message content and timing of questionnaire distribution are explored in Sections 4.4-4.5.

3.4 Support Website

In order to balance the competing requirements of detail (to enhance understanding and alleviate concerns) and brevity (to improve the chances of being read/acted upon) in communications, a dedicated website was launched to provide greater detail on the study (www.bizsurvey.org). This allowed communications to be more succinct, while offering a link to elaborate with more information. The website was specifically written in a conversational tone, tailored to both potential respondents and other interested parties which may have investigated participation. The aims of the study, along with the state of industry endorsement were included and kept up to date during the study and beyond. The website received approximately 60 visitors a week on average throughout the data collection period, with potential respondents using it to learn more about the survey, contact the research team, and potentially access the online questionnaire for completion. The use of a dedicated website also allowed for bespoke but very similar web links which could be created and distributed as required. For example, each supporting trade-association was given a unique URL, the use of which could be tracked. An analysis of this technique is included in Section 4.3.

4. Analysis and Discussion

4.1 Postal Versus Online

A total of 3,805 questionnaires were dispatched using both postal (hard-copy) and online means. Eight respondents specifically requested that a postal questionnaire was sent to them, and these were either firms that had been contacted electronically or had heard about the survey but had not yet been contacted directly. Such requests were handled in the usual postal manner, including cover letters (which noted the request) and pre-paid return envelopes. A number of responses were completely anonymous. In the case of online questionnaires, these are respondents who followed an anonymous link to the questionnaire (e.g. from the survey website) and chose not to provide their name; in the case of postal questionnaires, this required the respondent to both withhold their name and destroy the unique questionnaire ID number. The majority of responses could be tracked to specific individuals, either through unique postal questionnaire ID numbers or unique Qualtrics questionnaire links (sent out in personalised email messages). The totals of these methods are presented in Table 1.

Table 1: Questionnaire responses received by type

<table>
<thead>
<tr>
<th></th>
<th>Postal</th>
<th>Online</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sent</td>
<td>Received</td>
</tr>
<tr>
<td>Anonymous</td>
<td>n/a 1</td>
<td>n/a 54</td>
</tr>
<tr>
<td>Requested by respondent (%)</td>
<td>8 5 (62.50)</td>
<td>0 0</td>
</tr>
<tr>
<td>Tracked (%)</td>
<td>1977 307 (15.77)</td>
<td>1820 339 (18.63)</td>
</tr>
<tr>
<td>Total</td>
<td>1985 313</td>
<td>1820 393</td>
</tr>
</tbody>
</table>

Note: no respondents requested online links, but 13 respondents completed the questionnaire without prior contact.

Of the 706 total returned questionnaires, 313 were postal (44.33%), 339 (48.02%) were online and trackable, with a further 54 (7.65%) online responses which may be anonymous or those who were sent a postal questionnaire but completed online. This represents response rates of 15.77% for postal, 18.63% for tracked
online, and 18.55% overall. Given the difficulties inherent to our micro-firm context, we suggest this is a respectable return, although the concept of an ‘acceptable’ response rate is somewhat ambiguous and context specific (Mellahi and Harris, 2015; Saunders, Lewis and Thornhill, 2015; Brustbauer, 2016; Cassells and Lewis, 2017). We also note that five out of eight (62.5%) of respondents who requested a postal questionnaire actually completed it. While these respondents were clearly already somewhat engaged with the study given their request, this suggests (although the sample size is small) that providing a hard copy and distributing it when requested is a worthwhile strategy.

