Abstract: As travellers, we are usually aware that a map is not the territory it represents. However, as researchers, inquiring into practice, are we always aware of the domain within which that practice is situated? Descriptions of practice sometimes suggest that this is not the case. For example, do engineers actually believe that the models they develop and use are reflections of some reality? It is likely that an engineer never actually follows his models when developing an artefact or process. Similarly, we can ask ourselves whether we believe that a chef actually cooks by following a recipe. Possibly, only someone who does not know how to cook would think so. These idealised models are simply the basis for discussion/reflection and experimentation. It is sometimes the case, however, that descriptions of practice are produced based in a kind of rationality that suggests these misapprehensions are appropriate. In the context of research, can we say that postmodernism has any relevance? If, in the field of practice, only the uninitiated ever had illusions that the ‘grand theories’ of ‘modernism’ could be directly applicable, then informed research must recognize this also. To those with no illusions, such ‘grand theories’ were a basis for reflection and critique. Thus, to this extent we have always been ‘modern’ and still are. Rather than espousing a Postmodernist perspective, we might point to ‘Hypermodernism’ – a recognition that the ‘grand theories’ can only be used as metaphors, i.e. a basis for practical philosophy. By adopting such a stance, it is possible to avoid a false step of fighting ‘straw men’ and dismissing as worthless research that which could be useful material for reflection and learning when juxtaposed with other perspectives on practice. Models and explanatory frameworks within which research has been conducted need not be rejected as ‘modernist’ if there is recognition of their useful role as metaphor. At the same time, we suggest a need for a critically-informed approach to research which sheds light upon taken-for-granted assumptions and naïve rationalities, illuminating metaphor and stimulating reflection.

Keywords: metaphor; reflective practice; postmodernism; critical systemic thinking; contextual inquiry

1. Background

We are aware that a map is not the territory it represents. See, for example, Foucault’s (1973) reflections on Magritte’s work ‘Ceci n’est pas une pipe.’ (This is not a pipe). However, in some fields of life, an illusion can arise that models do in fact constitute reality. Some people may, for instance, believe that a chef cooks by following a recipe setting out ingredients and method of preparation and cooking. However, someone who knows how to cook would immediately point out that it is their experience and ‘feel’ for the process and context of cooking which is in play – the recipe is a guide only. In a similar way, we could ask ourselves whether engineers actually believe that the models they develop and use are reflections of some reality or are they simply the basis for discussion, reflection and experimentation? In our experience, an engineer does not necessarily approach her work as an exercise in applying models when developing an artefact or process. Such a view could be regarded by practising engineers as naïve and reductionist. Models are often developed under ideal conditions (e.g. in a laboratory) where the context of their application can be constrained. However, in the world of everyday practice, an engineer’s work is situated and contextual. Ideal models are useful, just as a map is useful to a traveller, but alone they are insufficient. Claudio Ciborra reflects upon this when he says:

‘A good example is the adventurous (and long) life of the Russian MIR space station. Up there, revolving in space, one could find, hand in hand, advanced, robust engineering solutions, rustic design, and widespread virtuoso tinkering … to keep the equipment and the system going as a whole.’
Imagine you are looking at planet Earth from a satellite in outer space (see Figure 1). What do you see – “The World”? In this particular view, the North American continent is most prominent and the United States seems to be the focus of attention. The satellite in question may, for instance, be in use to enable Americans to talk to each other by telephone. The viewer could interpret “The World” as a having a particular meaning, reflecting this viewpoint. The British Prime Minister was recently reported to have said “The World is becoming more global”. What could this mean? More global than what? In what context can this statement become meaningful? Some people reading this might think of the Earth as an oblate spheroid (that is what we were taught at school). However, unless the Prime Minister was addressing members of the Flat Earth Society when he made his statement, it seems unlikely that he referred to the shape of our planet. An atlas will frequently include maps depicting a number of different views of “The World”. These could show political boundaries, geological terrain, climatic regions, zones of common religion or language, etc. When talking about “The World”, an individual does not only denote its physical characteristics but their personal perceptions of other qualities – “Their World” as they experience it. Rather than “The World”, it is a description of “A World” – from a particular observer's point of view. As human beings, we have Weltanshauungen (images of the world) that frame what we see and experience (Checkland, 1981) and are not necessarily concerned with the planet we inhabit.

The focus of Postmodern discourse has been to criticise many theoretical perspectives from the modern world as naïve and unrelated to human experience of practice. This discourse appears to be grounded in an assumption that methods (as expounded in Grand Theories) were intended to be applied directly. However, we are unconvinced that the exponents of these methods deserved such a label of naivety. The relationship between methods as a concept and application of method in context is a delicate one, easily misrepresented in formal discourse. Postmodernism places emphasis on a discontinuity between the past, characterised by application of theory, and a present characterised by virtuality and ‘the death of reason’ (Power, 1990). It has been argued (see Burrell, cited in Grant et al 1998) that linearity is a key feature of Modernist discourse. This can be demonstrated in importance attributed to adherence to ‘coherence, order, regularity, prediction, and linguistic certainty’ (Grant, et al 1998, p.10). These are adjectives closely associated with the concept of method. Burrell points to efforts by Postmodernist researchers to move organizational discourse away from such linearity, e.g. Deleuze and Guattari's ideas of nomadism, in which a historic linearity is replaced by a geographic metaphor in organizational analysis (Burrell, 1998). A response to this rejection of method has resulted in some quarters in recourse to the narrative as a more fluid and less restrictive form in which discourse may be continued, liberating individual experience and the use of metaphor. However, this most Postmodernist of techniques has also been subject to criticism. Gabriel (2004) points to the seductive qualities of narrative. Theoretical frameworks, particularly those formulated in a spirit of Logical Empiricism, have tended to constrain the types of evidence legitimised in organizational discourse. However, freed from such restrictions, the power of imagery and the emotive qualities of described experience can be harnessed very persuasively. Writing of his reflections on poetic licence in storytelling, he states “I have long found this view that the truth of a story lies in its meaning rather than in its accuracy compelling. I have now developed serious doubts and have come to regard it as a comforting but inadequate rhetorical gesture where proper argument is called for. Could it be that a story deceives us precisely because its meaning rings true? Could it be that the more authentic a story seems, the more reason we have to approach it with caution?” (p.20). The balance between meaningfulness and verisimilitude in storytelling is highlighted here as problematic. It is interesting to compare this view to Claudio Ciborra's (2002) comments on the willingness of managers to
seize upon models, methods and other ‘apparitions’ in an attempt to deflect their discomfort with the uncertainties of organizational life.

Hans-Erik Nissen (2002), in discussing perceptions of software development failures, makes the point that investigators of such failure seldom point to technological shortcomings. Instead, it is social phenomena which are at the heart of the problem: issues of power, communication, individual and organizational learning processes and the ‘need to make embedded decision rules and their limitations visible.’ He goes on to point to an imbalance in impact of values in the process of creating software. Values of the producers are concerned with functionality and profitability, and to the extent to which use is considered it is by reference to notional ‘users’. Nissen makes the point that these ‘users’ are people who are unlikely to define themselves through their software use, but probably think of themselves as doctors, lawyers, carpenters, store managers, on one hand, and wives, fathers, friends or colleagues on the other. Their position in the relationship is further qualified by use of metaphors by developers which may have a side effect of degrading the self image of the so-called users, e.g. ‘software as an intelligent agent’ (in comparison to whom?); ‘computers as a mind’ (do people act upon formal logic?); ‘data systems as a conduit for information’ (hiding the enormous amount of mutual learning underpinning interpersonal communications which cannot possibly be replicated in human-computer interactions). Nissen’s conclusion is that individual software development practitioners would benefit from reflection upon this imbalance, and giving consideration to hidden conflicts that cannot be solved by elegant artefact design alone. This might improve chances of avoiding failure by recognition of ‘impossible tasks’. In his discourse, Nissen does not criticise formal methods of software engineering; indeed, he scarcely mentions them. His discussion is about those other aspects of practice which are critically impacting upon the way in which these methods can become feasible to create software that real human beings could perceive as ‘useful’ in the midst of life as it is lived.

Figure 2: “The business process”

In a similar way, it is possible to consider reasoning about business processes. It is possible to conceive of a business process as a transformation of inputs into outputs through a stepwise series of actions. Figure 2 shows a schematic representation of such thinking. This description is a model of a generic process. However, any particular individual engaged in work relating to the process will be interpreting the activity involved in her/his own way. It is most unlikely that any particular interpretation will coincide exactly with the formal representation contained in the model shown. “The” model should not be confused with a true representation of “a business process” as practiced.

In their description about their method for Object Oriented Analysis and Design, Mathiassen et al (2000, p. vii) explicitly state that: ‘this is not a book of recipes to be slavishly followed’. Instead, they suggest that they provide ‘a pedagogical presentation of the object-oriented approach’. They go on to suggest that they provide a comprehensive presentation of central principles’... including ‘essential aspects of analysis and design through a systematic presentation of the system’s problem domain, application domain, architecture, and components’. Their argument is that their presentation style is methodological, that their main emphasis is on principles, concepts, and ways of thinking'.

This book is the result of reflection on many years of development and practice to which Object Oriented methods have contributed. Similar points are made by Peter Checkland in relation to the Soft Systems Methology. In his original (1981) work, he suggests that rich pictures may be a useful tool for depiction of a problem situation unstructured. However, he issues a warning that such tools are not to be regarded as representations of reality but as working aids to reflection. In his later work (Checkland and Scholes, 1990), he puts forward a revised version of his original methodology. In doing so, he issues the warning that the earlier version, which had been described as comprising seven stages, has often been treated as a stepwise recipe for analytical work. This he condemned as misguided and unworkable. Further support for this position can be found in the field of management practice. For example, Morgan’s work (1986) sets out a range of metaphors which are in use as imagery for organizational life and work. He highlights a problem
that adherence to one metaphor can lead to the illusion that this is a representation of some reality. He suggests therefore that use of multiple metaphors can be helpful in setting an agenda for reflection, without being seduced into such an illusion.

Further warnings on a similar point come from Claudio Ciborra. He points to a phenomenon where organizational life becomes ‘disambiguated’ by adherence to ‘apparitions’ such as formalised models and methodologies, until ‘work is business process, and people are emotionless decision-makers who have to align their preferences and adjust to the changes rationally planned for them … The intricacies and uncertainties of ambiguity, hospitality, and hostility are ruled out from such a world of abstract organizations, but equally ruled out is the ‘organizingness’ of everyday business life’ (2000, p.31).

In the information systems field we can identify many problem areas of interest, including efforts to analyse both organizational and technical issues. For example, see descriptions of approaches in Avison and Fitzgerald (2002). However, focusing on sociological, economic or technical analyses may not by default lead to understanding or insight into what the concept ‘information system’ is or means. These aspects are used by Ciborra to explain why beliefs related to adoption of ‘neo-positivistic’ methodologies fail in information systems development praxis (Ciborra, 2002). Ciborra chooses to adopt a phenomenologically informed critical perspective. He relates information systems to the challenge of infrastructure incorporating the concept of ‘being in the world’ from the point of view of the individual person. In the flow of life as it is lived, human beings are forced to engage with their context and to develop and apply skills for living. This involves more than application of method, but requires understanding of situatedness of problems and contextualisation of skill and practice. This is the basis for Ciborra’s exploration of bricolage, tinkering and hacking as aspects of any design process (2000; 2002; 2004). Such reflections are not confined to the world of information systems or management alone. Similar thoughts have found expression in, for example, anthropology in the work of Margaret Mead. She suggests that improvisation, collaboration and creativity are to be found in everyday life (not only in the world of artists). This phenomenon can be observed in the contingencies of everyday improvisational interaction, in mundane conversation as well as in formal organisations (Mead, 1928).

Figure 3: “The organization”

Organizations can often be seen represented in a format similar to Figure 3, where distinct functions and lines of authority/responsibility are delineated. However, when contemplating this, we immediately reflect that on the existence of both formal and informal organizational structures. It at once becomes clear that such a standardized model of “The Organization” does not reflect our individual and contextual experiences of organizational practice. It is a metaphor only for an organization as it is experienced.

In organizational theory, it is suggested that individual human beings are capable to organise themselves and to accomplish collectively work of which they would not be capable do individually. When purposeful, complex and collaborative efforts are made to achieve a ‘design’, the nature of those efforts (i.e. method) tend to become visible and emergent only after the fact of production (Weick, 2001, p 58). Furthermore, on occasions, the nature of the efforts which have become successful never emerge to the point where they can be made explicit. See, for example, the account of efforts by Xerox engineers to solve complex faults which were not illuminated by instructions in the official manual. Experienced engineers collaborated in a ‘storytelling’ exercise and together brainstormed solutions which they were never able to fully articulate (Seely Brown and Duguid, 2000).

Lindblom and Cohen (1979) highlight the important role of interaction as a means to achieve problem resolutions. “… in many cases a solution to a problem can be found either analytically or interactively. Settling on one or the other … can also be done either through analysis or interaction, and so on. Settling on
one or the other is often accomplished through habit, tradition, customs, or routines, rather than explicit analysis of the problem of choice. How much thought is required to establish interactive problem solving varies from situation to situation” (Lindblom and Cohen, 1979, p.28). Sandberg and Targama (2007) point to a paradigm shift in management of organizations away from directing and controlling, towards sharing of values, culture and vision. However, they point out that this is often more in rhetoric than practice. Rodgers (2007) points to a paradox in that managers have formal authority to control business decisions but lack any control over the informal interpretations, expectations and competence of their staff. Successful management, they argue, lies in embracing this paradox rather than attempting to resolve it.

This sense of continuous efforts in bridging the gap between human preoccupation with theoretical explanation of experience and experience itself is recognisable in many fields of human endeavour. Malinowski (1922), for instance, describes how ethnography can be used to realise that accuracy is alien to human life, which never complies strictly with any rule. Regularities found by the ethnographer must be integrated with the exceptions that almost always are to be found in social phenomena. He calls this realisation ‘the imponderables of real life’. This can be described as a whole series of phenomena of great importance which consist of practice (actions, behaviours), which people apply in their relationship with their world and each other. We can characterise this as the gap between science and engineering; between music and musical performance; between the rules of a game and playing; between theorising and practice of theory.

Garfinkel (1967), in his discussions related to ethnomethodology, reflected over a constantly changing experience of contextual dependencies using the concept of ‘accountability’. He describes the principle of order on which the social world is based as the product of specific practices that people use to carry out their everyday lives and activities. He suggested that the continuous ‘living in the world’ is done by people effortlessly, and that peoples’ behaviour is always situated (not general). Life happens in a hermeneutic sense in the midst of experiences of contingencies that each situation requires, and to which people continuously and immediately respond. Living is situated and contextually dependent and people are capable of doing this effortlessly, without conscious difficulty. However, contextual analysis requires that efforts are made to surface sense-making processes (e.g. Sandstrom, 1985) The resulting social behaviour consists in many instances of very complex routines, which may become repeated and standardised. Ethnomethodology is then about making sense of, and reflecting about, how individuals in specific contexts and situations behave. It also addresses how people are able to act, and how this can produce practices which others can recognize and consider as appropriate. It is not about seeking abstract and generic principles. Garfinkel proposes intense participation and engagement, not observational distance. The driving insight is acceptance of ‘the unique adequacy of methods’. This calls for recognition of participation in the always-present (and changing) unique situation, context and participants.

All of these actors in diverse spheres of life and activity give us examples suggesting that a Modernist agenda was never as prescriptive as some people have suggested. Modernist methods need not be confused with Modernist practice. Even ‘Grand Theories’ can only be realised contextually. Confusion may arise in part due to an assumption of a Cartesian split inherent in some work based in Logical Empiricism (see discussion in Radnitzky, 1970). It is important to reflect that any observation can only be made by a particular observer; it is also relevant to emphasise the indivisibility of theory and practice in life as it is lived (Bednar and Welch, 2005).

1.1 Hypermodernism and Design Focus

Can we therefore say that Postmodernism has any relevance? If only the uninitiated ever had illusions that the ‘Grand Theories’ of Modernism were intended to be seen as directly applicable to any context, then to those with no illusions, they were a basis for reflection and critique (Table 1). If this is the case, a Postmodernist critique may be perceived as directed at straw men. Insofar as naïve adherence to Modernist theories of method can be seen, we might point to ‘Hypermodernism’ – our recognition that the ‘Grand Theories’ are only intended to be used as a basis for practical philosophy. Also in the ‘modern’ world people such as Gregory Bateson did argue for philosophy as practice when discussing approaches to problem inquiry, critical reflection and systemic thinking (e.g. Bateson, 1972).
Table 1: Relation to “Grand Theories”

<table>
<thead>
<tr>
<th>Relevance of Modernism?</th>
<th>Relevance of Post-modernism?</th>
</tr>
</thead>
<tbody>
<tr>
<td>To those with no illusions, ‘grand theories’ were never more than a basis for reflection and critique.</td>
<td>If, in the field of practice, only the uninitiated ever had illusions that the ‘grand theories’ of ‘modernism’ could be directly applicable, then informed research must recognize this also.</td>
</tr>
</tbody>
</table>

In design of information systems (see Table 2), focus is sometimes on artefact design (with aspirations which might be compared to those of the art deco movement). Design of hardware-intensive data systems would be an example where the focus of design is on the features incorporated into artefacts – what they can do, how they look, etc. Sometimes, focus is on the subtlety of these features, however. Aspirations here might be more comparable to those in the art nouveau movement. The usability of artefacts is the focus, e.g. design of a software interface which supports human interaction with the hardware – not just what can it do and how does it look, but how easy it is to use. Alternatively, a focus might be on the design of total work systems within which the features of the artefact could be relevant. Aspirations could be likened to an avant garde persuasion. Usefulness of artefacts in achieving goals within the life world of the people interacting with them is given attention.

Table 2: Metaphors in design

<table>
<thead>
<tr>
<th>Designers of hardware-intensive systems</th>
<th>Focus on use and artefact design</th>
<th>Aspirations comparable to art deco</th>
</tr>
</thead>
<tbody>
<tr>
<td>Designers of software-intensive systems</td>
<td>Focus on usability and subtlety of features</td>
<td>Aspirations comparable to art nouveau</td>
</tr>
<tr>
<td>Designers of human-intensive systems</td>
<td>Focus on usefulness and design for use in context</td>
<td>Aspirations comparable to the avant garde</td>
</tr>
<tr>
<td>Hermeneutic design (Improvisation)</td>
<td>Recognition that relationship to use can never be known by the designer. The relationship is that of a foreigner</td>
<td>Aspirations comparable to expressionism</td>
</tr>
</tbody>
</table>

However, in the latter case, ownership of the design process by the actors themselves becomes crucial to its successful outcome (e.g. Friis, 1991). A focus on design by a professional designer for notional, representative, or ‘expert’ users could not by default result in useful designs. In other words, any design practice is situational and contextually dependent. These metaphors (Table 2) are helpful in avoiding fights with straw men – other people whose focus is different may not be totally misguided.

2. Conclusion

When an experienced engineer is faced with a proposition that the success of a project can be assured by choosing and applying a method correctly, she might exclaim “Surely, no-one in their right mind really believes that people would actually do it like that!” This is a response appropriate to any prescribed recipe or methodology (e.g. Table 3). It is a familiar form of industrial action to have a ‘work to rule’; slavish adherence to the way things are officially supposed to be done are recognized as fundamentally disruptive to normal practice. Description of practice is thus only a caricature of ‘real’ professional practice, i.e. work carried out in context. Practice comes from a combination of structured and unstructured knowledge, applied by living human beings and therefore going beyond cognition. Ciборра refers to the term ‘Befindlichkeit’ here, to refer to a person’s feeling or being in the world. It is the relationships between different types of knowledge – explicit, tacit and affective - that are important. We have five senses, each of which has been trained and has a relationship to emotion and to memory. Control of some aspects of our behaviour may be unconscious (in the brain but not necessarily in the mind). Other behaviour is controlled through deliberate reflection on possible actions. The spread of different kinds of knowledge, emotion and use of the five senses, and the relationships between them, therefore permeate our experience of living. There is a constant play between senses, emotions and learned responses. We can, at times, choose to apply or not apply some aspects of our ‘knowing’ or feeling. We may choose to make efforts to curb ‘natural’ responses, sometimes successfully. It is also relevant to reflect that, as Maturana says, non-response is also a response (Maturana and Varela, 1980).

If we examine a process of design from conception through to creation, the means by which individuals make things happen become interesting. People may set in motion events or acts which have unintended, as well as intended consequences. These outcomes are a result of combining knowledge, skills, personality traits and attitudes of unique individuals through formal or informal processes or methods carried out systematically to bring about a desired outcome. This combination can be seen as a shadow falling between ideas and plans and their execution. The way in which plans evolve depends upon sense-making processes involving both cognitive and acted-out behaviour (Dervin, 1989). However, plans are carried out through
situated action and are subject to contextual dependencies which mean that individual destinies, though shaped by their sense-making processes and practices, may not be controlled by them. The emergent property from the play between these different aspects is ‘living’ (as opposed to ‘life’).

Table 3: Hyper-modernism

- A recognition that the ‘grand theories’ can only be used as metaphors, i.e. a basis for practical philosophy
- We suggest a need for a critically-informed approach to research which sheds light upon taken-for-granted assumptions and naïve rationalities, illuminating metaphor and stimulating reflection.

It is possible to see Postmodernist discourse as endlessly arguing that a map is not the territory it represents, when this appears to us to have been clear to many leading authorities in the ‘modern’ world already. There are many instances in which models and explanatory frameworks within which research has been positioned are performing a useful role as metaphor, and have not been adopted naively or prescriptively. However, we believe that there is a need for a critically-informed approach to research, i.e. one which specifically attempts to shed light upon taken-for-granted assumptions and naïve rationalities, illuminating metaphor and stimulating reflection. This ‘Hypermodernism’ includes a recognition that ‘Grand Theories’ may be useful as metaphors, i.e. a basis for practical philosophy.

References

Foucault, M. (1973). This is Not a Pipe. J. Harkness (translator). University of California

Elisa Chami-Castaldi, Nina Reynolds and James Wallace  
School of Management, University of Bradford, UK

e.chami-castaldi@bradford.ac.uk  
n.l.reynolds@bradford.ac.uk  
j.wallace1@bradford.ac.uk

Abstract: Response style bias has been shown to seriously contaminate the substantive results drawn from survey data; particularly those conducted using cross-cultural samples. As a consequence, identification of response formats that suffer least from response style bias has been called for. Previous studies show that respondents’ personal characteristics, such as age, education level and culture, are connected with response style manifestation.

Differences in the way respondents interpret and utilise researcher-defined fixed rating-scales (e.g. Likert formats), poses a problem for survey researchers. Techniques that are currently used to remove response bias from survey data are inadequate as they cannot accurately determine the level of contamination present and frequently blur true score variance. Inappropriate rating-scales can impact on the level of response style bias manifested, insofar as they may not represent respondents’ cognitions. Rating-scale lengths that are too long present respondents with some response categories that are not ‘meaningful’, whereas rating-scales that are too short force respondents into compressing their cognitive rating-scales into the number of response categories provided (this can cause ERS contamination – extreme responding). We are therefore not able to guard against two respondents, who share the same cognitive position on a continuum, reporting their stance using different numbers on the rating-scale provided. This is especially problematic where a standard fixed rating-scale is used in cross-cultural surveys.

This paper details the development of the Individualised Rating-Scale Procedure (IRSP), a means of extracting a respondent’s ‘ideal’ rating-scale length, and as such ‘designing out’ response bias, for use as the measurement instrument in a survey. Whilst the fundamental ideas for self-anchoring rating-scales have been posited in the literature, the IRSP was developed using a series of qualitative interviews with participants. Finally, we discuss how the IRSP’s reliability and validity can be quantitatively assessed and compared to typical fixed researcher-defined rating-scales, such as the Likert format.

Keywords: scale length, response styles, response bias, survey research, cross-cultural surveys, individualised rating-scale procedure

1. Introduction

Responses to survey questions include both attitudinal information and response bias. The latter can cause measurement error (Greenleaf 1992a), however, it is the former that is of interest. As such, response bias, that is a tendency to respond systematically to questionnaire items on some basis other than the content of those items, needs to be removed from the data before research results will reflect the construct of interest (Paulhus 1991). Response bias can occur due to the presentation of the construct or measurement instrument; or to respondents trying to portray themselves in a certain way. The former is known as response style, the latter as response set (Rorer 1965). Both types of response bias reduce the validity of research findings (Broughton and Wasel 1990). This paper shows how research instrument design might be used to address response style bias.

Response style bias can result in spurious findings, construct design can affect its manifestation, and it varies along with individual characteristics and across cultures (Berg and Collier 1953; Couch and Keniston 1960; Hamilton 1968; Lorr and Wunderlich 1980; Crandall 1982; Bachman and O’Malley 1984; Cheung and Rensvold 2000). Response style bias is of particular concern in cross-cultural research (Douglas and Craig 1983; Bachman and O’Malley 1984; Ross and Mirowsky 1984; Farh and Dobbins 1991; Chen, Lee et al. 1995; Javeline 1999; Stening and Everett 2001; van Herk, Poortinga et al. 2004).

This paper briefly reviews the types of response style bias; their measurement, current techniques for reducing their contamination of survey data; and their connection to rating-scale length. Subsequently, the paper details the development of the Individualised Rating-Scale Procedure (IRSP), which seeks to address the problem of response styles by means of a dynamic respondent-centred method of attitude measurement. The paper concludes by outlining how the technique’s reliability and validity can be tested.
2. Background

2.1 Response bias, data analysis and rating-scale length

Response styles can be divided into three main groups: those related to the use of particular scale points (e.g. extreme responding and mid-point responding), the spread of responses (e.g. index of dispersion), and the respondent’s reaction to item direction (e.g. acquiescence) (Diamantopoulos, Reynolds et al. 2006). Response style bias can affect both level and structural comparability in cross-cultural research (van de Vijver and Leung 1997), and can result in the appearance of differences between groups when no differences actually exist and/or can hide real differences between groups (Heide and Gronhaug 1992; Baumgartner and Steenkamp 2001). They can increase the association between variables so that significant relationships appear, yet response style bias can also decrease associations resulting in no apparent relationships in the data (Chun, Campbell et al. 1974; Lorr and Wunderlich 1980; Bardo and Yeager 1982b; Heide and Gronhaug 1992). Various techniques can be used to reduce response style bias (e.g., partial correlations, ANCOVA, etc), however, these generally require the researcher to estimate the extent to which response style bias is present.

Three main methods exist to estimate response style bias; using uncorrelated items to estimate response styles (Greenleaf 1992b); collecting both attitudinal and behavioural information as response style bias can then be estimated due to its lesser impact on more concrete (behavioural) information (Greenleaf 1992a); and estimating response styles from existing questionnaire items (Baumgartner and Steenkamp 2001). However, all have flaws: The first requires a large bank of uncorrelated items from which to draw (and assumes the items are uncorrelated across cultures), the last assumes that existing items on a single questionnaire would not share common variation. The problems with the second method of measuring response styles is that it is sometimes difficult to develop behavioural measures that directly relate to attitudinal constructs, and that doing so could greatly increase the length of any questionnaire (Chami-Castaldi, Reynolds et al. 2006).

Researchers aim to choose a scale length such that it is long enough to maximise the amount of information that can be collected, yet short enough to get accurate responses (Cox III 1980). One of the concerns when deciding on rating-scale length is its effect on response style bias. While a longer rating-scale is likely to lower extreme responding (Hui and Triandis 1989), it is also likely to increase scale attenuation (Wyer 1969). As such a balance is needed and six or seven response categories have long been considered appropriate (Miller, 1956). Nevertheless, if an ‘ideal’ rating-scale length could be used (for each respondent), then response styles would be less likely to manifest. Theoretical antecedents of response styles are either dispositional (characteristics of the individual) or situational (context or stimulus related) (Snyder and Ickes 1985; Baumgartner and Steenkamp 2001). The effects of these are considered below.

2.2 Response bias - dispositional and situational effects

Dispositional effects, of personality (Cronbach 1946; Cronbach 1950; Berg and Collier 1953; Lewis and Taylor 1955; Zax, Gardiner et al. 1964; Iwawaki and Zax 1969; Norman 1969; Merrens 1971; Crandall 1982); age (Osgood, Suci et al. 1957); education (Light, Zax et al. 1965); gender (Berg and Collier 1953; Lewis and Taylor 1955); culture (Smith 2004); and occupation and social class (see Hamilton (1968) for a summary of these studies) have been studied in relation to response style. The findings are varied and can appear contradictory. However, what can be said is that there appears to be a link between personal characteristics and response styles, and thus it can be hypothesised that:

H1: A respondent’s ‘optimum’ number of response categories will be related to the respondent’s characteristics.

Situational factors can discourage or encourage response style manifestation (Snyder and Ickes 1985; Baumgartner and Steenkamp 2001). As such the researcher’s measurement choices can directly impact on the data collected from respondents, they should be considered carefully (Rossiter 2002). Factors such as rating-scale length and rating-scale format have been shown to have significant effects on response style bias (Bardo and Yeager 1982b). As such, it is hypothesised that:

H2: Rating-scales with the number of response categories closest to the respondent’s ‘optimum’ number, will be least affected by response style bias.

Traditionally designed measurement instruments usually decide on a standardised rating-scale length for all respondents. Advances in scale development, however, indicate that this may not be necessary.
2.3 Respondent-defined rating-scales

There have been several researchers, in the past, that have argued the benefits of involving the respondent in the generation of more meaningful rating-scales (Kilpatrick and Cantril 1960; Battle, Imber et al. 1966; Donnelly and Carswell 2002; Nugent 2004).

Theoretically, it could be argued that there are three key ways in which a respondent could self-anchor a rating-scale:

- Verbally anchor the scale endpoints,
- Numerically anchor the rating-scale endpoints (i.e. defining the number of response categories they would like to use),
- Conceptually anchor the scale endpoints.

Previous studies have experimented with allowing respondents to conceptually anchor the scale endpoints, in that fixed numerical endpoints are shown to the respondent, before they are asked to anchor the two extreme endpoints with a meaningful scenario (specified by the researcher).

Kilpatrick and Cantril (1960) describe their self-anchoring scale approach as one in which each respondent is asked to describe, in terms of his/her own perceptions, the top and bottom of the dimension on which scale measurement is desired, and then to employ this self-defined continuum as a measuring device. Nugent (2004, p. 171) asked respondents “to imagine a thermometer-type instrument that measures the magnitude of ... depression, with higher scores indicating a greater intensity problem with depression and lower scores indicative of a lower magnitude problem.” Respondents conceptualised their maximum and minimum depression intensities by imagining what, for them, would be the most/least depressing scenario they could picture. Bloom et al. (1999) defined this as an individualised rating scale. Measures from these conceptually-anchored rating-scales were found to be reliable (Battle, Imber et al. 1966; Morrison, Libow et al. 1978) and valid (Battle, Imber et al. 1966; Bond, Bloch et al. 1979; Mintz, Luborsky et al. 1979).

These studies demonstrate that respondents can self-anchor a rating-scale, conceptually. Should respondents also be capable of verbally-anchoring and numerically-anchoring their own rating-scales, this would maximise the meaningfulness of the rating-scale. Allowing each respondent to individualise their own rating-scale should, theoretically, account for both dispositional and situational antecedents of response style, and as such result in measures that more accurately reflect respondents’ actual opinion (Viswanathan, Sudman et al. 2004). As such, it is hypothesised that:

\[ H_3: \text{Measuring constructs using respondent-defined rating-scales will produce more valid measures.} \]

The theoretical issues surrounding the use of individualised rating-scales to minimise measurement error has been considered above. However, before these hypotheses could be tested, it was necessary to determine the feasibility of such a technique and develop a working version.

3. Methodology – Development of the individualised rating-scale procedure (IRSP)

Existing methods of self-anchoring rating-scales do not use the three theoretically possible methods – conceptually, verbally and numerically anchoring – simultaneously. The proposed method, would have respondents independently anchor the numerical as well as verbal endpoints. As such it was necessary to determine the method’s feasibility and develop suitable instructions so that respondents could successfully generate personally meaningful response categories.

When examining the feasibility of individualised rating-scales, the concept of agreement/disagreement was chosen to be the focus of measurement. This concept is frequently used in surveys; when measuring respondents’ opinions towards products/services (i.e. to what extent they agree/disagree with items), when measuring respondents’ psychological characteristics (i.e. Likert-type rating-scales measuring the extent of agreement with a statement representing the ‘self’). If an individualised rating-scale procedure (IRSP) were feasible, it would be necessary to compare its performance against a typically-used measurement tool. Given Likert formats are frequently used to measure agreement/disagreement, it was deemed practical and useful that the substantive focus for the IRSP would be on its ability to measure respondents’ agreement/disagreement with items.
3.1 Phase one – Qualitative exploration & development

To assess the feasibility of the technique and aid its development, fifteen semi-structured in-depth interviews using different verbal and numerical anchoring methods, were conducted. These provided a rich qualitative ground for exploring the order of instructions, phrase structure, semantics and pictorial aids, that yielded the most meaningful rating-scales for respondents.

3.1.1 Main findings

In brief, the results of the interviews indicated that the technique was feasible by:

- Including a visual aid (horizontal line with neutral marked at its centre) to help respondents define their ‘ideal’ rating-scale.
- Giving respondents a conceptual definition of the agreement and disagreement endpoints.
- Having respondents verbally anchor their endpoints, resulting in them having personally-meaningful representations of the extreme positions.
- Having respondents decide on the number of response categories they want to use.
- Having respondents practice using their individualised rating-scale (IRS) by rating a small batch of uncorrelated items (Greenleaf’s (1992b) 16 uncorrelated items were used, given their particular suitability), with the option to revise their numerical endpoints. This task allowed respondents to assess the meaningfulness and distinctness of their response categories before proceeding to the main survey.

3.1.2 The question of comparable data analysis

Prior to this preliminary testing, the issue of comparable data analysis was considered. Given that, even if it were theoretically possible for respondents to numerically (and verbally) anchor their own rating-scales, we considered an important question: ‘How can researchers analyse responses obtained from a group of respondents that each have their own uniquely defined number of response categories?’ This problem was deemed to be solvable if certain controls were implemented:

- That the concept of agreement/disagreement would be measured on a bipolar rating-scale with a neutral point at its centre,
- That the neutral point be numerically anchored at ‘0’, with its conceptual definition defined to the respondent. This is to ensure that each respondent is given the same conceptual definition of the neutral point (i.e. the same point of origin).
- That the endpoints of the scale represent the respondent’s conceptual agreement/disagreement extremes (i.e. that the endpoints to the right/left of neutral represent ‘the most they (the respondent) could possibly agree/disagree with a statement’. This is necessary in order to capture a respondent’s entire continuum of cognitive agreement/disagreement.
- That it is assumed that the distance between adjacent intervals on the rating-scale is equal.

The implementation of these controls would allow scores from rating-scales of different lengths, to be converted into a common scores-index, enabling comparable data analysis. This is analogous to a researcher converting data recorded in different units of measurement into a common unit of measurement, such as imperial units (feet, inches) into metric units (metres, centimetres). Figure 1 illustrates this point.

![Figure 1: Translation of individualised rating-scale ratings into the IRSP score-index](link)
Notice that in this example, Respondent 1 has chosen verbal anchors that, for him/her, represent his/her conceptual extreme points on the agreement/disagreement continuum. Whilst Respondent 2 has chosen different verbal anchors, conceptually, they represent the same extreme positions. This is an important control in order for differing rating-scales to produce data that is comparable. Here, if Respondent 1 rated their particular agreement with a statement (using their IRS) as ‘3’, this score would be converted into the IRSP score-index of ‘1’. Should they report a score of ‘-2’, this would be converted to a ‘-0.67’ in the IRSP score-index.

When considering data comparability in this context, it is useful to reflect upon an observation made by Kilpatrick and Cantril (1960: 4),

“One may question the method by asking, are the data comparable from person to person, or group to group, when one allows each individual to anchor the scale on his own terms? Our position is that they are psychologically directly comparable, that is, the scale level selected by one person or group (average of selections) can be specifically and meaningfully said to be higher, lower, or equal to the scale level of some other individual or group, because the frames of reference of the replies are in fact similar psychologically [...] Meaningful numerical comparisons are then possible.”