While the postal questionnaire itself had a healthy response rate, the tracked online response rate was higher, and greater still when including anonymous responses. However, there were no differences in types of responses between online and postal in terms of the demographics or type of respondents. Furthermore, when conducting our applied analyses there were no statistically significant differences in the overall findings between the online or postal responses. We therefore conclude that ultimately, offering a postal questionnaire is useful (and was preferred in a small number of cases), but the online questionnaire was the fall-back option for potential respondents. The cost of printing questionnaires and return envelopes was over £480, with further outgoing and return postage costs in excess of £2,000. We note that there is some variation in the return postage costs as some respondents (needlessly) added stamps to their freepost envelope, while this often reduced the cost to us, it sometimes actually increased the costs due to the incorrect stamps being added which resulted in a fine and handling charges. We therefore recommend that in addition to printing the address, freepost envelopes print “no stamp required” on the top right of the envelope. In addition, there is a further, largely incalculable but not insignificant time-cost related to creating the postal packages and entering questionnaire ID numbers into the database. Moreover, the responses of each returned postal questionnaire had to be entered manually into the dataset. By comparison, email addresses for online recipients were exported from the database into Qualtrics, messages were sent, and this was all logged in the database within minutes. Responses had to be checked, but not transposed into the dataset. While there is a cost to the various applications in use in the online distribution process (Qualtrics and Microsoft Office, as well as STATA for analysis), through the use of standard university licences, the direct cost to the study was negligible and most of these would have been required even if the questionnaire had been entirely postal in nature.

Given the monetary cost and time implications, future studies should not consider ‘online only’ as a handicap, most especially given the higher response rate for that medium. We also note that across all analyses we have undertaken, there has been no statistical difference between the results from the postal and online questionnaires, indicating that the questionnaire mode has no discernible impact on the responses given.

4.2 Trade-Association Support and Direct Membership Access

There are several ways that trade-associations can provide support for studies such as ours. Firstly, as detailed above, they can help with the refinement of the survey questions, providing feedback on understanding and suitability based on their experience. In our case, the main associations were approached, with several suggestions made to improve wording, which were incorporated before the survey was launched (not all suggestions were accepted in the interests of academic independence).

Once the survey is launched they can act as a source of a list of potential respondents, allowing researchers to have direct access. This could be achieved by formally providing a restricted mailing list of members. However, such availability will be unlikely in many contexts, particularly if a formal endorsement of the study is not forthcoming. It can also be achieved when membership lists are provided publicly. In our case of the accommodation sector, membership lists were sometimes publicly available through a centralised website or guidebook, such as the Farm Stay UK website which details all Farm Stay UK members.

Trade-associations can also support studies through formal means, the most prominent being official endorsement, whereby the association publicly backs the study. In our case, four associations ultimately endorsed the study and we highlighted these endorsements in all communications and on the support website. As part of this endorsement, the trade-associations implicitly approved of the contents of the (already launched) questionnaire. A further form of support is by trade-associations communicating with their members on behalf of the study, such as by sending information about the study to members. This may be in the form of an article in existing publications, or a focussed, dedicated message. While it carries an implicit endorsement, it may or may not include a full formal endorsement. Ultimately, three associations sent communications about our study to their members without a formal endorsement. The form of support we received from several trade-
associations is presented in Table 2, along with the number of firms held in the database and the number of respondents.

Table 2: Trade-Association support and membership response rates

<table>
<thead>
<tr>
<th>Trade-Association</th>
<th>Endorsement</th>
<th>Communications sent on behalf</th>
<th>Membership access</th>
<th>Number in the sampling frame</th>
<th>Number of respondents (% of database)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Farm Stay UK</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>604</td>
<td>202 (33)</td>
</tr>
<tr>
<td>BedPosts</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>533</td>
<td>67 (13)</td>
</tr>
<tr>
<td>BH&amp;HPA</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>557</td>
<td>94 (17)</td>
</tr>
<tr>
<td>The Caravan Club</td>
<td>✔</td>
<td></td>
<td>✔</td>
<td>1101</td>
<td>133 (12)</td>
</tr>
<tr>
<td>AA</td>
<td></td>
<td></td>
<td>✔</td>
<td>332</td>
<td>33 (10)</td>
</tr>
<tr>
<td>Bed and Breakfast Association</td>
<td></td>
<td></td>
<td></td>
<td>n/a*</td>
<td>35 (n/a)</td>
</tr>
<tr>
<td>(BBA)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>British Hospitality Association</td>
<td></td>
<td></td>
<td></td>
<td>n/a*</td>
<td>8 (n/a)</td>
</tr>
<tr>
<td>(BHA)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VisitEngland</td>
<td></td>
<td></td>
<td></td>
<td>n/a*</td>
<td>52 (n/a)</td>
</tr>
</tbody>
</table>

*Note: as these trade-associations do not publicise a membership list, they were not used to populate the database.