Given this method of measurement is dynamic and respondent-specific, it was clear that a computer program was required, to further develop the technique. This would provide respondents with the facility to use interactive visual aids, with their individualised rating-scales being presented to them on screen, to be used to rate survey statements. There was enough data from phase one, to inform the creation of the IRSP software.

### 3.2 Phase two – Creation of the IRSP software & further development

Using bespoke online survey software, phase two aimed to (a) determine whether more meaningful response categories could be developed using IRSPversion1 or IRSPversion2, (b) ensure that respondents found the resultant survey user-friendly, and (c) ensure data capture was accurate. Sixteen in-depth interviews were conducted using verbal protocols followed by retrospective debriefings (Taylor and Dionne 2000). Respondents were asked to use an online survey to create an individualised rating-scale (using either IRSPv1 or IRSPv2) and rate items using their individualised rating-scale (IRS) whilst ‘thinking aloud’ (verbal protocol), and were subsequently interviewed about their experiences with the exercise (retrospective debrief). Given that response styles have been linked with dispositional characteristics, these surveys used several psychological scales, and provided the opportunity to test the suitability of these scales for future IRSP development.

#### 3.2.1 Main findings

Listed are some of the findings that were particularly pertinent or interesting:

- The IRSP survey was viable.
- Respondents carrying out IRSPv2 seemed to produce more personally-meaningful rating-scales, than respondents carrying out IRSPv1.
- E.g. The mysterious attraction to ‘±10’ numerical endpoints (mentioned in Chami-Castaldi, Reynolds et al. 2006), only occurred with respondents that carried out IRSPv1.
- Insights were gained, where respondents chose different verbal anchors for the endpoints of agreement and disagreement.
- Insights were gained, with respondents who chose not to use an equal number of response categories for the positive (agree) and negative (disagree) sides of the neutral point.
- Allowing respondents to practice using their IRS on Greenleaf’s 16 uncorrelated items, proved a valuable part of the process for several reasons;
  - Respondents could ascertain the ease-of-use of their IRS.
  - Respondents could reflect on their responses to the items, and give second thought as to whether their response categories are distinctly meaningful to them, before proceeding to the rest of the survey. This stage allowed respondents to ‘reduce’ or ‘increase’ their rating-scale lengths.
  - The verbal protocols showed that respondents who had a tendency to acquiesce, for example, became conscious of it when reflecting on their use of their IRS on Greenleaf’s items. In some of the retrospective debriefs, these respondents said that after realising
this, they subsequently tried to be more honest and accurate about their opinions when rating items in the main survey.

- Qualitative insights gained after each protocol-debrief interview, frequently led to immediate improvements/modifications for both IRSPv1 and IRSPv2. This meant that IRSP development and data collection was done concurrently.

Due to word limit constraints, only some of these findings will be discussed in more detail.

The quantitative data collected from this small sample of 16 respondents, together with the protocol-debrief qualitative insights, appeared to suggest that the IRSPv2 was more effective than the IRSPv1 (in having respondents produce more meaningful response categories). The mean number of response categories for those that used the IRSPv1 was 11.38, with a standard deviation of 7.05. Whereas the mean for those that used the IRSPv2 was 8.88, with a standard deviation of 5.35. Figure 2 helps to illustrate this point.

![Figure 2: Respondents' chosen no. categories on versions 1 and 2 of the IRSP](image)

On the whole, respondents were able to easily anchor their own verbal endpoints on the agree/disagree continuum in both versions of the IRSP. There were only three respondents who chose peculiar verbal endpoints, shown in the table below (respondents 13, 5 and 11).
Table 1: Verbal endpoints inputted by respondents (phase two)

<table>
<thead>
<tr>
<th>Resp.</th>
<th>Age</th>
<th>Gender</th>
<th>First Language</th>
<th>Ethnicity</th>
<th>National Identity</th>
<th>Verbal Endpoint* (agree)</th>
<th>Verbal Endpoint* (disagree)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>22</td>
<td>Male</td>
<td>French</td>
<td>Other</td>
<td>French</td>
<td>totally</td>
<td>totally</td>
</tr>
<tr>
<td>4</td>
<td>24</td>
<td>Male</td>
<td>English</td>
<td>White</td>
<td>English</td>
<td>definitely</td>
<td>completely</td>
</tr>
<tr>
<td>6</td>
<td>22</td>
<td>Female</td>
<td>English</td>
<td>White</td>
<td>English</td>
<td>absolutely</td>
<td>Absolutely</td>
</tr>
<tr>
<td>8</td>
<td>21</td>
<td>Male</td>
<td>English</td>
<td>White</td>
<td>English</td>
<td>strongly</td>
<td>Strongly</td>
</tr>
<tr>
<td>9</td>
<td>25</td>
<td>Male</td>
<td>English</td>
<td>White</td>
<td>English</td>
<td>absolutely</td>
<td>Absolutely</td>
</tr>
<tr>
<td>13</td>
<td>20</td>
<td>Female</td>
<td>Polish</td>
<td>White</td>
<td>Polish</td>
<td>minimaly</td>
<td>slightly</td>
</tr>
<tr>
<td>15</td>
<td>21</td>
<td>Female</td>
<td>Gujarati</td>
<td>Asian - Indian</td>
<td>British</td>
<td>totally</td>
<td>highly</td>
</tr>
<tr>
<td>16</td>
<td>19</td>
<td>Female</td>
<td>English</td>
<td>Black - African</td>
<td>English</td>
<td>totally</td>
<td>completely</td>
</tr>
</tbody>
</table>

IRSPv2

<table>
<thead>
<tr>
<th>Resp.</th>
<th>Age</th>
<th>Gender</th>
<th>First Language</th>
<th>Ethnicity</th>
<th>National Identity</th>
<th>Verbal Endpoint* (agree)</th>
<th>Verbal Endpoint* (disagree)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>21</td>
<td>Male</td>
<td>Punjabi</td>
<td>Asian - Pakistani</td>
<td>British</td>
<td>totally</td>
<td>fully</td>
</tr>
<tr>
<td>3</td>
<td>24</td>
<td>Male</td>
<td>English</td>
<td>White</td>
<td>British</td>
<td>totally</td>
<td>totally</td>
</tr>
<tr>
<td>5</td>
<td>26</td>
<td>Male</td>
<td>Chinese</td>
<td>Asian - Chinese</td>
<td>British</td>
<td>ok</td>
<td>not ok</td>
</tr>
<tr>
<td>7</td>
<td>18</td>
<td>Female</td>
<td>English</td>
<td>White</td>
<td>English</td>
<td>totally</td>
<td>totally</td>
</tr>
<tr>
<td>10</td>
<td>22</td>
<td>Male</td>
<td>German</td>
<td>White</td>
<td>English</td>
<td>German</td>
<td>totally</td>
</tr>
<tr>
<td>11</td>
<td>22</td>
<td>Male</td>
<td>English</td>
<td>White</td>
<td>English</td>
<td>Agree</td>
<td>Strongly</td>
</tr>
<tr>
<td>12</td>
<td>29</td>
<td>Female</td>
<td>Swedish</td>
<td>White</td>
<td>Swedish</td>
<td>defenently</td>
<td>totally</td>
</tr>
<tr>
<td>14</td>
<td>20</td>
<td>Female</td>
<td>English</td>
<td>White</td>
<td>English</td>
<td>definitely</td>
<td>completely</td>
</tr>
</tbody>
</table>

In their retrospective debriefs, respondents 13 and 5 felt that they had misunderstood the verbal anchoring instructions as a result reading the instructions improperly and possibly the language barrier (given that their first languages are Polish and Chinese, respectively). Respondent 11 apologised for his seemingly unusual “agree” verbal endpoint, explaining that he had rushed through the exercise and had not read through all the instructions. Nonetheless, these three respondents (together with all the other respondents) indicated that when rating statements, they treated the endpoints on each side of the continuum as ‘the most they could possibly agree/disagree with a statement’.

Insights into how some of the respondents conceptually anchored their endpoints before choosing their verbal anchors, proved interesting. Our prompts were worded carefully, so as to encourage respondents to choose personally meaningful verbal labels (i.e. that they would most naturally use). Several respondents talked about the specific scenarios they had pictured when choosing their verbal anchors. For example, Respondent 6 was asked what she was picturing when she chose the word ‘absolutely’, and whether she pictured herself talking to somebody. She said “Talking to my mum! [Laughs].” She confirmed that she pictured herself saying “I absolutely agree” to something her mum might say. She felt she would equally use the word ‘absolutely’ when expressing her extreme disagreement in a conversation with her mother.

Where respondents chose to use different verbal anchors to represent their agreement/disagreement extremes (respondents 4, 15, 16, 2, 12, 14 in Table 1), their reasons for doing so were explored in the retrospective debriefs. It was clear that the verbal anchors they generated were personally meaningful to them, when you consider some of the reasons they gave;
Respondent 16

“*I completely agree*” doesn’t sound like something I would say, whereas, “*I completely disagree*” is something I would say.” She indicated that both her chosen verbal endpoints represented the most she could possibly agree/disagree with something, however, she felt that she would naturally place different adverbs before ‘agree’ and ‘disagree’.

Respondent 4

“When I saw ‘agree’, - I just - ‘definitely’ sprung to mind. I think it’s the way I talk…maybe I associate ‘definitely’ with more positive things. And then, um, when I saw the ‘disagree’ side of things – I just – thought of another word really. I suppose ‘completely’ just sprung to mind, I don’t know if I associate that with being more firm and disagreeing…it was just my opposite.”

It is worth noting that all respondents who chose different verbal endpoints indicated that both of their endpoints (although different) represented, for them, their extremes on the agreement/disagreement cognitive continuum. This echoed some of the findings from phase one of our research study (Chami-Castaíldi, Reynolds et al. 2006).

This is a particularly interesting finding, given that respondents have been shown to interpret standardised verbal anchors in different ways; intensity, quality (Rohrmann 2003). Rohrmann (2003) highlighted the disadvantages of using standardised (i.e. for all respondents) verbal anchors; inferior measurement quality and proneness to cultural biases. He emphasised the need to create rating-scales using verbal anchors which reflect the cognitions of respondents.

In phase one, it was discovered that some respondents desired numerically-imbalanced rating-scales, in that either the positive (agree) or negative (disagree) side of the continuum had more intervals than on the other side. We ensured that the IRSP software, allowed such respondents to individualise their own rating-scales in this manner. See respondents 4, 16 and 14 in Table 2.
Table 2: Numerical endpoints inputted by respondents (phase two)

<table>
<thead>
<tr>
<th>Resp.</th>
<th>IRSPv1</th>
<th>Verbal Endpoint (agree)</th>
<th>Verbal Endpoint (disagree)</th>
<th>Numerical Endpoint (agree)</th>
<th>Numerical Endpoint (disagree)</th>
<th>Modified Numerical Endpoint (agree)</th>
<th>Modified Numerical Endpoint (disagree)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>totally</td>
<td>totally</td>
<td>8</td>
<td>-8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>definitelly</td>
<td>completely</td>
<td>3</td>
<td>-3</td>
<td>3</td>
<td>-2</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>absolutely</td>
<td>Absolutely</td>
<td>2</td>
<td>-2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>strongly</td>
<td>Strongly</td>
<td>2</td>
<td>-2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Fully</td>
<td>Fully</td>
<td>10</td>
<td>-10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>minimally*</td>
<td>slightly*</td>
<td>3</td>
<td>-3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>totally</td>
<td>highly</td>
<td>10</td>
<td>-10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>totally</td>
<td>completely</td>
<td>3</td>
<td>-4</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

IRSPv2

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>totally</td>
<td>fully</td>
<td>3</td>
<td>-3</td>
<td>2</td>
<td>-2</td>
</tr>
<tr>
<td>3</td>
<td>totally</td>
<td>totally</td>
<td>3</td>
<td>-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>ok*</td>
<td>not ok*</td>
<td>6</td>
<td>-6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>totally</td>
<td>totally</td>
<td>5</td>
<td>-5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Fully</td>
<td>fully</td>
<td>3</td>
<td>-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Agree*</td>
<td>Strongly</td>
<td>2</td>
<td>-2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>definitely</td>
<td>totally</td>
<td>4</td>
<td>-4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>definitely</td>
<td>completely</td>
<td>6</td>
<td>-5</td>
<td>5</td>
<td>-4</td>
</tr>
</tbody>
</table>

a This respondent opted for a numerically-imbalanced scale, in that they desired more response intervals for disagreeing then for agreeing.

b This respondent opted for a numerically-imbalanced scale, in that they desired more response intervals for agreeing then for disagreeing.

c This respondent initially defined their IRS as [-3→0→3], which is numerically balanced. He subsequently modified it, after rating Greenleaf’s items, to a numerically-imbalanced scale [-2→0→3].

* This respondent opted for a numerically-imbalanced scale, in that they desired more response intervals for disagreeing then for agreeing.

Interestingly, respondents 14 and 16 defined a numerically-imbalanced IRS right from the start, whereas respondent 4 initially defined his extreme disagree at ‘-3’ and his extreme agree at ‘3’, providing him with a numerically-balanced IRS with seven categories (a typical length with Likert formats). Respondent 4, after completing Greenleaf’s items, realised that whilst he was using all of his varying levels of agreement, he didn’t think any ‘finer’ than two stages when it came to rating his level of disagreement with something. Respondent 14 also chose to modify her IRS (from [-5→0→6] to [-4→0→5]), before proceeding to the main survey. The process of practicing the use of her IRS on Greenleaf’s items, led her to evaluate that whilst she did not need quite so many response categories, she still desired more “shades of grey” when ‘agreeing’ than when ‘disagreeing’. This highlights another potential problem with having respondents use researcher-defined fixed rating-scales to rate statements. It is clear that some respondents gradate their levels of agreement to a greater/lesser extent than with disagreement. This has implications for the instrument validity of researcher-defined fixed rating-scales, and lends further support to the argument for individualised rating-scales.

Some respondents became aware of their tendency to adopt a particular response style, after using their IRS to rate Greenleaf’s 16 uncorrelated items, and being prompted to reflect upon their responses (a pictorial aid with instructions assisted them with this reflection). One particularly interesting case was respondent 4. When prompted to reflect on the meaningfulness of his [-3→0→3] IRS (after using it to rate Greenleaf’s statements), not only did he modify his numerical endpoints (feeling “more comfortable” with a numerically-imbalanced [-2→0→3] IRS), he stated that prior to this point (in his retrospective debrief) that “I didn’t realise I was such an agreeing person.” When asked whether his realisation of this affected the way he completed
the rest of the survey, he responded “I looked at the statements [psychological items] like I did with the first set [referring to Greenleaf’s items] and just thought, “do I completely agree with this?” Um, I didn’t want it to be as neutral as before. I wanted to be a bit more assertive with my answers.”

When comparing his spread of responses on Greenleaf’s uncorrelated items using his [-3…0…3] IRS (Figure 3), to his spread of responses on the subsequent items using his modified [-2…0…3] IRS (Figure 4), it would seem that his tendency to acquiesce was reduced.

3.3 Phase three – IRSP software development & pilot test
A pilot survey with 51 respondents, 24 using IRSPv1 and 27 using IRSPv2, was conducted. This phase was necessary in order to;

- provide a live test for the robustness of the IRSP survey software (i.e. respondents simultaneously completing the survey).
- provide additional data to decide whether IRSPv1 or IRSPv2 would be chosen for the reliability and validity assessment.
- examine how effectively respondents were able to independently carry out the online survey.

Unfortunately, these findings cannot be discussed within the scope of this paper.

4. Future development
In response to Baumgartner and Steenkamp’s (2001) call for forms of measurement that are resistant to response style bias, this paper has considered the impact of rating-scale length on response styles. The literature affirms that there is no single ‘optimum’ rating-scale length for all situations, as it is dependent on the respondent in question (Bonarius 1971; Hui and Triandis 1989; Si and Cullen 1998). This work has established the feasibility of respondents defining their own individualised rating-scales. However, while it may be possible to have individualised rating-scales, the reliability and validity of this measurement method has not yet been established.

Further quantitative research is necessary in order to test the hypotheses presented earlier in this paper (H₁ to H₃). The use of a multi-group experimental design is planned, together with the inclusion of previously...
valued psychological constructs in the online survey. This would help determine test-retest reliability, internal consistency, discriminant, convergent and nomological validity (Bagozzi 1994).

To conclude, if individualised rating-scales can be shown to reduce response style bias, then it would be possible to greatly improve research instrument design. Proving this method is legitimate, would be particularly advantageous for cross-cultural researchers where response style bias is especially problematic.

Acknowledgements

We would like to thank Patrick Oladimeji for building the IRSP Survey Software without which we could not have developed the IRSP technique further. His impressive programming skills and hard work have provided us with the tools necessary to continue developing the IRSP and to test it in a forthcoming online survey.

References


A Framework for Mixed Stakeholders and Mixed Methods

Barbara Crump and Keri Logan
Massey University, Wellington, New Zealand
b.j.crump@massey.ac.nz
k.a.logan@massey.ac.nz

Abstract: Balancing stakeholder expectations and requirements is frequently a challenge for the ethical researcher contracted to evaluate government-funded community projects. Invariably these projects involve people from diverse backgrounds with their own agenda and expectations for the project. This was the scenario for adopting a mixed-method evaluation of Wellington’s Smart Newtown community computing project where free Internet access as well as some computer skills training was made available at the newly-established computing centres. The four-year, multiple stakeholder evaluation project involved qualitative and quantitative approaches, situated within a five-purpose conceptual framework of: triangulation, complementarily, development, initiation, and expansion. The framework provided a robust platform that ensured a systematic and thorough approach in both collection and analysis of data. In this paper we describe the application of each “purpose” of the framework to the different data sets that resulted in an objective, impartial evaluation which was subsequently used for deciding future directions of publicly-funded community computing centres.

Keywords: Mixed method, evaluation, community computing, triangulation

1. Introduction

The acceptance of difference and representation of the diverse perspectives of multiple partners and stakeholders engaged in a research evaluation project can be a difficult task for the researcher contracted to conduct an independent evaluation. Invariably the projects involve people from diverse backgrounds; some will be the target recipients of the completed project, others will represent community interest groups and trusts, local and national governments and corporations. Each group brings its own agenda and expectations as to the outcome of the evaluation and will differ in their abilities to promote their views and being heard.

A strategy to achieve a balance so that a greater diversity of divergent views are heard, questions are answered that other methodologies cannot, and stronger and better inferences are provided (Teddle & Tashakkori, 2003), is to use a mix of qualitative and quantitative methods. Many social science research projects have adopted a mixed method design and in recent years this approach has also gained acceptance in business and information systems research (Bryman & Bell, 2007; Trauth, 2001). However, there is a need for “a clear differentiation of alternative purposes for mixing qualitative and quantitative methods” (Greene, Caracelli, & Graham, 1989, p. 255) and clarity of understanding of why we use mixed methods (Teddle & Tashakkori, 2003).

In this paper we describe how Greene et al’s (1989) five-purpose conceptual framework was applied to the mixed-method design of a longitudinal evaluation of the Smart Newtown Project. This New Zealand digital divide project involved the establishment of five computing centres in Newtown, a low socio-economic suburb in New Zealand’s capital city, Wellington. The centres variously provided computers in homes, free Internet access and free introductory computing lessons. The goal for the two-stage, multiple-partner project was to narrow the digital divide for groups identified by the New Zealand government as likely to be excluded and disconnected from an information society (Servon, 2002). These groups include Maori and Pacific peoples, those on low incomes, sole parents, older people, people with no or low qualifications or poor literacy, and the unemployed (Connecting Communities: a Strategy for Government Support of Community access to Information and Communications Technology, 2002). The Newtown suburb has a high representation of these groups. The focus was therefore on the community sector, identified in the New Zealand’s Digital Strategy (“The Digital Strategy: Creating our Digital Future,” 2005) as one of the three “agents of change” necessary for raising people’s awareness of the potential of information and communications technology. The next section backgrounds the Smart Newtown evaluation.

2. The Smart Newtown study

Smart Newtown’s five community computing centres were established at different locations within Wellington’s Newtown suburb. Free computing facilities and Internet access were provided on a drop-in basis and at two of the centres introductory computing lessons were available. The periods of operation in terms of remaining open to the community varied and at the time of writing just one centre remains open.
The aim for the Smart Newtown Project was to narrow the digital divide and had specific objectives of:

- Improving educational achievement and interest in participation in further education
- Improving short and long term employment prospects
- Developing ICT skill levels among disadvantaged groups
- Enhancing economic and social benefits for the wider community
- Strengthening intra-family relationships and cooperation
- Extending social networks and greater community interaction
- Improving opportunities for residents’ expression of cultural heritage
- Improving flow of information between home and school.

The project was based on a partnership model with stakeholders representing the Wellington City Council, the Wellington Regional Economic Development Agency (WREDA), Wellington Region 2020 Communications Trust, three tertiary education institutions, Fujitsu New Zealand Ltd, community organisations and volunteers and members of the Newtown community. WREDA and Massey university funded the evaluations and Fujitsu financed a computer centre in the Newtown Park Flats (a six-block, council-subsidised, high-rise apartment complex).

A cooperative relationship was envisioned between people or groups who agreed to share responsibility for achieving the project’s goals. While representatives of each stakeholder group were united in a commitment to the success of the project (to a greater and lesser extent), each had his/her own reasons for engaging with the project. As well, they had varying responsibilities and accountabilities and they brought with them their own perspectives and backgrounds that framed their expectations of the project.

For the evaluation two principal manifestations of this partnership model were first, the commercial, contractual aspect where the researchers were ‘employed’ to undertake research on behalf of the funder (WREDA). The second was the development of collaborative, iterative partnership processes between the researchers, individuals and groups within the Newtown community, employees of Fujitsu New Zealand Ltd, a multi-national corporation, and stakeholders who included individual computer users within the Newtown suburb and visitors from outside the community. Figure 1 shows the diversity of the stakeholders.

**Figure 1:** Stakeholders of the Smart Newtown project

Evaluation of the project was in two stages: the 18-month Pilot Project followed by the second 30-month Post Implementation Review. To capture data that would contribute to a meaningful and useful evaluation as well as adopt appropriate data collection methods for eliciting information from such disparate groups a research design that used mixed methods involving qualitative and quantitative approaches was used. Creswell and Clark (2007) recommend the use of a notation system and visual diagrams to communicate the “complexity inherent in mixed methods designs” (p. 40). We have used Ivankova, Creswell and Stick’s (2006) Visual Model for Mixed-Methods Sequential Explanatory Design Procedure to show the phases, procedures and products involved in our mixed-methods design but have included a fourth column that describes the “Sample” (see Figure 2). The centres operated on a drop-in basis and it was therefore impossible to define a “population” from which to draw a sample. Potentially the population was the entire Newtown community.
Barbara Crump and Keri Logan

The Visual model follows Ivankova et al.’s “Ten Rules” (p. 15) for drawing visual diagrams that include limiting the diagram to one page, using capitalized or lowercase letters to designate priority of quantitative and qualitative data collection and analysis, drawing boxes for the collection, analysis and interpretation of results and use of concise language.

Figure 2 shows that over the four-year period two surveys were conducted. First, the User Survey was developed and administered over the early months of the centres’ operations. The low response rate for users of one centre prompted the development and administration of a second survey, that targeted the residents of the high-rise flat (apartment) complex. This we termed the Non-User Survey.

Qualitative approaches included participant and non-participant observation, numerous formal, informal semi-structured interviews, focus groups, and casual “conversations” that occurred during observation at the computing centres. These “conversations” were particularly suited, as a data gathering approach, to the immigrant English language group and the head injury group where formal communication was difficult and structured interviews could be viewed as intimidating. Computing users e-mailed anecdotes of their computing experiences. The variety of data gathering approaches suited a stakeholder sector that would otherwise not have had a voice in the evaluation. At least one of the three researchers attended the monthly communication meetings where documents relevant to the project were tabled and these, together with Memoranda of Understanding between Fujitsu and the Wellington City Council, minutes of meetings, the user registration and attendance log contributed to the data collection.

The components shown in the visual model (Figure 2) are helpful in understanding the mixed methods design for the Smart Newtown Project. The phases and procedures of the design were guided by Greene, Caracelli and Graham’s (1989) five-purpose conceptual framework for mixed method evaluation which are described in the next section.

<table>
<thead>
<tr>
<th>Phase</th>
<th>Procedure</th>
<th>Product</th>
<th>Sample Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quantitative (Users) Data Collection</td>
<td>Paper-based survey ($n = 118$) Centre co-ordinators and researchers administered survey Observation ($n = approx. 32$) Development interview questions 6 Interviews: purposeful selection 23 Conversations</td>
<td>Numeric data</td>
<td>Computer centre users; most, but not exclusively, from Newtown suburb.</td>
</tr>
<tr>
<td>1st Qualitative Data Collection</td>
<td></td>
<td>Text data (interview transcripts and notes, environment description, computer room typology)</td>
<td>Computer users, Centre coordinators, volunteers Fujitsu employees: management and technicians City Council housing representatives and councillors</td>
</tr>
<tr>
<td>Interim Analysis of User Quantitative Data</td>
<td>Low response rate initiated early analysis and decision to develop and administer survey to non-users of Newtown Flats' centre.</td>
<td>32 usable surveys. Descriptive statistics</td>
<td>Majority of computer users from one centre.</td>
</tr>
<tr>
<td>Quantitative (Non-users Data Collection)</td>
<td>Face-to-face survey ($n = 125$ non-users)</td>
<td>Numeric data</td>
<td>Population = 159 residents then living in the Newtown Park Flats where one of the computer centres was located.</td>
</tr>
<tr>
<td>Quantitative Analysis of Data</td>
<td>Excel statistical functions</td>
<td>Descriptive statistics</td>
<td>As above</td>
</tr>
</tbody>
</table>
3. A supportive framework

After analysing 57 mixed method evaluations Greene, Caracelli and Graham (1989) developed a conceptual framework describing how mixed-method evaluation may be used. This framework is based on the similarity of the methods used, the phenomena researched, the paradigms used, the status applied to each method, the timing of the individual methods and whether they are implemented independently or interactively. The five purposes of mixed-method designs are Triangulation, Complementarity, Development, Initiation and Expansion.

Triangulation is achieved by intentionally using more than one method of gathering and analyzing data about the same phenomenon in order to seek convergence and corroboration and to eliminate the inherent biases from using only one method (Denzin 1988; Greene et al. 1989). Such methods should be implemented simultaneously and independently to provide triangulation. Complementarity seeks elaboration, enhancement, illustration and clarification of the results from one method with the results from the other method. Thus quantitative and qualitative results are used to measure overlapping but different phenomena. It is best undertaken when each method is implemented interactively and simultaneously. Development uses the results from one method to help develop or inform the other. Here, each method is implemented sequentially and the results from one method, for example interviews, may be used to develop a questionnaire to assess the same phenomena. Initiation looks for paradox, contradiction and new perspectives in the hope of discovering why such contradictions exist. It requires an iterative approach and Greene et al (1989) suggest that “mixing paradigms in this design is acceptable and even encouraged” (p. 268) because of the ability to maximize the possibility of discovering inconsistencies. They suggest that “purposeful initiation” (p. 268) is probably uncommon. Expansion seeks to extend the breadth and range of inquiry by using different methods for different inquiry components. The researchers’ conclusion that very few of the 57 studies integrated the different method types at the level of data analysis led to a later paper (Caracelli & Greene 1993) that identified four major strategies for analysing, in an integrative fashion, the content of a mixed-method framework.
In the next section we discuss how Greene et al’s (1989) conceptual framework and the “purposes” supported the Smart Newtown research design in the application of the multiple tools of our mixed-method research design.

4. Applying the framework

Prior to development of the research design for the pilot evaluation of the two-stage project a conscious decision was made to use a variety of methods for three clear purposes – that is, Complementarity, Development and Triangulation. It was not until inconsistencies between two different data sets from two of the methods that the purpose of Initiation became a part of the research design and prompted changes to the frequency and timing of the observation method. At the conclusion of the pilot evaluation further funding made possible an evaluation of stage two, the post-implementation phase, and it was then that the fifth design purpose of Expansion produced a more comprehensive evaluation. The application of these purposes is discussed in the following sections and on reflection, now that the project evaluation has finished, we believe that the different characteristics of each purpose contributed to a better understanding and management of the “real-world conditions” (Patton 1990 p. 42) and balancing the politics and tensions that inevitably arose in this multi-stakeholder project.

4.1 Complementarity

The longitudinal evaluation, over a four-year period, allowed the time for illustration, enhancement elaboration, and clarification of the results from one method to the other. Observation proved to be an important method and the focus was broad. We wanted to gain a holistic view of the entire project and there were many frequent, informal visits, at random times to the centres. Depending on what centre and what session was being observed the role of the researcher was as an onlooker, as an outsider (for instance during teaching sessions) or, more usually, as a part-participant (Patton 1990). For example, while observing the computing classes where the immigrant English language group were taught how to connect to the Internet for communicating with friends and family in their homeland we helped when individuals had difficulties with the technology. Observations were overt in that all participants were aware of our role in the project and that observations were being made. The benefits of the observation periods (usually 40 to 70 minutes at a time) included:

- becoming well acquainted with the coordinators and many of the volunteers at the centres as well as some of the regular centre-users
- building trust with the project stakeholders and partners that led to more relaxed interviews and conversations which elicited broader and deeper information.

During the observation sessions informal, “conversation” type interviews occurred and proved to be useful in many situations. For example, they were critical for eliciting the perceptions of participants attending the Head Start programme. These were people who had varying degrees of head injuries and as a result their concentration span was often very limited. Conversations were also useful with the project partners where perceptions and opinions (often not expressed at the formal monthly communication meetings) were expressed.

At the same time formal interviews were undertaken with users, volunteers and partners whereby questions were asked on different aspects of similar phenomena, thus drawing a richer and more in-depth picture of the situation. The formal interviews often complimented the earlier observations and conversations and also enabled us to clarify points raised by participants thus providing ‘the other side of the story’.

4.2 Development

Quantitative data, in the form of a questionnaire, were collected to provide a demographic portrait of participants, to discover what aspects of computing individual users found to be relevant and useful, the frequency and level of use of the centres’ facilities and whether they attended classes or were sufficiently knowledgeable to achieve their objectives independently. The questionnaire was given to users by the volunteers/tutors and was (outside of meetings with those partners involved in developing and supporting the project) the first step in the collection and analysis of data from the intended users. The data gathered from these questionnaires were instrumental in the development of the in-depth interview questions and informed the evaluators’ conversations with participants. Thus, the Development design purpose was congruent with Greene et al’s (1989, p. 267) recommendation for the “sequential timing of the implementation of different methods”.

www.ejbrm.com ISSN 1477-7029
4.3 Triangulation

A major rationale for our use of a mixed-method design was an expectation that there would be convergence in the analysis of the results and a conscious effort was made to assess the same conceptual phenomena. For that reason both formal interviews and informal conversations were undertaken with volunteers, users, and other partners, combined with regular observations at the centres. The different methods were implemented simultaneously and independently conforming with Green et al's (1989) description of triangulation. We combined all four types of Denzin’s (1978) definition; that is data, investigator, theory and methodological triangulation. Data came from a multitude of sources; quantitative and qualitative approaches and the three researchers who brought different perspectives and analytical skills in the analysis and interpretation of the data.

Most of the data collection could be classified as being “within-methods” (Denzin 1978 p. 301) triangulation. That is, many different techniques were used to gather qualitative data which enabled us to be more confident of the internal validity of the results. However, "between-methods" triangulation also occurred whereby data on the same specific questions were gathered from questionnaires and analysed together with the qualitative data generated from interviews on the same topic using an open-ended format (Tashakkori & Teddlie 1998). The various methods produced some convergent results. For instance, findings from the analysis of the interview, conversation and observational data at the centres indicated high usage of the computers and attendance at classes at the Pacific Island Community and Cultural Centre. Here all participants enjoyed learning how to use basic computing functions, the Internet and e-mail. New migrants and settlers in particular appreciated being able to access up to the minute news from their home country and in their own language.

There were also some differences discovered. At the computing centre in the Newtown Park Flats the volunteer coordinator had co-opted another volunteer to help with tutoring the introductory computing classes. These classes were conducted on behalf of a local tertiary institution that received Government funding. The tutors received a small payment for overseeing the self-paced lessons and, where necessary, assisting the learners. Formal interviews with the volunteer tutors indicated that enrolled students, and attendance at classes, were high. During the first six months of classes our observational visits confirmed reasonable attendance. However, in subsequent months classes were frequently cancelled and when they did run there were often no students and on other occasions, just one or two in attendance. Despite this, the register continued to show high enrolments. Informal conversations with the second volunteer tutor with whom one of the evaluators had established a trusting relationship, revealed a different and paradoxical story. He reported that most classes had no students attending, despite many listed on the enrolment register.

4.4 Initiation

As mentioned earlier Initiation was not a planned purpose for the evaluation of the Smart Newtown Project. However, to unravel the contradictions of our data there was a stronger emphasis on the on-going observations, achieved by increasing the frequency of our visits at the scheduled class times. In this way we were able to assess attendance rates rather than rely on reports from the volunteer tutors. Informal conversations with participants at the “drop-in” times (out of scheduled classes) were also useful in our effort to seek and understand the veracity of what was happening. The two methods of “conversations” and observation (and subsequent follow-up with the tertiary provider) clarified the paradox. There were no systems in place that provided checks and balances to ensure accurate and true enrolment status or evaluation of the course. This situation was in direct contrast to the other major centre where the coordinator ran well-attended classes based on a freeware learning programme. The classes at this centre received very positive evaluations and had a transparent enrolment system.

The combination of the different methods revealed a situation that was ethically unsound. It appeared that government-funding, based on fictitious enrolments, was supporting a community computing programme. This knowledge presented a dilemma for the researchers. We had a relationship of trust; we felt morally obligated to WREDA who had contracted the research, and also to the wider community because taxpayers' money was being mis-used. One of the researchers contacted the tertiary provider and suggested evaluations and some observation may reveal irregularities. However these suggestions were received with little interest and after taking advice no further action was taken by us apart from indicating in our final report that our data showed problems with reliable opening and very poor, to no attendance of the classes. The centre closed a few months later due to wider abuse of the facilities (Crump 2006).
The Initiation purpose provided a fuller understanding of the inconsistencies in our original data. We gained new perspectives of the enrolment paradox through the “recasting of questions or results from one method with [the] questions or results from the other method” (Greene et al. 1989, p. 259). It was not a planned “purpose”, which accords with Greene et al’s (1989) comment that “purposeful initiation” (p. 268) is rare.

The Initiation purpose, again unplanned for, was further demonstrated when we broadened our evaluation to include a second survey (Non-user) of the majority of residents who were not using their centre’s free computing facilities. This was an unplanned survey as the feasibility study had indicated high interest in the residents’ intention to use the Newtown Park Flats’ centre yet very few questionnaires were being completed by them. Data resulting from the observation and conversation interviews confirmed many were not taking advantage of the computing facilities within their complex. We therefore responded by developing, testing and administering the “non-user” questionnaire, to elicit from the residents their reasons for not using the facilities. Eighty two percent of respondents to this survey said they did not use the centre and 25 percent indicated they had a computer in their apartment. The main reason for non-use was simply because they were “not interested”. This led to a conclusion that the facility had not been “sold” to residents and that publicity programmes should be activated. It also uncovered the fact that not all people saw engagement with ICTs as a positive force that would transform the quality of their life (Crump & McIlroy 2003).

4.5 Expansion

Because the original research contract was for the 18-month pilot project we did not anticipate the scope and breadth of the research that eventuated when funding enabled evaluation of phase two. Greene et al. (1989) state that the most frequently cited reason for using a mixed-method research model for the 57 studies they reviewed was Expansion. They believe that this indicates researchers are mixing their methods in order to extend the range, depth and breadth of their findings. Their findings suggest that quantitative and qualitative methods are used to assess different components within the same study, thus leading to expansion of the project.