Table 2 suggests that trade-association support through communication (such as an article in their publications) is not enough to engender a sufficient level of response for quantitative analysis. Direct communication between the researchers and potential respondents seems to be a key requirement.

Farm Stay UK was the only trade-association to completely engage with our survey: they endorsed it, contacted their membership on behalf of the study, and provided contact information for their membership (through their public website) allowing for direct communication. This resulted in the highest trade-association response rate (33%). Following this, the trade-associations allowing direct access to members generally provided a higher response rate than those that did not. While it is not possible to calculate the response rates for those without membership access, if we surmise that communications were indeed sent to the maximum stated membership (which is likely an overestimation and copies of apparent messages were not provided to the researchers in all cases), the response rates for the BBA, BHA, and VisitEngland would each equal less than 1%. This is notably smaller than the cases where direct membership access was possible. Indeed, from this, we suggest that direct membership access, even if through public sources, is of greater advantage for response rates than either trade-association endorsement (formal or otherwise), or communications sent on behalf of the study. However, endorsement AND membership access likely yields the best response rate.

4.3 Source Tracking Links

To further ascertain the value of the various means through which potential respondents could reach the questionnaire, we developed a technique to track each individual source. Through the domain of the support website a number of unique and individually identifiable links to the online questionnaire were created. This meant that (for example) a response via the study support website could be identified separately from a response via the address on the postal questionnaire. Individual links were also given to each support trade-association to be used in direct messages to their members and/or in feature articles detailing the survey. These links and the number of responses using them is included in Table 3. Note that responses made via links in emails (which were individually tracked through Qualtrics) and postal replies are not included in the table.
Table 3: Responses following unique links

<table>
<thead>
<tr>
<th>Source of weblink</th>
<th>Responses via the link</th>
<th>Number of respondents reporting membership</th>
</tr>
</thead>
<tbody>
<tr>
<td>Postal questionnaire</td>
<td>67</td>
<td>n/a</td>
</tr>
<tr>
<td>Study support website</td>
<td>31</td>
<td>n/a</td>
</tr>
<tr>
<td>BedPosts*</td>
<td>10</td>
<td>16</td>
</tr>
<tr>
<td>Visit England*</td>
<td>5</td>
<td>52</td>
</tr>
<tr>
<td>Links requested by specific individuals</td>
<td>1</td>
<td>n/a</td>
</tr>
<tr>
<td>Farm Stay UK</td>
<td>0</td>
<td>202</td>
</tr>
<tr>
<td>Bed &amp; Breakfast Association</td>
<td>0</td>
<td>35</td>
</tr>
<tr>
<td>British Hospitality Association</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>Total</td>
<td>114</td>
<td>313</td>
</tr>
</tbody>
</table>

*Note: these trade-associations had to be manually entered into the ‘other’ box.
Note: respondents could state membership of any number of trade-associations, but would only have responded once and hence used one weblink.

Table 3 demonstrates that this source link alone is not a suitable method for tracking how respondents accessed the questionnaire, given the disparity between the number of links used and the number of members in a group. This is to be expected as multiple entryways were encouraged to ensure ease of access to the questionnaire. For example, a respondent who was sent a postal questionnaire could have: i) used the postal questionnaire; ii) used the weblink included in the postal questionnaire; iii) visited the support website and followed the link there; or iv) picked a link from one (or more) of the trade-association publications/communications they were sent.

While there is much uncertainty as to how respondents came to be aware of the survey, it seems clear that reliance on third-party publications is unlikely to yield a high response rate, especially when viewed in conjunction with trade-association endorsement levels (see Section 4.2 and Table 2). For instance, we believe the BBA sent an email to over 8,000 members and associates (the exact figures were not offered), yet none followed the bespoke link they were provided.