In their discussion of the Expansion purpose Green et al note that frequently qualitative data is used to assess outcomes whereas quantitative data is used to assess implementation. The development of the Non-user questionnaire permitted quantitative data to be used to assess both outcomes and implementation and the phenomena which was being investigated was distinct from other phenomena within the study. The Smart Newtown study incorporated the different approaches used in the qualitative and quantitative methods, particularly at both the discovery and the analysis stage and led to the Initiation purpose described above.

Greene et al (1989) note Expansion as the design option that provides the most flexibility of the five purposes and this proved so for the Smart Newtown Project. We believe the longitudinal nature of the two-stage evaluation and the continued involvement of many of the partners and stakeholders were elements which proved conducive to an Expansion purpose.

Green et al also discuss studies that originally have a quantitative design but due to insufficient or inadequate data are “expanded” to include qualitative approaches, thereby becoming a mixed-method design. This was not the case in the Smart Newtown Project as expansion began early in the Newtown Project when the low usage at the Newtown Park Flats became apparent and while the Non-user Survey was an unplanned component, the mixed-method design had been planned from the beginning for other aspects of the project, thus reflecting a “multitask” intent from the start.

5. Conclusion

The complexity of this two-stage, longitudinal evaluation project involving multiple partners and participants was a major factor in planning the use of a multi-dimensional research design. No single method was sufficient to ensure a systematic and thorough approach to both the collection and analysis of the evaluation data or that reflected the voices of the disparate group of participants.

The Visual Model for Mixed-Methods Sequential Explanatory Design Procedures is based on Creswell and Clark’s (2007) Explanatory Design-participant selection model (p. 85-86). Our study conforms with their choices for this design in that the first quantitative phase is followed by the qualitative phase, the results of which connect to the first phase and the participants were carefully selected so that the qualitative research aim could be best addressed. The scholars note that “investigators typically place greater emphasis on the quantitative methods ….” (p. 72) of these designs but for our study the multiple tools used in the qualitative
approach provided rich content that explained paradoxes and surprising results (Morse, 1991) of the quantitative data. Therefore, priority or weight was given to the qualitative data collection and analysis in the study.

The mixed method approach facilitated responsiveness to unexpected events and results within the project thus enabling evaluation of the project's social and economic objectives. We believe the strength of this evaluation design lies in its flexibility which accommodated the two main strands of the “partnership” model as applied to this project - the commercial and contractual, and the collaborative, iterative approach between the researchers and the various individuals and groups. Greene et al’s (1989) five-purpose conceptual framework guided the design and implementation of the evaluation. However it was not until the dynamism and changes within the project revealed unexpected results that all five purposes became applicable to our study.

We have demonstrated from our discussion that synthesising the results of the data gathered from the different methods led to a greater understanding of the interactions and outcomes of the different computing centres. As noted by Teddlie and Tashakkori (2003, p. 35) “the ultimate advantage of using mixed methods is in the quality of inferences that are made at the end of a series of phases/strands of study”. This proved so for this digital divide study.

References
The Digital Strategy: Creating our Digital Future. (2005) [online], www.digitalstrategy.govt.nz
Active Exploration of Emerging Themes in a Study of Object-Oriented Requirements Engineering: The “Evolutionary Case” Approach

Linda Dawson
Monash University, Caulfield East, Australia
Linda.Dawson@infotech.monash.edu.au

Abstract: The evolutionary case approach provides a framework for qualitative case study research in information systems (IS). It uses revelation, reinforcement, reflection and re-examination to explicitly explore emerging themes in interpretive case study research.

The method is based on the progressive development of a theoretical model grounded initially in the literature and then refined using sequential case studies grounded in practice. The method addresses the gap which often separates data from conclusions in qualitative case study research by documenting the “revealed” and “reinforced” changes in the theoretical model as it evolves from the empirical data.

The paper provides an illustrative study of the use of models in object-oriented requirements engineering to demonstrate the use of the evolutionary case approach.

Keywords: Case study, action research, qualitative, object-oriented, requirements engineering

1. Introduction

The evolutionary case approach provides a framework for qualitative case study research in information systems (IS). This framework was specifically developed to assist exploratory theory-building research using case studies of IS in practice. It uses revelation, reinforcement, reflection and re-examination to explicitly explore emerging themes in interpretive case study research.

The method is based on the development of a theoretical model grounded initially in the literature and then refined using sequential case studies grounded in practice. The method addresses the gap which often separates data from conclusions in qualitative case study research by documenting the “revealed” and “reinforced” changes in a theoretical model as it evolves from the empirical data.

This approach is different to traditional case study approaches which are usually based on replication of situations or units of analysis with the aim of studying each case in isolation and then using within or cross case analysis to discover common themes.

The paper uses an illustrative study of the use of models in object-oriented requirements engineering practice to demonstrate the use of the evolutionary case approach.

2. The evolutionary case approach

The evolutionary case approach is built on the same foundations (Carroll et al., 1998) as the “structured-case” approach (Carroll and Swatman, 2000). It is similar to structured-case in that it is an iterative, theory-building approach based on refining a conceptual framework or theoretical model. It differs from structured-case in that it is designed to explicitly explore emerging concepts in an evolving theoretical model based on reinforcement, revelation, reflection and re-examination, in order to understand the relationship between theory and practice. It also differs from structured-case in that there is no “… literature-based scrutiny of the research findings (p236)” based on Eisenhardt’s (1989) “enfolding of the literature” since this can be considered to be theory-testing rather than theory-building.

Case study research aims to examine a phenomenon in its natural context (Yin, 1994, Cavaye, 1996) and aims to contribute to knowledge by relating findings to generalisable theory (Cavaye, 1996). The case study research method is one of the most popular methods in information systems research and is well-suited to understanding the interactions between information technology-related innovations and organisational contexts (Darke et al., 1998). Qualitative interpretive case study research involves more than an examination of qualitative data. It involves looking for a deep understanding of why behaviours occur and why people do what they do. It also attempts to understand the opinions, beliefs, experiences, expectations, etc which influence the behaviour of participants through interviews (Myers and Newman, 2007), participation and/or
observation (Walsham, 1995). Researchers need to “let the data tell the story” but when analysing qualitative case study data it is often difficult to demonstrate the links between the data and conclusions.

The action research method involves collaboration between the researcher(s) and the participant(s) and intervention by the researcher(s) in the situation being studied (Susman, 1983, Myers, 1997, Baskerville and Wood-Harper, 1998, Avison et al., 1999). This often involves a cycle of feedback aimed at increasing the understanding of a given social situation (Hult and Lennung, 1980). Galliers (1992) describes action research as an interpretivist subset of the case study approach where the presence of the researcher affects the situation being studied.

The grounded theory approach seeks to develop theory that is grounded in data that has been systematically gathered and analysed (Strauss and Corbin, 1990). A key element of the grounded theory approach is that the theory should be developed with a reflexive 'back and forth' interplay between data collection and analysis (Myers, 1997, Urquhart, 1998) without any preconceptions about the outcome until all the data has been collected and analysed.

The evolutionary case approach incorporates elements of case study research in that it relies on interviews in the field, it is non-interventionary (and non-participatory), and it uses multiple sequential cases. It incorporates elements of action research in that it is iterative (each case has its own cycle), and it relies on learning and reflection on accumulated findings, both within individual case cycles and between cycles of multiple sequential-cases, for refining interview scripts and the theoretical model. And it incorporates elements of grounded theory in that the data is analysed as it is collected and revisited before final conclusions are drawn.

![Figure 1: The evolutionary case cycle](image)

Central to the evolutionary case approach is the concept of a theoretical model initially grounded in the literature and describing a common understanding of “what might be happening” in practice, which is progressively refined based on empirical data grounded in practice. The term “theoretical model” is used here to mean a representation of a theory where the definition of a theory is based on “A conception or mental scheme of something to be done, or of the method of doing it; a systematic statement of rules or principles to be followed.” (OED, 2004) or “conjectures, models, frameworks, or bodies of knowledge” (Gregor, 2006) and contains features of a theoretical model as defined by Dubin (1976) and Bacharach (1988) i.e. the interactions or relations between defined units or concepts within a set of boundaries or constraints depicting a limited portion of the world.
The evolutionary case approach is particularly suited to interview-based data collection and the subsequent discussion in this paper is based on this. In each case cycle the researcher seeks to refine the current version of the theoretical model by:

- looking for reinforcement of concepts already contained within the theoretical model
- revelation – identifying new areas for exploration and potential reinforcement
- learning and reflection on data collected so far
- re-examining previous transcripts to find any further reinforcement of an emerging theme

The researcher is active in the data collection. Leading questions are encouraged in order to facilitate reinforcement and semi-structured, open-ended questions are used to facilitate revelation. Exploration of these revelations is incorporated into revised interview scripts which are used in the next cycle. Reinforced concepts are retained in the evolving theoretical model. The process is ongoing but concluded when there has been enough reinforcement for a representative model of the research domain being investigated to stand alone or when theoretical saturation has been reached (Eisenhardt, 1989). Therefore, the outcome of the research method is a revised theoretical model (with several revisions during the process) which represents an theory which has evolved about the area being investigated.

2.1 The research cycle

As shown in Figure 1, the formulation of an initial theoretical model and research questions are based on the definitions and common understanding established from the relevant literature. The theoretical model together with the research questions is used to set up, plan and initiate the subsequent research cycle. This approach is adapted and extended from Miles and Hubermann (1994) who call this “focussing the collection of data”, p 16-25.

After the evaluation of findings at any iteration, the current accumulated findings and learning are reflected upon. This reflection activity provides two things: possible further refinements of the theoretical model and possible refinements to the interview scripts in order to explore any emerging categories.

As the number of cases increases and the accumulated data increases, an important part of the reflection process involves the re-examination of previous cases. This re-examination allows the potential reinforcement of emerging categories from the previous data.

The next section provides an illustrative example of the use of the evolutionary case approach.

3. Example: the use of object-oriented models in requirements engineering practice

The move towards the use of object-oriented methods for information system development led to a need for the development of object-oriented approaches to requirements engineering. However, little is understood, or reported on the basis of research, of the use of object-oriented methods by practising professionals in the production of requirements specifications for commercial or industrial sized projects (Burton-Jones and Meso, 2006, Dobing and Parsons, 2008). An understanding of what successful professional developers do, how they do it and why they do what they do needs to be investigated (Burton-Jones and Meso, 2006, Dobing and Parsons, 2006, Grant and Reza, 2007).

3.1 Themes from the literature review

This project brings together concepts from requirements engineering and object-oriented methods. Preliminary findings have been reported in Dawson and Swatman (1999) and Dawson and Darke (2002).

3.1.1 Requirements engineering

Requirements engineering refers to the early stage of the systems development process which manages the identification and documentation of system requirements (Loucopoulos and Karakostas, 1995, Macaulay, 1996, Sommerville and Sawyer, 1997, Hull et al., 2005). Various frameworks and models of the requirements engineering process have been suggested in the literature. Highly cited frameworks include Pohl (Pohl, 1994), Loucopoulos and Karakostas (1995) and Macaulay (1996). For this project a review of the literature resulted in the selection of Loucopoulos and Karakostas (1995) framework from which to develop a theoretical model for object-oriented requirements engineering. In this framework the requirements
The engineering process can be broken down into three sub-processes, elicitation, specification and validation, which deal with two external entities, the user and the problem domain. See Figure 2.

![Figure 2: A framework for requirements engineering processes: Loucopoulos and Karakostas (1995) page 21](image)

### 3.1.2 Object-oriented methods

Object-oriented methods model systems as a set of communicating “objects”. Identifying and designing objects is part of the object-oriented systems development process. Object-oriented models include characteristics and attributes of objects modelled as data structures (static properties) and behaviour modelled as methods or operations (dynamic properties), together with concepts of generalization, inheritance and aggregation. For example, a banking system might include “customer” objects and “account” objects and account objects may have various operations including “withdraw”, “deposit” and “check balance”, etc.

Many claims have been made about the object-oriented paradigm (Coad and Yourdon, 1991, Budd, 1997, Jacobson et al., 1999). These claims include:

- ease of understanding object-oriented models due to a consistent underlying representation throughout the development process
- the ability to model the behaviour of objects (encapsulation of data and process)
- ease of modification and extensibility of object-oriented models.
- ease of reuse of object components from previously designed systems
- superior data abstraction facilities including inheritance and polymorphism

When studying object-oriented methods in requirements engineering, the interest is in how experienced analysts and developers actually use such methods in “real-world” system specification. Researchers need to consider whether the benefits of reuse, abstraction and reduction in complexity outweigh any difficulties in using to object-oriented methods.

### 3.2 An initial theoretical model

The initial theoretical model (Figure 3) is based on Loucopoulos and Karakostas (1995) and common characteristics of published, well known object-oriented methods (Coad and Yourdon, 1991, Rumbaugh et al., 1991, Jacobson et al., 1992, Graham, 1994, Henderson-Sellers, 1997, Avison and Fitzgerald, 2006, Booch et al., 2007) particularly the concept of separate static models (which describe objects, their characteristics and the relationships between them) and dynamic models (which define states of objects, state transitions, message passing and event handling). This model represents the three processes of the requirements engineering process: elicitation, object-oriented modelling and validation which interact with users and a specific problem domain. The elicitation process clarifies user requirements within some problem domain. These requirements are then modelled within that domain using various techniques and these models are then validated against the original requirements within the problem domain. The initial model explicitly shows the static and dynamic models that are produced during object-oriented modelling.
3.3 Research questions

Research questions based the initial theoretical model were:

- Are there three identifiable processes of elicitation, modelling and validation in object-oriented requirements engineering practice?
- Is elicitation an iterative process in object-oriented requirements engineering practice?
- Which static and dynamic models are developed in object-oriented requirements engineering practice and how they are used and by whom?

3.4 The participants

Participants for this study were recruited through industry. All the participants were currently working in the field of object-oriented requirements specification. See Table 1.

Table 1: Background information for each consultant

<table>
<thead>
<tr>
<th>Case</th>
<th>Title</th>
<th>Years in RE</th>
<th>Client</th>
<th>Project</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Operations manager</td>
<td>3.5 yrs</td>
<td>Federal Govt</td>
<td>Complex Technical</td>
</tr>
<tr>
<td>2</td>
<td>Principal Consultant</td>
<td>15 yrs</td>
<td>State Govt</td>
<td>Web based transactions</td>
</tr>
<tr>
<td>3</td>
<td>Senior Consultant</td>
<td>12 yrs</td>
<td>Telecommunications</td>
<td>Fault Mgt System</td>
</tr>
<tr>
<td>4</td>
<td>Director &amp; partner</td>
<td>22 yrs</td>
<td>Software developer</td>
<td>Insurance</td>
</tr>
<tr>
<td>5</td>
<td>Consultant</td>
<td>14 yrs</td>
<td>Software developer</td>
<td>Life Insurance</td>
</tr>
<tr>
<td>6</td>
<td>Technical Manager</td>
<td>12 yrs</td>
<td>Software developer</td>
<td>Stockbroking</td>
</tr>
</tbody>
</table>

4. Findings

Evaluating and analysing rich qualitative data in a rigorous manner is difficult and "…no single qualitative data analysis approach is widely accepted" (Neuman, 1994). The approach in this project was to use an evolving set of categories to structure the qualitative data as it was gathered. Firstly, a set of seed categories (Miles and Huberman, 1984, Fitzgerald, 1997, Wynekoop and Russo, 1997) was formulated based on the initial theoretical model and these were used to formulate the initial structured interview script. In each interview other categories and sub categories emerged and were incorporated into interview scripts for
investigation in the following cases. So the number of categories grew as the case studies continued. A detailed representation of the evolution of a theoretical model is presented in tables showing reinforced and revealed themes and documented revisions of the theoretical model.

The presentation of case data is based on illustrated narrative style, or an oral narrative told in the first person, as described by Miles and Huberman (1994) and Myers (1997) and as used in Fitzgerald (1997) and Urquhart (1998). This approach as described by (Miles and Huberman, 1994) does not resort to explicit coding but looks for “... key words, themes, and sequences to find the most characteristic accounts.”

4.1 The reinforcement of the model

4.1.1 Three processes

One of the major concepts embodied in the theoretical model is that there are three identifiable processes in the object-oriented requirements engineering activity. Those processes are elicitation, object-oriented modelling and validation of the models. The specific questions addressing this concept taken from the interview scripts are:

- Is elicitation explicitly undertaken and when does it start? When does it end?
- When does modelling begin? That is, when do you start drawing object models?
- Do you think it is necessary to validate the specification once the models have been produced?
- When does the validation process start?

The three processes of elicitation, modelling and validation were identified in all six cases. In Case 1 there were three stages within the elicitation process. Elicitation started with a "blast-off" meeting and continued with regular interviews with the users and meetings of the project team. There was little traditional modelling although requirements cards were used to represent requirements and their characteristics. Validation was based on walkthrough techniques.

In Case 2 elicitation was initiated by interviews with upper management to "scope the project". Existing documentation and data models were used as a starting point for interviews and setting up use case scenarios. Modelling was based on standard object models and use cases. Validation was based on walkthrough techniques.

In Case 3 elicitation was done by interview and prototyping with a small group of selected users and a specific "subject matter expert". Modelling was based on object/class models and comprehensive textual use cases. Modelling and prototyping were supplemented with a case tool called Software Through Pictures (STP). Validation was based on walkthroughs and revisiting the use cases and the prototype with the user group and subject matter expert.

In Case 4 elicitation was done explicitly where possible but because this project was a generic package the users were not captive to the organisation. For the purposes of elicitation, in-house business analysts and pre-sales people played the role of users in a client organisation. Both entity-relationship modelling and UML models were used to produce requirements models. The consultant also used some use cases together with ad hoc diagrams and rich pictures as models with the users. Validation was done as walkthroughs with the prototype and role-playing based on use cases.

In Case 5 elicitation was done by interviews with an emphasis on building trust and rapport. There was no use of specific modelling techniques or notation. The specification was based on a document which "embodied concepts that could be directly implemented by experienced system developers". This appeared to be based on a "programmer-oriented" approach to system development. Specific validation was not seen as necessary if the team has "built the right thing" although some acceptance testing was seen as useful.

In Case 6 elicitation was explicitly undertaken using interviews and involved gap analysis and the production of a business requirements document. Modelling was based on entity-relationship models and UML notation though not a lot of models were produced. Use cases were not used extensively because there are not a lot of tools to support them. Use cases were seen as useful in the validation process where the team can "verbalise or walk through a scenario". Prototype demonstrations were also considered useful for validation.
The analysis of findings showed that although the methods used in each process in each case were different, the three processes could be identified in all cases and so the existence of three processes was empirically validated.

4.1.2 The explicit use of feedback in elicitation

The explicit use of feedback in the elicitation process of object-oriented requirements engineering as depicted in the initial theoretical model as a double-headed arrow was actively explored with the following question taken from the interview scripts:

- Is knowledge elicitation iterative? That is, do you go back to the users several times?

In Case 1 there was feedback and elicitation was iterative. On-site meetings were held every second day with the clients and these meetings were interspersed with in-house project team meetings.

In Case 2 elicitation was an iterative process. The analyst would go back to the client, on average, three times so that the whole transaction specification took about a week. Information gathering usually took place on an initial half-day with some follow-up over the next two days. A week after the start the specification was given back to the client for review and the analyst walked them through it.

In Case 3 elicitation was considered to be "highly iterative" and included contact with the subject matter expert every couple of days.

The generic nature of the package being developed in Case 4 meant that there were no actual users available and elicitation was done using role-playing. It is not clear whether the role-playing involved iteration.

In Case 5 knowledge elicitation was seen as iterative and based on feedback. The team went back to the users several times where some piece of information triggered the need to explore some new feature or aspect.

In Case 6 users were interviewed several times to clarify points and produce the business requirements document. Different groups of users may be involved in each iteration including large groups and subsets of key users.

The analysis of findings showed that in all but one case the analyst saw the elicitation process as iterative and based on feedback for acquiring and clarifying requirements. The feedback loop in elicitation as shown in the theoretical model was empirically validated.

4.1.3 Identification and use of static and dynamic models

A major concept in the literature regarding object-oriented modelling is that there is a need for both static and dynamic models for the representation of object-oriented concepts. The nature of, and types of, static and dynamic models produced in object-oriented requirements engineering were actively explored using the following questions taken from the interview scripts:

- Which models are produced during specification?
- Do you produce class models, use case models or interaction models?
- Would you categorise models as static or dynamic?

Case 1 did not make extensive use of object-oriented models although these models were available as part of the in-house methodology. The main vehicle for representing requirements was requirements cards and the modelling that was done was based on flow charts. In Case 2 both static object/class models and dynamic interaction models were produced although the dynamic models were put into the appendices of the specification document and were rarely used with the users or for requirements specification.

In Case 3 static object/class models and dynamic interaction models were developed but the interaction models were used later in the system development process rather than in requirements specification. Case 4 included entity-relationship models and static object/class models. The dynamic models included state-transition models and interaction models.

In Case 5 models were viewed as "a set of concepts" and these models were classed as dynamic by the analyst. In Case 6 several types of static and dynamic models were used including entity-relationship
models, object models, state transition models and interaction models "... the static model, their object model, class diagram, however you want to describe it, is the core. Everything starts there. ... And then from that springs forth collaboration diagrams and interaction diagrams."

Table 2: Reinforcement of initial concepts in the theoretical model

<table>
<thead>
<tr>
<th>Three processes</th>
<th>Feedback in the elicitation process</th>
<th>Explicit static models</th>
<th>Explicit dynamic models</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case 1</td>
<td>reinforced</td>
<td>reinforced</td>
<td>reinforced</td>
</tr>
<tr>
<td>Case 2</td>
<td>reinforced</td>
<td>reinforced</td>
<td>reinforced</td>
</tr>
<tr>
<td>Case 3</td>
<td>reinforced</td>
<td>reinforced</td>
<td>reinforced</td>
</tr>
<tr>
<td>Case 4</td>
<td>reinforced</td>
<td>reinforced</td>
<td>reinforced</td>
</tr>
<tr>
<td>Case 5</td>
<td>reinforced</td>
<td>reinforced</td>
<td>reinforced</td>
</tr>
<tr>
<td>Case 6</td>
<td>reinforced</td>
<td>reinforced</td>
<td>reinforced</td>
</tr>
</tbody>
</table>

The analysis of findings showed that in four of the six cases models were produced and that they were specifically defined as static and dynamic models. Overall, the characteristics of the initial theoretical model were empirically validated by the results contained in the data from the six cases (see Table 2). There was no evidence for removing any of the major concepts from the initial theoretical model. The main characteristics analysed in this section were reinforced in five of the six cases where the five cases were different for each concept.

4.2 The evolution of the model

The following section outlines the findings of each of the six case studies with reference to evolution of the theoretical model. The major emergent themes reinforced by subsequent cases were:

- Evidence of the use of use case models as distinct models for requirements representation
- Evidence of mental modelling by analysts during elicitation before any models were committed to paper
- Evidence of the use of separate formal and informal models where informal models were the only models shown to users by analysts when discussing requirements.

4.2.1 Theme 1: Use cases

Jacobson (1992) defined a use case as:

“When a user uses the system, she or he will perform a behaviourally related set of transactions in a dialogue with the system. We call such a special sequence a use case. p 129” Jacobson, 1992

Simply, a use case can be thought of as a description of a transaction within the system. It can take a graphical or textual form. The use of use case models (both textual and graphic) as requirements models distinct from the static and dynamic models emerged as a category worth exploring in Case 2.

Although use cases or scenarios are described as ways of modelling requirements and transactions for object-oriented systems in several methodologies, they are difficult to classify as specifically static or dynamic models. Use cases can come in textual forms, graphical forms or both. Use case models could be classified as static models because of their textual characteristics or they could be classified as dynamic models based on the graphical, process-oriented representation. As a result use case models were not explicitly shown in the initial theoretical model. The use of use case models (both textual and graphic) as requirements models distinct from the static and dynamic models emerged as a category worth exploring in Case 2.

After Case 2 the questions incorporated in subsequent interview scripts exploring the use of use case models were:

- Do you use use case models at this stage [elicitation]? What form do they take, textual or graphical?
- Could you comment on the role and/or importance of use cases in the specification process?
- Could you comment on the relationship between use cases and static and dynamic models?

On reflection and re-examination Case 1 used a methodology that potentially included use cases but the analyst did not use them in the project under study. In Case 2 the use of simplified use cases as models in requirements specification emerged as a significant technique. “The generic object model is a standard
object model and the rest of the methodology is built around Jacobson use cases. The focus is on use cases. There is a standard process modelled as a use case flow diagram and a use case [structured dialog] and the client has to tick the box if there is a good fit.”

Subsequently, Case 3 and to a lesser extent Cases 4 and 6 gave evidence of the use of use cases for requirements modelling. “We saw the need for a static type of model and a dynamic type of model and use cases… the use case model was more stand alone and really became the functional statement that went behind or reinforced the UR [user requirements] prototype … really most of the first phase of the work developed our use case model and a prototype to go with it.” [Case 3]

“Well, that’s … according to the theory they [use cases] are the backbone – you start with them and they go right through to the testing and so forth and I think that’s reasonable. Generally speaking I think it’s a good way to use them.” [Case 4]

“We do [use use cases] a little … when we have got a prototype and some rules and requirements [and] the users now start saying well, ‘what if?’ … We verbalise or walk through a scenario and then say well how does this system handle this.” [Case 6]

Use cases were used to model requirements in four of the six cases and were explicitly added to the theoretical model – version 2 (see Table 3 and Figure 4).

4.2.2 Theme 2: Mental modelling

The concept of mental modelling during elicitation by an analyst emerged in Case 3 as the following selection of quotes from the transcript illustrates: “…there were fragments of that [business] model getting developed in a couple of people’s heads for probably three months…I would say that it was being done largely privately and it was not written down until the last minute when it was just a dump”.

“You will be listening very carefully [in project meetings] and collecting and cataloguing constraints and refining the abstractions in your mind.”

This concept was explored in subsequent case study interviews with the following question:

- Do you start developing mental models during elicitation?

Overall four of the six analysts believed that they were continually “modelling in the mind” during the elicitation process and that these mental models were further refined in the mind before they were communicated to others (users or fellow analysis team members) or before they were committed to paper.

In Case 4 creating mental models involving key objects was seen as a natural part of requirements elicitation and modelling “I think that I do immediately start thinking of key objects during requirements gathering, not in any formal way, they just pop into one’s head. I don’t agree with the implication … that identifying objects and ‘building mental models of the system’ are mutually exclusive. One can help the other.”

In Case 5 mental modelling was perceived as an integral part of abstraction "This is really about how people think. Some of us more conscious of the models, others not."

In Case 6 producing mental models during elicitation was seen as a natural way of thinking for that particular analyst although he thought it was a personal thing and that other analysts might not work that way “I do that. … That’s not true of some of the others. You can tell when you look at their work that they’ve not actually thought about any form of underlying structure at all. All they’ve really done is try to gather the business requirements.”

Mental modelling emerged as significant in Case 3 and was subsequently reinforced in Cases 4, 5 and 6 and was explicitly added to the theoretical model – version 3 (see Table 3 and Figure 4).

4.2.3 Theme 3: “Formal” and “informal” models

As described above, when asked when the modelling process began, most analysts said they were building models in their heads long before any formal models were written down.

Evidence of the use of “informal” models (simple use case models, ad hoc pictures, diagrams, animations etc) instead of “formal” notated models (object/class diagrams) for communicating the specification to clients and users emerged in Case 4. Subsequent reflection and re-examination of the transcripts of Cases 1, 2 and 3 revealed further evidence of this “separation of models”. For the purpose of discussion formal models and informal models are distinguished in the following way. Formal models are considered to be those models.
that require training in order to be understood or explained. That is, models that contain specific, often graphical notations such as UML models, interaction models or state models. Informal models are considered to be models that can be understood and explained without specific training. In this category are natural language models including text descriptions, use case scripts, ad hoc diagrams and interactive demonstration models as often produced for prototypes.

Questions regarding types of models which were part of the original interview script were:
- Which (how many) models are produced during specification?
- Who uses them? Who are they produced for?
- Which models, if any, are shown to the user? Which models are used internally by the development team?

Questions regarding types of models asked in interviews subsequent to Case 4 were:
- Which (how many) models are produced during specification? Do you produce class models, use case models or interaction models?
- Do you use informal models (pictures etc) to communicate with the users?
- Do you use use case models at this stage? What form do they take?

In the case studies there seemed to be two different kinds of modelling taking place. Firstly, there were the informal models that were used to communicate with the users. Secondly, more formal models were developed which were not shown to the users because it was believed that the users would not understand them. The formal models were developed primarily for design purposes and were private to the analyst or team of analysts. In effect these formal models were the analysts’ internal version of informal models.

In Case 1 requirements cards were used directly with the users to represent or “model” requirements and it was not made clear which models, if any were used in the design and implementation process “The card is pretty much self documenting … straight into the actual requirement spec. So once you have the cards complete, a lot of the hard work is done … we cannot write the requirements, they must tell us the requirements … we work with them on the cards.”

In Case 2 only simplified use case diagram and dialogues were shown to the users when describing requirements. Models based on formal notation were considered too complex for users to understand “We tell them [the users] that the model is technical mumbo jumbo … you know I wouldn’t show them a data model either … the closest I’ve gotten is working with this type of flow diagram (use case flow diagram) … they can follow that pretty well but they don’t usually have the patience to really work through the interaction diagrams or the model. It just takes too much explanation.”

In Case 3 use case scripts and a prototype were used to develop the requirements with users and formal object models were not shown to users “… unless the 'users' were IT-literate people, which most aren’t. (Do you believe it is not necessary to show them?) I believe it is not only not necessary, but potentially dangerous. It is the analyst’s job to perform the use case to business object model translation.”

In Case 4 the analyst was explicit about using various ad hoc diagrams, pictures, PowerPoint simulations and some use cases to communicate the requirements to users “I mean if you draw a picture and that doesn’t make any sense to them then you draw another one … A requirements specification has to be in terms that they understand and those three mechanisms we’ve already mentioned are the way: the use case, the ad hoc diagrams and the dynamic screen simulations … I am not saying one should do away with the formalisms. They are a powerful aid to one’s own understanding and analysis but they are not a good tool for feeding requirements back to the users. It's much better to bend the formalism to the user than the user to the formalism.”
Table 3: The evolution of the theoretical model

<table>
<thead>
<tr>
<th>Case 1</th>
<th>use case models</th>
<th>mental models</th>
<th>formal and informal models</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case 2</td>
<td>revealed</td>
<td></td>
<td>informal models for users reinforced by reflection and re-examination</td>
</tr>
<tr>
<td>Case 3</td>
<td>reinforced</td>
<td>revealed</td>
<td>reinforced by reflection and re-examination</td>
</tr>
<tr>
<td>Case 4</td>
<td>Weakly reinforced</td>
<td>reinforced</td>
<td>revealed</td>
</tr>
<tr>
<td>Case 5</td>
<td>reinforced</td>
<td>reinforced</td>
<td>weakly reinforced</td>
</tr>
<tr>
<td>Case 6</td>
<td>Weakly reinforced</td>
<td>reinforced</td>
<td>reinforced</td>
</tr>
</tbody>
</table>

The consultant in Case 5 did not use models or explicit diagrams. Most of the specification was based on text-only documents including a formal specification document.

In Case 6 the most important objective was to get the users or clients to sign off on the specification “…we would have a formal walkthrough with the users where we go through the requirements, … and probably the prototype and we would get formal sign off”. In this case the prototype was the most used tool for validation and use cases were only used for exceptions or special cases “…as we were looking at the requirements document we had the prototype running and projected up on a big screen and we actually walked through the prototype in relation to the requirements”.

Table 3 shows the evolution of the theoretical model in terms of revelation, reinforcement reflection and re-examination. Figure 4 shows version 4 of the theoretical model incorporating, use case models (version 2), mental modelling (version 3) and the separation of formal and informal models in the object-oriented requirements engineering process. The feedback from the validation process for formal modelling in the problem domain is via the ad hoc mapping of informal models to formal models for use in design and implementation.

Figure 4: Final theoretical model

The final theoretical model (Figure 4) now embodies all the concepts of the initial theoretical model which were proposed at the start of the evolutionary cycle based on the literature regarding object-oriented requirements engineering, together with the concepts that emerged from the sequential case studies. This revised theoretical model provides a theoretical representation of the object-oriented requirements engineering process grounded both in the literature and in professional requirements engineering practice.
5. Conclusion

The evolutionary case approach provides a framework for qualitative case study research which actively explores emerging themes. Emerging themes that are contained in the current version of the theoretical model are actively explored in interviews and themes that are reinforced by empirical data become part of the next version of the theoretical model. An ongoing process of reflection and re-examination further refines the data and the model.

This framework was specifically developed to assist exploratory theory-building research using case studies of IS in practice. The theoretical model and the associated tables documenting the emergence and reinforcement of themes provides structure and support for theory building within the specific constraints of interpretive, qualitative, interview-based data collection.

This structure and support for the research process makes contributions in two ways. It provides productivity gains by providing a specific framework and process based on the development of a theoretical model for representing the current state of the area under study and it contributes to quality outcomes by providing clear documentation of the process and results.

The illustrative example of the use of models in object-oriented requirements engineering demonstrates that the evolutionary case approach provides a research environment for truly explorative theory building in an evolutionary and active way.

References

Booch, G. et al. (2007) Object-Oriented Analysis and Design with Applications, Addison-Wesley, Reading MA.
Budd, T. (1997) An Introduction to Object-Oriented Programming, Addison-Wesley, Reading MA.


Best Practices in Project Management Through a Grounded Theory Lens

Svetla Georgieva¹ and George Allan²
¹Technical University of Sofia, Bulgaria
²University of Portsmouth, UK
george.allan@port.ac.uk

Abstract: Grounded Theory is still regarded with skepticism as a research methodology by many academics so this paper has been specifically constructed using recent research to illustrate the rigour of the Grounded Theory Methodology. An investigation into ‘best practice amongst European Project Managers used the Grounded Theory Methodology to analyse interview data. This paper reports on that practical application of the Grounded Theory Methodology and contributes to the growing literature on the use of Grounded Theory as a research methodology. The three individual data analysis methods of Open-Coding, Constant Comparison, and Memorising are shown in operation and can be seen producing the findings. These findings are then operated on with the Grounded Theory method of Theoretical Coding and some deeply influential conclusions come out of this data analysis. The whole Grounded Theory methodology can be seen in operation and the concepts and categories can be seen emerging from the data. The contribution to the discipline of computing will be published elsewhere in technical journals but the research method is the main aim of this paper in this conference.

Keywords: Grounded theory methodology, open coding, theoretical coding, conceptualization, project management

1. Introduction

Project teams have varying experiences as a result of the skills/lack of skills displayed by their project manager (PM). Evidence is typified in this quote from a Bulgarian project team member who stated that:

“In my first job as a Web Developer in a Bulgarian company I had some really bad experiences with my PM. The last project which my team worked on was a failure. My colleagues and I worked without clear specification of the project requirements. As a result we developed software modules and then we had to redo them several times. The project missed its deadline although we worked overtime during the evenings and the weekends. At that time I didn’t know much about project management but I knew that things should not happen like that if a project is to succeed (Georgieva, 2007).

After this practitioner moved to another project with a different PM none of the above problems occurred. She realized that there are good as well as poor PMs. If the PM understands and uses good practices then projects can finish successfully. The result for the customers is delivery on time of the required products that work: the result for practitioners is a much better working experience.

The phrase ‘best practice’ is used by many professional bodies to encompass those practices that lead to success in that profession. There is not one professional body recognized as representative of all practitioners in USA, Europe, Asia, Australasia or Africa. Nor should there be, when we consider that those differing cultures will have differing needs and ways of thinking. Projects in differing cultural backgrounds have to be organized and operated in that cultural setting. However, research in any setting will also be valuable in allied settings. Projects the world over share some commonalities that practitioners and PMs recognize as the processes in project management. These include the processes of requirements gathering and analysis, design, implementation, system integration, going-live and project close-down. The main question addressed in the following research was: given that a project manager follows the usual processes in a project, what are the ‘best practices’ used by PMs that result in really successful outcomes? This paper reports on research in UK and Eastern Europe that investigated this question and attempted to establish the characteristics of best practice.