4.4 Message Type and Endorsement Totals

We next explore the messages that were sent to prospective respondents, using correlations and an OLS regression model to investigate a number of factors which may have affected the response rate (i.e. the dependent variable). The model was specified as:

\[
\text{Response Rate} = B_0 + B_1 \text{Message method} + B_2 \text{Survey period} + B_3 \text{Message type} + B_4 \text{Number of trade - association endorsements} + \epsilon_i
\]

Dependent variable - Response Rate: the percentage of each set of messages sent to potential respondents that resulted in a completed questionnaire response (i.e. if 100 messages are sent out in a particular wave, how many of these resulted in a submitted questionnaire response).

Control variables - Message Method: a categorical variable indicating if the message is postal (0) or email (1). Survey Period: a categorical variable indicating if the message was sent during the first (0) or second (1) half of the six-month data collection period, thereby representing seasonal variation (as many of the target firms are closed during the ‘off-season’). Message Type: a categorical variable indicating the general content of the message, invitation (0), first follow-up (1), second follow-up (2), or third follow-up (3).

Independent variable - Number of Trade-Association Endorsements: indicating the number of formal trade-association endorsements which had been made and were therefore detailed at the time in each message (a raw number from 0-4).

Table 4 presents the correlations between these variables, then Table 5 presents the regression results, first with the control variables, then with the independent variable.
Table 4: Variable correlations

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Response rate</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Message method</td>
<td>-0.24</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Survey period</td>
<td>-0.27</td>
<td>0.24</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Message type</td>
<td>-0.23</td>
<td>0.43*</td>
<td>0.22</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>5. Number of Trade-Association endorsements</td>
<td>-0.11</td>
<td>0.07</td>
<td>0.65*</td>
<td>0.18</td>
<td>-</td>
</tr>
</tbody>
</table>

*p <0.05

There is no statistically significant correlation between response rate and trade-association endorsements, lending further credence to the notion (detailed in Section 4.2) that endorsement alone does not drive responses. There is an expected, though unimportant, correlation between message type and message method, which simply reflects the aforementioned point that all of the follow-up messages were via email. There is also an expected correlation between trade-association endorsements and survey period, given that the number of endorsements increased over time. Given such correlation might indicate multicollinearity, the VIFs for each variable were checked, and none were bigger than 1.84 which is well within the acceptable range (Pallant, 2011; Tabachnick and Fidell, 2013).

Table 5: OLS regression. Dependent variable: Message Response Rate

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>12.42*** (3.01)</td>
<td>11.64** (3.64)</td>
</tr>
<tr>
<td>Message method</td>
<td>-2.01 (3.32)</td>
<td>-1.80 (3.42)</td>
</tr>
<tr>
<td>Survey period</td>
<td>-3.37 (3.18)</td>
<td>-4.46 (4.23)</td>
</tr>
<tr>
<td>Message type</td>
<td>-0.85 (1.44)</td>
<td>-0.92 (1.48)</td>
</tr>
<tr>
<td>Number of Trade-Assocation endorsements</td>
<td></td>
<td>0.58 (1.44)</td>
</tr>
<tr>
<td>R²</td>
<td>0.12</td>
<td>0.12</td>
</tr>
<tr>
<td>Adjusted R²</td>
<td>0.01</td>
<td>-0.03</td>
</tr>
<tr>
<td>F (3, 24)</td>
<td>1.06</td>
<td></td>
</tr>
<tr>
<td>F (4, 23)</td>
<td></td>
<td>0.81</td>
</tr>
</tbody>
</table>

*p <0.1, ** p <0.05, *** p <0.001. Note: Standard Error in parentheses

We note that none of the variables included in the model are statistically significant. While this would typically be cause for concern in econometric studies, in this case it simply suggests that the response rate of each message is not associated with the message method, type, or the time of year. This is a non-trivial result, which not only lends credence to the validity of the data we collected for the study of regulation in micro-firms, but also suggests future researchers should not be fixated on any one particular approach for raising response rates. Furthermore, not only is there also no association between the response rate and the number of trade-association endorsements, but introducing this variable has both no effect on the R², and also causes the adjusted R² value to become negative, indicating that trade-association endorsement really does not explain the response rate to our messages (Wooldridge, 2009; Hair, Black and Babin, 2010; Warner, 2012). This further supports our other results so far, that endorsement alone does not influence respondents towards completing a questionnaire.

4.5 Number of Messages and Time to Completion

Respondents were sent between one and five communications during the survey period, the summary statistics of which, along with those for the number of days between the last communication and completion, are included in Table 6.

Table 6: Summary statistics for number of messages sent and the number of days to questionnaire completion

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of messages before completion</td>
<td>2.55</td>
<td>1.07</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Number of days between last message and completion</td>
<td>7.82</td>
<td>8.87</td>
<td>0</td>
<td>57</td>
</tr>
</tbody>
</table>

Note: that postal questionnaires necessarily took longer to return as they had to be posted out, completed, posted back, and logged; whereas online responses could be sent and then completed in a matter of minutes, and were automatically logged immediately.
This demonstrates that while 17% of respondents completed the questionnaire after receiving only one message, it generally required one or two further messages before respondents actually undertook the questionnaire to completion. In most cases, the second message was the one which included either the postal questionnaire or a link to the online questionnaire; although we note that many of the pre-notice messages included a link to the support website, from which respondents could choose to undertake the questionnaire right away. Indeed, 73% of respondents completed the questionnaire after a total of three messages, which in most cases will have taken the form of “pre-notice > invitation > follow-up”. The fourth and fifth messages brought in successively fewer responses each round which is to be expected but suggests some respondents do respond to ‘nagging’.

Table 6 also provides the mean number of days between the final communication with each respondent and their completion of the questionnaire. The mean delay of almost a week could be partly related to the time needed for postal questionnaires to be returned, but likely also demonstrates how busy the target owners-managers can be. Despite this, 25% of responses were received on the day of the last communication, which were all via the online questionnaire. Generally, the longer return times, such as the longest at 57 days, were respondents who were sent a postal questionnaire and not sent any follow-up messages because they had no email address for such messages. Only 15% of responses took longer than 14 days to return. This demonstrates that multiple messages are required to boost a response rate, but after a total of three communications, there will be diminishing returns. It also demonstrates that an appropriately prolonged survey period will likely improve results as it allows even the busiest firms to participate when it suits them.

4.6 Unsolicited Comments Box

While not related to response rates, we found an additional area of interest in our questionnaire construction, which we believe to be of use for future researchers. The final question on the questionnaire was an open comments box, with the question stem of “is there anything else you would like to tell us about any of the issues raised in this survey?”. This allowed respondents an opportunity to comment on any aspect of the issues raised or the survey process itself. While the 197 comments were interesting in and of themselves, they were also put to use in our analysis of regulation in micro-firms. In lieu of wide-scale supplementary interview data, which could not be collected at the time, these comments were formally analysed using a constant-comparison thematic method, as defined by Krueger and Casey (2009). Using this technique, comments were coded relating to their content; these comments were compared with each other to limit duplication (thus, the constant-comparison). The codes were later compared again and grouped together to form themes, which were then used for interpretive analysis, with direct quotes included to provide illustrative examples of the themes. The resulting 10 themes were then included to provide a flavour of the sentiments of respondents and to suggest possible causation for the results of various statistical analyses. Such an approach is in keeping the various uses for such ‘final comments’ that have been explored in prior literature, as summarised by Cynthia and Matthias (2016). We therefore recommend that future researchers provide a means for respondents of closed-questionnaires (or even questionnaires with some open-questions) to comment freely. These remarks should then be analysed, not necessarily in lieu of interview data, but ideally to support it by, for instance, suggesting areas to investigate during in-depth interviews.