2. Research methodology

The intellectual considerations involved in the performance of best practice reside inside the minds of good PMs. Therefore data collection had to be by interview in order to bring out the key points inside PMs’ minds ready for analysis. Data analysis used Grounded Theory to open-code, conceptualize and categorize by constant comparison and memoing (Glaser and Strauss, 1967; Charmaz, 2001; Urquhart, 2001; Urquhart and Fernández, 2006). The next two sections describe the principles of the data collection and the data analysis in this research.
2.1 Data collection principles

There is a problem embedded in data collection by interviews – that of finding ‘good’ PMs who embed ‘best practice’ into their work. There were PMs in three organizations in Bulgaria and 3 in the UK who were known to practise ‘good’ project management and would therefore yield useful information. The main limitation was in obtaining data from Bulgarian participants. Due to time and financial constraints it was not possible to visit them for face-to-face interviews. Interviews were conducted by Skype. For an interview to be good it is necessary for the interviewer not only to hear the interviewees’ responses but also to see reactions. This may have had an effect on the subsequent analysis but it is felt that this was not a major factor at this stage and future research could extend the range of this work.

Initial contact was made to each interviewee explaining the research aim and setting the scene for data collection. This served to explore the willingness of PMs and assess the feasibility of using that company and contact. Data was collected in a subsequent in-depth interview with the possibility of a third short interview to clarify any points of uncertainty.

Six PMs from 3 Bulgarian and 3 UK companies were interviewed. Participants came from TyrusBlu Ltd., Sciant Ltd and Amexis Ltd. in Bulgaria. TyrusBlu is a technology intensive company, delivering software and engineering solutions and services to enterprise customers worldwide. There were two PMs from Sciant which is the largest outsourcing company in Bulgaria providing complete software solutions international corporate customers. Amexis is a developing company which specializes in healthcare systems development, business intelligence and data warehouse design and support.

In the UK, interviews were conducted with PMs from Zurich Finance Group, an insurance-based financial services provider with headquarters in Zurich, Switzerland and offices in North America, Europe, Asia Pacific and Latin America (the PM selected for interview was based in London and responsible for developing medium-sized projects to support financial business processes); Zurich Training Centre in southern England responsible for development and maintenance projects on tight time schedules with inadequate resource provision; Portsmouth University where the business focus is on the provision of Services to approximately 15,000 customers requiring to support a wide variety of job functions.

Each interview consists of noise and information (Allan, 2005). It is important to distinguish between noise (which adds nothing to the research) and information (which is important to the research). This important information contains the key points relevant to the research area. The researcher must distinguish between ‘noise’ and ‘key points’ in the data and collect only key points relevant to the research area. This is necessary for meaningful analysis.

Nearly all qualitative research methods recommend that interviews be pre-planned to follow a structured or semi-structured framework that facilitates data analysis. However, the selected research method was Grounded Theory (the methodology is explained in the next section) which does not work with predefined or preconceived structured questions (open or closed) as these will tell the interviewee what to think and bias their responses but interviews need to be guided to avoid time-wasting. The following areas of concern were used to facilitate keeping interviews in focus without biasing the responses from interviewees:

- What the PM does when starting a new project including starting the planning process;
- How plans are updated during progress;
- Dealing with customer’s requirements and inevitable changes;
- Work task definition and distribution to team members;
- Monitoring and controlling work products;
- Dealing with changes;
- Meeting deadlines.

These areas were used as focus prompts in a way that encouraged interviewees to talk about what they did in the areas and how they did those things. This opens a discourse rich in data without the interviewer biasing the discussion.

With the above principles in mind a series of interviews was arranged for data collection.
2.2 Data analysis principles

Data was analyzed using the Grounded Theory Methodology (GTM) of Glaser and Strauss (1967) because, compared to other methods for qualitative data analysis, GTM has a number of significant advantages.

1. Grounded Theory is a complete methodology which consists of six different methods for data analysis:
   - Open coding of the data for conceptual understanding.
   - Constant comparison of codes, concepts, and categories as they emerge from the data.
   - Memoing for clarity of thought.
   - Discovery of the Core Category which becomes the focus for selective coding.
   - Theoretical coding that investigates the links between categories.

These five methods are applied systematically to the data collected (Allan, 2006). Using them systematically makes the analysis coherent and rigorous leading to holistic conclusions.

2. Grounded Theory does not require a preconceived hypothesis. Traditional research designs usually start with a literature review to form a hypothesis which is then tested by experimentation in the real world (Allan, 2005). On the other hand, in Grounded Theory “the rule for generation of theory is not to have any pre-set hypotheses” (Glaser and Strauss, 1967, p. 194). In this research project, the little evidence about Best Practices in project management found in the literature does not allow a valuable preconceived hypothesis to be developed. Therefore, Grounded Theory is the appropriate method for data analysis in this project.

3. Conceptualization vs. description and explanation: GTM goes beyond the methods for qualitative data analysis described in the previous section because it tries to avoid descriptions and explanations by using conceptualization of the processes underlying the data (Glaser, 1978; Allan, 2006). The emphasis is on the most important property of conceptualization for Grounded Theory – “concepts are abstract of time, place and people” (Glaser, 2001, p. 10). By identifying the relationships between concepts at different levels of abstraction, Grounded Theory develops an explanation of why events are happening the way they are.

The explanation of real-world events that emerges from using GTM is grounded in the data and therefore called a grounded theory. What does it mean for a theory to be grounded in the data? Glaser and Strauss (1967) pointed out that a discovered grounded theory combines concepts and hypotheses that have emerged from the data. This means that the emerging theory is valid and reliable and does not require further proving or testing in the real world because it comes directly from the real world data itself.

This paper is structured as follows: section two focuses on the data collection followed by a section on data analysis. When using GTM, concepts and categories emerge from the data analysis, which means that findings and conclusions come directly out of the data and these are presented in a final section.

3. Data collection by interview

It is a fact that GTM does not work with predefined questions but there is a need to save time by keeping each interview within the research focal area. Readers are reminded that the aim was PM best practice and the following serve as area of concern to facilitate each interview: what the PM does when a new project starts, including starting the planning process and how this is updated during progress; dealing with customer’s requirements and inevitable changes; work task definition and distribution to team members; monitoring and controlling work products; dealing with deadlines and changes.

The objective of each interview was to search the minds of the PMs for characteristics of good PM practices. Relevant information sorted from the interview ‘noise’ was labelled, first with the initial letter of the company name (e.g. T, B, S etc.), then with a number representing the interview (e.g. T1), and finally, with a number, representing the data point relevant to PM practice (e.g. T1.1, T1.2, etc.).

There were 19 Key Points that emerged from the first interview; these are listed in the appendix and the end of this paper. The defining characteristics of the Grounded Theory Methodology is that immediately each point is collected it is compared with previous points looking for similarities and connections. Key point T1.2 was compared with T1.1 and straight away we notice that something relating to the customer is emerging.

- T1.1. Establishing good relationship with the customer
- T1.2. Developing and showing the customer something small/a small project
Similarly T1.3 was compared with T1.1 looking for connections
T1.3. Establishing everyday communication – regular meetings, conference calls, reports
And then compared with T1.2 etc.

Here are the remaining key points but it must be stressed that this is where many GT researchers make a mistake – they analyse their data into a long list of meta-data without constantly comparing each new data item with previous data items to look for connections.

4. Data Analysis using Grounded Theory
The data analysis method was the Grounded Theory Methodology (GTM) as developed as a research method by Glaser and Strauss (1967).

4.1 Emergence of concepts
It can be seen from the characteristics of the above key points that there are commonalities. For example, the following key points refer to establishing, improving and maintaining the relationship with the customer:
T1.1. Establishing good relationship with the customer
T1.3. Establishing everyday communication – regular meetings, conference calls, reports
T1.4a. Establishing very strong everyday communication
T1.18. Talking with the customer in advance in case you are going to miss a deadline
T1.19. Being always ready to give an explanation when saying something to the customer

The relationship with the customer is established in the beginning of the project (T1.1). Everyday communication with the customer (T1.3, T1.4a) starts in the beginning of the project and has to continue during the project life. The PM has to talk with the customer many times during the project (T1.18), he should always be ready to give the customer explanations (T1.19). Therefore, the relationship with the customer, once established, has to be developed, improved and maintained during the project life. This can be represented by Concept 1 – Establishing and developing the relationship with the customer. ‘Developing’ was chosen instead of ‘improving’, because the process of developing includes the process of improving, of making the relationship better. ‘Developing’ was chosen instead of ‘maintaining’, because before developing the relationship, its current level has to be maintained.

- Concept 1 – Establishing and developing customer relationship

Other key points were grouped together and the following concepts emerged:
- Concept 2 – Building customer confidence emerged from key points T1.2, T1.3, T1.4, T1.4a, T1.8, T1.18 and T1.19.
- Concept 3 – Communicating with the customer emerged from key points T1.3, T1.4a, T1.10, T1.18 and T1.19.
- Concept 4 – Managing customer requirements emerged from key points T1.10 and T1.11.
- Concept 4 has a necessary property which is Gathering and assessing customer requirements emerged from key points T1.5 and T1.6.
- Concept 5 – Managing the team: emerged from key points T1.13 and T1.14. and has a property ‘Internal communication’ emerged from key points T1.5 and T1.13.

This is how concepts emerge from the data.

4.2 Emergence of categories
Reflecting on the concepts it becomes clear that some of them share common characteristics and can be grouped in higher level groups called categories. For example, the concept of Establishing and developing customer relationship; the concept of Building customer confidence; the concept of Communicating with the customer; the concept of Gathering and assessing customer requirements are all related to the customer. Establishing and developing the relationship with the customer, building his confidence, communicating with him, gathering, assessing and managing his requirements can be grouped into the category Managing the customer, because all of these processes are important for project success and cannot be left to happen on their own. The project manager is the one who manages these processes, the PM manages the customer. Therefore, the five concepts belong to the category of Managing the customer.
The following concepts came from the data analysis: Managing customer requirements; Gathering and assessing customer requirements; Managing the team; Internal communication; Planning; Testing; Monitoring the progress. These are related to the project work. Customer requirements must be gathered, assessed and managed in order to have a clear understanding of what the project work is about. Managing the team is also directly related to the amount and technical difficulties of the project work. Without internal communication, that management is impossible. Planning, testing and monitoring the progress of the project work are some of the main processes used to manage the project work. Therefore, these seven concepts can be grouped together in the Category of Managing the project work:

![Diagram of concepts and categories]

Figure 1: Concepts and categories from the analysis of interview T-1

These emergent codes, concepts and categories were used as a basis for coding the next interview as well as keeping an open mind so that more codes, concepts and categories emerged. There is insufficient space here to include the full data analysis but it is available on request to the lead author. Suffice it to say that the analysis gathered great momentum as the research progressed and existing concepts were confirmed with more and more evidence from subsequent interviews while further concepts emerged. This is typical in GTM and is part of the excitement generated in this research methodology.

The result of analysing the data from six interviews and four follow-up interviews is represented in Figure 2 below. Seven categories emerged: Managing the Customer; Managing the Project Work; Managing Communications; Managing Stakeholders; Awareness (which encapsulates Managing Awareness); Motivating; Knowledge Transfer. These Categories are shown in bold in Figure 2. The codes, concepts and categories that emerged from this first interview will form the basis for the next and subsequent interview data analysis.
5. Developing a theory of project management best practice

At this stage in the GTM analysis we look for relationships and connections between the core category and the other categories. This is known in grounded theory as theoretical coding, a term coined by Glaser and Strauss (1967). However, the four categories of Managing the Customer, Managing the Project Work,
Managing Communications and Managing Stakeholders are well known and well established in the literature so no further comment will be made in this text other than stating the obvious that they must be included in project management best practice. The three categories of Awareing, Motivating and Knowledge Transfer are interesting and less well documented in connection to project management and this section will focus primarily on these in this research report. A full analysis of the relationships between all the categories is available from the authors.

5.1 Awareing

From Figure 2 it can be seen that all the concepts in Awareing also belong in other categories. This ties Awareing to all but two of the other categories which shows its interconnectedness and thereby its strong position in project management best practice. However, in practice little notice is taken of making people aware or of the PM being aware. Thus, two types of awareness can be distinguished: the PM's awareness and other people's awareness (more appropriately called Stakeholder's awareness).

The PM's awareness consists of four parts: PM's awareness of herself/himself, PM's awareness of stakeholders, PM's awareness of project issues and the PM's awareness of Stakeholder's awareness. Best practice occurs when the PM considers all four awarenesses.

Project stakeholders include development team members, customers, top management, other PMs on other projects within the organization, suppliers and subcontractors. The concept of awareing encapsulates the fact that a good PM has to manage the awareness of all these people. In particular, the PM has to manage the awareness of the customers by appraising them at all times of risks, problems and difficulties, project progress, need for communication and need for clear requirements. It is not enough for the PM to inform customers of these issues, the PM has to make the customer aware. This is another factor in project management best practice.

Similarly, the PM has to make the development team aware of the project's importance, context and benefits, of the risks, of the need for internal communication, of the customer's changes to requirements, and of the actual project work. All this comes from the analysis and can be seen in Figure 2.

5.2 Motivating

The category of Motivating has four concepts regarded now as properties of motivating. These are Responsibilising Self-motivation; Awareing the team and Leadership. Responsibilising encapsulates delegating responsibility (as well as project work) and managing how team members handle that responsibility. An ability to do this is regarded as a contribution to project management best practice. Any absence of self-motivation in a PM must have a negative influence in project work and project success, whereas good self-motivation will contribute to project success. This is a subject worthy of future research.

Awareing the project team of difficulties and problems appeared to help the project workers avoid problems rather than having to fix them after they had occurred. Good leadership, good responsibilising and high PM motivation will motivate the development team, whereas poor leadership coupled with poor responsibilising and low PM motivation will inhibit a team's motivation. Therefore good leadership coupled with good responsibilising and high PM motivation are key to project success and therefore elements in best practice. However, the three elements of good leadership, good responsibilising and high PM motivation are tough concepts and difficult to achieve in practice.

There emerged a strong connection between motivation and being aware of motivation. The relationship between these two parts defines the connection between motivation and managing awareness. The two parts are absorbed by each other, their concepts transfer by osmosis.

5.3 Knowledge transfer

The PM has an important role in knowledge transfer (KT) because it is the PM who has to make KT happen. When KT does take place the PM has to be aware of the possible difficulties in the actual process of the transfer as well as the accuracy of the knowledge itself. When KT does not happen, the PM has to be aware of the lack of transfer and the need for KT to happen.

Knowledge transfer can be informal in the work place between team members or between the PM and technical staff or higher management. Alternatively, KT can be formal in the form of minutes of meetings and
official documents. From this research came the concept of unused knowledge, that is where KT did not take place for undisclosed reasons. Therefore KT has properties of informal, formal, and unused. The area of KT and its properties would make interesting future research projects.

The main findings form this section are that there is unused KT and that KT is often informal.

6. Conclusions

From this research with practicing PMs, seven categories emerged that encapsulate their best practices: Managing the Customer, Managing the Project Work, Managing Communications, Managing Stakeholders, Awareness (which encapsulates Managing Awareness), Motivating, and Knowledge Transfer. Focusing on the last three and the connections between them a number of conclusions can be drawn as contributing elements of best practice in project management.

1. Being aware of individual stakeholder’s awareness is necessary for the PM to be able to manage projects well. When a PM manages the customer’s awareness, this increases the probability of success and reduces risk.

2. The more diligently a PM manages the development team’s awareness, the more successful will be the project work and the higher the probability of project success. Good communications play a vital role here because communications help awareness.

3. The PM’s awareness of motivation is key to motivating everyone else including the customer. It is not enough for a PM to motivate the team, the PM has to monitor and manage the state of the team’s motivation. Good leadership combined with good responsibilising and high PM motivation will increase everyone else’s motivation and lead to reduced risk of project failure. However, good leadership, good responsibilising and high motivation are all tough concepts which do not occur naturally and need to be developed.

4. When the process of knowledge transfer is managed well, knowledge will flow, accumulate and build up and this will promote better management in the team, the project work, the customer, other stakeholders and every aspect of project management. Knowledge transfer increases motivation and is an essential element in good leadership.

5. Motivation, knowledge transfer and managing awareness are inextricably linked together. They and their properties (Figure 2) constitute elements of project management best practice as a result of this research. Future research should build on this to clarify their interlinked roles and discover other properties of project management best practice.

Acknowledgments

We acknowledge the voluntary help of the 6 Project Managers and thank them for all their time, effort and interest in this research. The findings were fed back and have led to innovative thinking and changes. We also thank Christine Dawe of the University of Chichester for proof-reading this manuscript and her valuable advice on points of grammar.

References

Allan, G.W., (2005), The Use of Grounded Theory as a Research Method: warts & all, in: Ray Hackney (Ed.) Proceedings of the International Symposium on Research Methods, Tampa, USA
Atkinson, R., (1999), Project management: Cost, time and quality, two best guesses and a phenomenon, its time to accept other success criteria, International Journal of Project Management, 17, 337-342
Georgieva, S., (2007), Project Management paper, University of Portsmouth, internal publication
Appendix

List of Key Points from Interview 1

T1.1. Establishing good relationship with the customer
T1.2. Developing and showing the customer something small/a small project
T1.3. Establishing everyday communication – regular meetings, conference calls, reports
T1.4. Building the customer's confidence
T1.4a. Establishing very strong everyday communication
T1.5. Discussing and assessing the functional requirements with the team of developers
T1.6. Developing detailed Software Requirements Specification
T1.7. Researching about the new things in the project – new technologies, techniques, methods
T1.8. Implementing the most important parts of the new application first
T1.9. Tightly defined deliverables – every 2 or 3 weeks
T1.10. Saying 'no' to the customer instead of always agreeing with him
T1.11. Managing and changing the customer's wishes, ideas and requirements
T1.12. Changing project plan mechanism
T1.13. Going personally to each developer to see at what stage of the task he/she is at the moment
T1.14. Defining tasks according to the developer's skills set
T1.15. Developing and using a Test Plan
T1.16. Testing before each deliverable and at the end of the project
T1.17. Using a bug-tracking system
T1.18. Talking with the customer in advance in case you are going to miss a deadline
T1.19. Being always ready to give an explanation when saying something to the customer
Structural Equation Modelling: Guidelines for Determining Model Fit

Daire Hooper¹, Joseph Coughlan¹, and Michael R. Mullen²
¹Dublin Institute of Technology, Dublin, Republic of Ireland
²Florida Atlantic University, Florida, USA
daire.hooper@dit.ie
joseph.coughlan@dit.ie
mullen@fau.edu

Abstract: The following paper presents current thinking and research on fit indices for structural equation modelling. The paper presents a selection of fit indices that are widely regarded as the most informative indices available to researchers. As well as outlining each of these indices, guidelines are presented on their use. The paper also provides reporting strategies of these indices and concludes with a discussion on the future of fit indices.