5. Limitations

This paper has a number of limitations in addition to the possible transient nature of our discoveries and the associated suggestions for possible ‘best practice’. To start, our insights derive not from a dedicated study with a primary focus on exploring different distribution methods, which would have used slightly different methods, but from detailed reflection of what we found when doing our primary study on micro-firms and regulation. This was made possible by recording detailed data on our questionnaire distribution practices and the associated responses secured. Had this been a dedicated methodological study entirely focussed on distribution methods, we undoubtedly would have undertaken the process slightly differently. Furthermore, this means the analysis herein is based on one single study and so there is no opportunity for direct comparison with similar studies. Nevertheless, there is still value in our findings given the paucity of methodological insight for researchers investigating micro-firms and trade-association endorsements.

A number of other limitations were also created due to our efforts to favour multiple survey entry points to ease access to the questionnaire (to enhance response rates), which is especially important when studying micro-firms (and why including non-essential questions such as “how did you find the survey” is not possible). For
example, while we have explored the number of endorsements, we could not explore the effect of no endorsements or different combinations of endorsements, as this would have required multiple versions of communications and the website (which were not possible).

Similarly, every potential respondent was offered the opportunity to enter the prize draw. Exploring the impact of this would again have required more detailed tracking, multiple communications, and websites. This was not pursued due to resource limitations (requiring us to focus on our primary study) and the fact that there already exists a considerable literature on the impact of prize draws (although not on micro-firms in particular) (e.g. Cobanoglu and Cobanoglu, 2003; Dennis Jr, 2003; Alderson and Morrow, 2004; Gendall and Healey, 2008; Groves and Peytcheva, 2008; Snyder and Elliard, 2012). In detail, we opted to provide a prize draw of four Amazon vouchers with values ranging from £25 to £100. In order to qualify, respondents had to complete the entire questionnaire (or the relevant sections where some could be left incomplete) and explicitly demonstrate their wish to be considered for the prize by entering their contact details at the end of the questionnaire. A total of 339 respondents (out of 706 total responses received) were ultimately identified as being eligible for the draw. Only one respondent outright rejected the prize draw, using the unsolicited comments box to complain about the choice of Amazon vouchers (for personal reasons relating to Amazon’s policies). Similarly, a number of individuals did not complete the name box, thereby deliberately being ineligible for the draw. We therefore have no reason to suspect that incentivising via a prize draw had a detrimental impact on response rates and could well have had a positive impact.

6. Conclusions and Suggestions

The aim of this paper was to provide future researchers with greater insight in what did and did not affect response rates in our context of English accommodation micro-firms (and likely, micro-firms in other industries) and other wider aspects. We accept that there are a number of limitations, especially since this was a secondary part of our study, so we therefore present tentative suggestions for consideration. However, given the unusual micro-firm context, the novel inclusion of trade-association endorsement, and the use of the unsolicited comments box, there is still much to be learned from our efforts, especially given the dearth of literature in this area. We therefore hope that future researchers will not only seek to include these difficult-to-research firms in their own work, but that they will also enjoy greater success when doing so.

Table 7 presents a summary of the suggestions made following our review of our own survey experience. In addition, we recommend that future researchers include similar small reflective exercises in their own studies, as we have done, and report the results in some form. Such small-scale tests not only provide rich information which may be of use to other researchers, but they can also be used to further explore the data which has already been collected for a given study.

Table 7: Suggestions following our study

<table>
<thead>
<tr>
<th>Paper Section</th>
<th>Suggestions</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>While many researchers may feel a simple mailing list will suffice, a dedicated database of potential respondents allows for a greater level of detail and tracking. This offers a centralised hub for managing communications and provides the opportunity to investigate the data collection process, as we have done here. Additionally, by using participant codes, the database and dataset (of responses) can be stored independently, reducing the risk of data breach. We used Microsoft Access, but any similar database application would work.</td>
</tr>
<tr>
<td>3.4</td>
<td>A dedicated website for a study can perform a number of beneficial tasks. Firstly, it permits researchers to provide extra detail on the study, thereby allowing communications with prospective supporters and respondents to be more succinct. Secondly, it allows for the easy creation of branded (short) links for use by study partners. Finally, it acts as a central hub for a study, both at the data gathering stage and beyond, thereby demonstrating the impact of research and legitimising the researchers involved to potential respondents/endorsers.</td>
</tr>
<tr>
<td>4.1</td>
<td>Providing a hard copy questionnaire (even if the focus is an online questionnaire) and distributing it when requested, will likely be beneficial.</td>
</tr>
<tr>
<td>4.1</td>
<td>Given the monetary cost and time implications, future studies should not consider ‘online only’ as a significant handicap, especially given the higher response rate we found for that medium.</td>
</tr>
<tr>
<td>4.1</td>
<td>In addition to printing the address, free-post envelopes should also print “no stamp required” on the top right of the envelope, to reduce the number of respondents who add (often incorrect) postage stamps.</td>
</tr>
</tbody>
</table>
| 4.2           | Direct access to potential respondents (or members of trade-associations), even if through public sources, is of greater advantage for response rates than either trade-association endorsement (formal or
otherwise), or communications sent on behalf of the study. However, endorsement AND membership access likely yields the best response rate.

4.3 While useful for publicising a study, reliance on third-party publications (e.g. VisitEngland’s Quality Edge magazine) is unlikely to yield a high response rate. This follows our suggestion that direct access to potential respondents yields a higher response rate.

4.4 While it may not be true in all contexts and industries, we found that the response rate to messages (to potential respondents) is not influenced by the method (postal or email), type (invitation or follow-up), or time of year.

4.4 There is also no association between the number of trade-association endorsements and the response rate of a message (to potential respondents), further supporting our other results, that endorsement alone does not influence respondents towards completing a questionnaire.

4.5 Multiple messages are required to boost a response rate, but after a total of three communications, there will likely be diminishing returns. A prolonged survey period will likely improve results as it allows even the busiest firms to participate at a time that suits them best.

4.6 We suggest that future researchers provide a means for closed-questionnaire respondents to speak freely. These comments should then be analysed, not necessarily in lieu of interview data, but to support it and to suggest areas to investigate during in-depth interviews.

Acknowledgements

The authors would like to thank all of the survey respondents, and those who helped in the refinement and distribution of the survey, including: David Weston at the Bed and Breakfast Association; Andy Woodward at Farm Stay UK; Martin Couchman and Julia Svetlosakova at the British Hospitality Association; Sharron Orrell and Jane Darragh at VisitEngland; Gené Jeffrey at BedPosts; Jamie Hurst at The National Caravan Council; Gregory Yeoman at the Tourism Society; and Ros Pritchard at the British Holiday and Home Parks Association. The responsibility for the work, however, remains entirely with the authors.

References


BedPosts. 2014. Conversation with Gené Jeffrey, committee member of BedPosts. (Personal communication, November 2014)


Book review: Pat Bazeley’s ‘A practical introduction to Mixed Methods for Business and Management’

Review by EJBRM Editor, Ann Brown
DOI: 10.34190/JBRM.18.1.005

A Practical Introduction to Mixed Methods for Business & Management
Author: Pat Bazeley
Publisher: Sage, Los Angeles
Year of publication: 2019

Mixed Methods is a research methodology that integrates several types of data (both qualitative and quantitative) for a research project, using more than one analytical technique. This book gives an excellent introduction to the methodology, focusing on the issues involved in the handling of multiple approaches. The readership is assumed to have experience of business research and at least the skill to apply some of the current research methodologies. A mixed methods researcher needs not only an awareness of the main research methodologies, but also the ability to learn how to use them or has access to collaborators who do have the appropriate complementary expertise.

Mixed methods is a particularly powerful method for business and management research. Business problems typically involve specialist discipline facets embedded in and significantly affected by the social context. In such a situation both the hard quantitative data measuring performance and the qualitative information explaining the community responses can throw light on the problem and its interpretation. In chapter 1, the author makes a strong case for the great potential value of combining more than one research method. As she states (P13), using these methods ‘to assess the same or different facets of an entity or an experience can lead to enriched, elaborated and/or expanded understanding of that phenomenon’. The effective integration of the insights developed from each and every approach used is needed to obtain this value.

Chapter 2 is the core of the book. It addresses the question as to whether mixed methods is appropriate for your research problem. The five component model (figure 2.1) offers an excellent overview of the major elements of any research project design. This model suggests that the final decision to adopt (or not to adopt) mixed methods comes when the three elements of goals (setting the direction of the research), conceptual framework (for the research subject) and research questions (actionable and specific) have been clarified.

A valuable feature of this book is the use of such wonderful case examples. As each key concept or activity is introduced a case example accompanies it as illustration. The author uses a research project carried out by herself to show how each of the steps described for developing a mixed methods study might work out in practice. Indeed the book greatly benefits from the long-time personal experience of the author in using this type of research approach.

The book has an easy to read writing style. The five chapters follow a well thought out logical sequence of topics. Chapter 1 establishes what can be achieved with mixed methods. Chapter 2 answers the question – Do we have a mixed methods research project? Chapters 3 and 4 address the design of the various aspects of such projects - the choice of data types and samples, data collection, the methods of analysis of the various data elements both individually and in combination so as to achieve the full value of the method through effective synthesis. Chapter 5 views the research project as a whole and brings together all the elements discussed in the previous chapters. I highly recommend this book to aspiring mixed methods researchers.
EJBRM Volume 18 Issue 1, 2020

Editorial by the Editor: Ann Brown

Invitation to our readers to send in their opinions and views by letter to the editor
One of the papers in this issue raises an issue central to academic writing and reviewing – plagiarism - and proposes a controversial change to current accepted practice (Re-imagining Academic Writing to De-incentivise Plagiarism by Anthony Stacey). The paper proposes a paradigm shift in how the academic community might manage referencing. We feel that many of you will have strong views on this proposal and we want to hear from you; indeed, we are always interested in well-argued points of view. All your letters will be read by the Associate Editors and selected ones will be published in a later issue.

This issue also includes a book review of a remarkable book on Mixed Methods (Pat Bazeley’s ‘A Practical Introduction to Mixed Methods’). We would like to publish more book reviews. If you are aware of a recently published book on research methods, which you would recommend to our readership, please send us the reference with your reasons for proposing it.

The research papers include papers on:
- Data handling
- Case application of mixed methods
- Theory development

Data handling - Two of these four papers focus on ways to enhance the collection and analysis of data, using case material to illustrate their ideas. These are:
- A new framework for identifying the key themes of qualitative data emerging from the initial analysis (Alistair J Campbell)
- Ways to increase response rates of surveys in the specific projects involving micro-firms in England (Marc E Betton, J Robert Branston and Philip R Tomlinson)

Case application - an example of how to apply Mixed Methods which also offers a critical assessment of the approach (Oluka Pross Nagitta and Marcia Mkansi).

Theory development - The fourth paper proposes a method of formative scale development – a theoretical paper for the handling of quantitative data (Dawn Hall and Joshua D Shackman)

Alistair Campbell addresses the problem of making sense of the ramifying set of nodes and themes that usually emerge from the first attempt at analyzing qualitative data. His proposed framework aims to determine a hierarchy of themes by using the attributes that are found in the database itself and hence use the expert knowledge collected in the database. The framework is applied to two mini cases.

Response rates for surveys are a key measure for establishing the credibility of findings. The paper by Marc E Betton and colleagues gives us an in depth review of the survey process adopted for their research project on Micro-firm owner managers’ views of regulation in England. The aim is to establish which, if any, of their actions to get more responses from the micro-firms’ managers offered the best results.

Oluka Pross Nagitta and Marcia Mkansi describe their application of mixed methods in their research project to identify the key factors in the successful management of the supply chain of medicines for hospitals in Uganda. They make the case for the value of mixed methods for problems of this nature – complex embedded supply chains. This paper is an excellent example of how to use mixed methods as it not only describes the application of a mixed methods approach in detail but assesses the strengths and weaknesses of using it.

The paper by Dawn Hall and Joshua Shackman contributes to theory creation for quantitative data. Two approaches to scale creation – reflective and formative - are compared and ‘an empirical example of the hybrid approach of measuring a scale formatively for the first level and reflectively at the second order level’ is then proposed.

ISSN 1477-7029