Keywords: Structural equation modelling, fit indices, covariance structure modelling, reporting structural equation modelling, model fit

1. Introduction

Structural Equation Modelling (SEM) has become one of the techniques of choice for researchers across disciplines and increasingly is a ‘must’ for researchers in the social sciences. However the issue of how the model that best represents the data reflects underlying theory, known as model fit, is by no means agreed. With the abundance of fit indices available to the researcher and the wide disparity in agreement on not only which indices to report but also what the cut-offs for various indices actually are, it is possible that researchers can become overwhelmed by the conflicting information available. It is essential that researchers using the technique are comfortable with the area since assessing whether a specified model ‘fits’ the data is one of the most important steps in structural equation modelling (Yuan, 2005). This has spurred decades of intense debate and research dedicated to this pertinent area. Indeed, ever since structural equation modelling was first developed, statisticians have sought and developed new and improved indices that reflect some facet of model fit previously not accounted for. Having such a collection of indices entices a researcher to select those that indicate good model fit. This practice should be desisted at all costs as it masks underlying problems that suggest possible misspecifications within the model.

This paper seeks to introduce a variety of fit indices which can be used as a guideline for prospective structural equation modellers to help them avoid making such errors. To clarify matters to users of SEM, the most widely respected and reported fit indices are covered here and their interpretive value in assessing model fit is examined. In addition to this, best practice on reporting structural equation modelling is discussed and suggests some ways in which model fit can be improved. In recent years the very area of fit indices has come under serious scrutiny with some authors calling for their complete abolishment (Barrett, 2007). While such a drastic change is unlikely to occur any time soon, the shortcomings of having stringent thresholds is becoming more topical within the field of structural equation modelling (Kenny and McCoach, 2003; Marsh et al, 2004).

2. Absolute fit indices

Absolute fit indices determine how well an a priori model fits the sample data (McDonald and Ho, 2002) and demonstrates which proposed model has the most superior fit. These measures provide the most fundamental indication of how well the proposed theory fits the data. Unlike incremental fit indices, their calculation does not rely on comparison with a baseline model but is instead a measure of how well the model fits in comparison to no model at all (Jöreskog and Sörbom, 1993). Included in this category are the Chi-Squared test, RMSEA, GFI, AGFI, the RMR and the SRMR.

2.1 Model chi-square ($\chi^2$)

The Chi-Square value is the traditional measure for evaluating overall model fit and, ‘assesses the magnitude of discrepancy between the sample and fitted covariances matrices’ (Hu and Bentler, 1999: 2). A good model fit would provide an insignificant result at a 0.05 threshold (Barrett, 2007), thus the Chi-Square statistic is often referred to as either a ‘badness of fit’ (Kline, 2005) or a ‘lack of fit’ (Mulaik et al, 1989) measure. While the Chi-Squared test retains its popularity as a fit statistic, there exist a number of severe
limitations in its use. Firstly, this test assumes multivariate normality and severe deviations from normality may result in model rejections even when the model is properly specified (McIntosh, 2006). Secondly, because the Chi-Square statistic is in essence a statistical significance test it is sensitive to sample size which means that the Chi-Square statistic nearly always rejects the model when large samples are used (Bentler and Bonnet, 1980; Jöreskog and Sörbom, 1993). On the other hand, where small samples are used, the Chi-Square statistic lacks power and because of this may not discriminate between good fitting models and poor fitting models (Kenny and McCoach, 2003). Due to the restrictiveness of the Model Chi-Square, researchers have sought alternative indices to assess model fit. One example of a statistic that minimises the impact of sample size on the Model Chi-Square is Wheaton et al’s (1977) relative/normed chi-square ($\chi^2/df$). Although there is no consensus regarding an acceptable ratio for this statistic, recommendations range from as high as 5.0 (Wheaton et al, 1977) to as low as 2.0 (Tabachnick and Fidell, 2007).

2.2 Root mean square error of approximation (RMSEA)

The RMSEA is the second fit statistic reported in the LISREL program and was first developed by Steiger and Lind (1980, cited in Steiger, 1990). The RMSEA tells us how well the model, with unknown but optimally chosen parameter estimates would fit the populations covariance matrix (Byrne, 1998). In recent years it has become regarded as ‘one of the most informative fit indices’ (Diamantopoulos and Siguaw, 2000: 85) due to its sensitivity to the number of estimated parameters in the model. In other words, the RMSEA favours parsimony in that it will choose the model with the lesser number of parameters. Recommendations for RMSEA cut-off points have been reduced considerably in the last fifteen years. Up until the early nineties, an RMSEA in the range of 0.05 to 0.10 was considered an indication of fair fit and values above 0.10 indicated poor fit (MacCallum et al, 1996). It was then thought that an RMSEA of between 0.08 to 0.10 provides a mediocre fit and below 0.08 shows a good fit (MacCallum et al, 1996). However, more recently, a cut-off value close to .06 (Hu and Bentler, 1999) or a stringent upper limit of 0.07 (Steiger, 2007) seems to be the general consensus amongst authorities in this area.

One of the greatest advantages of the RMSEA is its ability for a confidence interval to be calculated around its value (MacCallum et al, 1996). This is possible due to the known distribution values of the statistic and subsequently allows for the null hypothesis (poor fit) to be tested more precisely (McQuitty, 2004). It is generally reported in conjunction with the RMSEA and in a well-fitting model the lower limit is close to 0 while the upper limit should be less than 0.08.

2.3 Goodness-of-fit statistic (GFI) and the adjusted goodness-of-fit statistic (AGFI)

The Goodness-of-Fit statistic (GFI) was created by Jöreskog and Sorbom as an alternative to the Chi-Square test and calculates the proportion of variance that is accounted for by the estimated population covariance (Tabachnick and Fidell, 2007). By looking at the variances and covariances accounted for by the model it shows how closely the model comes to replicating the observed covariance matrix (Diamantopoulos and Siguaw, 2000). This statistic ranges from 0 to 1 with larger samples increasing its value. When there are a large number of degrees of freedom in comparison to sample size, the GFI has a downward bias (Sharma et al, 2005). In addition, it has also been found that the GFI increases as the number of parameters increases (MacCallum and Hong, 1997) and also has an upward bias with large samples (Bollen, 1990; Miles and Shevlin, 1998). Traditionally an omnibus cut-off point of 0.90 has been recommended for the GFI however, simulation studies have shown that when factor loadings and sample sizes are low a higher cut-off of 0.95 is more appropriate (Miles and Shevlin, 1998). Given the sensitivity of this index, it has become less popular in recent years and it has even been recommended that this index should not be used (Sharma et al, 2005). Related to the GFI is the AGFI which adjusts the GFI based upon degrees of freedom, with more saturated models reducing fit (Tabachnick and Fidell, 2007). Thus, more parsimonious models are preferred while penalised for complicated models. In addition to this, AGFI tends to increase with sample size. As with the GFI, values for the AGFI also range between 0 and 1 and it is generally accepted that values of 0.90 or greater indicate well fitting models. Given the often detrimental effect of sample size on these two fit indices they are not relied upon as a stand alone index, however given their historical importance they are often reported in covariance structure analyses.

2.4 Root mean square residual (RMR) and standardised root mean square residual (SRMR)

The RMR and the SRMR are the square root of the difference between the residuals of the sample covariance matrix and the hypothesised covariance model. The range of the RMR is calculated based upon the scales of each indicator, therefore, if a questionnaire contains items with varying levels (some items may range from 1 – 5 while others range from 1 – 7) the RMR becomes difficult to interpret (Kline, 2005). The
standardised RMR (SRMR) resolves this problem and is therefore much more meaningful to interpret. Values for the SRMR range from zero to 1.0 with well fitting models obtaining values less than .05 (Byrne, 1998; Diamantopoulos and Siguaw, 2000), however values as high as 0.08 are deemed acceptable (Hu and Bentler, 1999). An SRMR of 0 indicates perfect fit but it must be noted that SRMR will be lower when there is a high number of parameters in the model and in models based on large sample sizes.

3. Incremental fit indices

Incremental fit indices, also known as comparative (Miles and Shevlin, 2007) or relative fit indices (McDonald and Ho, 2002), are a group of indices that do not use the chi-square in its raw form but compare the chi-square value to a baseline model. For these models the null hypothesis is that all variables are uncorrelated (McDonald and Ho, 2002).

3.1 Normed-fit index (NFI)

The first of these indices to appear in LISREL output is the Normed Fit Index (NFI: Bentler and Bonnet, 1980). This statistic assesses the model by comparing the $\chi^2$ value of the model to the $\chi^2$ of the null model. The null/independence model is the worst case scenario as it specifies that all measured variables are uncorrelated. Values for this statistic range between 0 and 1 with Bentler and Bonnet (1980) recommending values greater than 0.90 indicating a good fit. More recent suggestions state that the cut-off criteria should be NFI ≥ .95 (Hu and Bentler, 1999). A major drawback to this index is that it is sensitive to sample size, underestimating fit for samples less than 200 (Mulaik et al, 1989; Bentler, 1990), and is thus not recommended to be solely relied on (Kline, 2005). This problem was rectified by the Non-Normed Fit Index (NNFI, also known as the Tucker-Lewis index), an index that prefers simpler models. However in situations where small samples are used, the value of the NNFI can indicate poor fit despite other statistics pointing towards good fit (Bentler, 1990; Kline, 2005; Tabachnick and Fidell, 2007). A final problem with the NNFI is that due to its non-normed nature, values can go above 1.0 and thus be difficult to interpret (Byrne, 1998). Recommendations as low as 0.80 as a cutoff have been preferred however Bentler and Hu (1999) have suggested NNFI ≥ 0.95 as the threshold.

3.2 CFI (Comparative fit index)

The Comparative Fit Index (CFI: Bentler, 1990) is a revised form of the NFI which takes into account sample size (Byrne, 1998) that performs well even when sample size is small (Tabachnick and Fidell, 2007). This index was first introduced by Bentler (1990) and subsequently included as part of the fit indices in his EQS program (Kline, 2005). Like the NFI, this statistic assumes that all latent variables are uncorrelated (null/independence model) and compares the sample covariance matrix with this null model. As with the NFI, values for this statistic range between 0.0 and 1.0 with values closer to 1.0 indicating good fit. A cut-off criterion of CFI ≥ 0.90 was initially advanced however, recent studies have shown that a value greater than 0.90 is needed in order to ensure that misspecified models are not accepted (Hu and Bentler, 1999). From this, a value of CFI ≥ 0.95 is presently recognised as indicative of good fit (Hu and Bentler, 1999). Today this index is included in all SEM programs and is one of the most popularly reported fit indices due to being one of the measures least effected by sample size (Fan et al, 1999).

4. Parsimony fit indices

Having a nearly saturated, complex model means that the estimation process is dependent on the sample data. This results in a less rigorous theoretical model that paradoxically produces better fit indices (Mulaik et al, 1989; Crowley and Fan, 1997). To overcome this problem, Mulaik et al (1989) have developed two parsimony of fit indices; the Parsimony Goodness-of-Fit Index (PGFI) and the Parsimonious Normed Fit Index (PNFI). The PGFI is based upon the GFI by adjusting for loss of degrees of freedom. The PNFI also adjusts for degrees of freedom however it is based on the NFI (Mulaik et al 1989). Both of these indices seriously penalise for model complexity which results in parsimony fit index values that are considerably lower than other goodness of fit indices. While no threshold levels have been recommended for these indices, Mulaik et al (1989) do note that it is possible to obtain parsimony fit indices within the .50 region while other goodness of fit indices achieve values over .90 (Mulaik et al 1989). The authors strongly recommend the use of parsimony fit indices in tandem with other measures of goodness-of-fit however, because no threshold levels for these statistics have been recommended it has made them more difficult to interpret.
A second form of parsimony fit index are those that are also known as ‘information criteria’ indices. Probably the best known of these indices is the Akaike Information Criterion (AIC) or the Consistent Version of AIC (CAIC) which adjusts for sample size (Akaike, 1974). These statistics are generally used when comparing non-nested or non-hierarchical models estimated with the same data and indicates to the researcher which of the models is the most parsimonious. Smaller values suggest a good fitting, parsimonious model however because these indices are not normed to a 0-1 scale it is difficult to suggest a cut-off other than that the model that produces the lowest value is the most superior. It is also worth noting that these statistics need a sample size of 200 to make their use reliable (Diamantopoulos and Siguaw, 2000).

5. Reporting fit indices

With regards to which indices should be reported, it is not necessary or realistic to include every index included in the program’s output as it will burden both a reader and a reviewer. Given the plethora of fit indices, it becomes a temptation to choose those fit indices that indicate the best fit (see Appendix A for a summary of some key indices discussed herein). This should be avoided at all costs as it is essentially sweeping important information under the carpet. In a review by McDonald and Ho (2002) it was found that the most commonly reported fit indices are the CFI, GFI, NFI and the NNFI. When deciding what indices to report, going by what is most frequently used is not necessarily good practice as some of these statistics (such as the GFI discussed above) are often relied on purely for historical reasons, rather than for their sophistication. While there are no golden rules for assessment of model fit, reporting a variety of indices is necessary (Crowley and Fan 1997) because different indices reflect a different aspect of model fit. Although the Model Chi-Square has many problems associated with it, it is still essential that this statistic, along with its degrees of freedom and associated p value, should at all times reported (Kline, 2005; Hayduk et al, 2007). Threshold levels were recently assessed by Hu and Bentler (1999) who suggested a two-index presentation format. This always includes the SRMR with the NNI (TLI), RMSEA or the CFI. The various combinations are summarised in Appendix B below. Kline (2005) speaks strongly about which indices to include and advocates the use of the Chi-Square test, the RMSEA, the CFI and the SRMR. Boomsma (2000) has similar recommendations but also advises for the squared multiple correlations of each equation to be reported. Based on these authors guidelines and the above review it is sensible to include the Chi-Square statistic, its degrees of freedom and p value, the RMSEA and its associated confidence interval, the SRMR, the CFI and one parsimony fit index such as the PNFI. These indices have been chosen over other indices as they have been found to be the most insensitive to sample size, model misspecification and parameter estimates.

6. How to improve model fit

Given the complexity of structural equation modelling, it is not uncommon to find that the fit of a proposed model is poor. Allowing modification indices to drive the process is a dangerous game, however, some modifications can be made locally that can substantially improve results. It is good practice to assess the fit of each construct and its items individually to determine whether there are any items that are particularly weak. Items with low multiple r² (less than .20) should be removed from the analysis as this is an indication of very high levels of error. Following this, each construct should be modelled in conjunction with every other construct in the model to determine whether discriminant validity has been achieved. The Phi (ϕ) value between two constructs is akin to their covariance, therefore a Phi of 1.0 indicates that the two constructs are measuring the same thing. One test which is useful to determine whether constructs are significantly different is Bagozzi et al’s (1991) discriminant validity test. The formula for this is: parameter estimate (ϕ value) ± 1.96 * standard error. If the value is greater than 1.0 discriminant validity has not been achieved and further inspections of item cross-loadings need to be made. Items with high Lambda-Y modification indices are possible candidates for deletion and are likely to be causing the discriminant validity problem. By deleting indiscriminant items fit is likely to improve and is advantageous in that it is unlikely to have any major theoretical repercussions.

A further way in which fit can be improved is through the correlation of error terms. This practice is generally frowned upon (Gerbing and Anderson, 1984) as it means that there is some other issue that is not specified within the model that is causing the covariation. If a researcher decides to correlate error terms there needs to be strong theoretical justification behind such a move (Jöreskog and Long, 1993). Correlating within-factor error is easier to justify than across latent variable correlations, however it is essential that the statistical and substantive impact are clearly discussed. If a researcher feels they can substantiate this decision, correlated error terms is acceptable, however it is a step that should be taken with caution.
7. The future for fit indices

As a final point it must be noted that while fit indices are a useful guide, a structural model should also be examined with respect to substantive theory. By allowing model fit to drive the research process it moves away from the original, theory-testing purpose of structural equation modelling. In addition, fit indices may point to a well-fitting model when in actual fact, parts of the model may fit poorly (Jöreskog and Sörbom, 1996; Tomarken and Waller, 2003; Reisinger and Mavondo, 2006). Indeed, the area of fit indices ‘rules of thumb’ is highly topical at the moment with some experts in the area calling for a complete abandonment of fit indices altogether (Barrett, 2007). Others are less hesitant to denounce their usefulness but do agree that strictly adhering to recommended cutoff values can lead to instances of Type I error (the incorrect rejection of an acceptable model) (Marsh et al, 2004). Although the debate is ongoing, it is unlikely that fit indices will become defunct any time soon and for this reason the above literature review remains pertinent to those using SEM.

References


**Appendix A**

**Table 1: Fit indices and their acceptable thresholds**

<table>
<thead>
<tr>
<th>Fit Index</th>
<th>Acceptable Threshold Levels</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Absolute Fit Indices</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chi-Square $\chi^2$</td>
<td>Low $\chi^2$ relative to degrees of freedom with an insignificant $p$ value ($p &gt; 0.05$)</td>
<td></td>
</tr>
<tr>
<td>Root Mean Square Error of Approximation (RMSEA)</td>
<td>Values less than 0.07 (Steiger, 2007)</td>
<td>Has a known distribution. Favours parsimony. Values less than 0.03 represent excellent fit.</td>
</tr>
<tr>
<td>GFI</td>
<td>Values greater than 0.95</td>
<td>Scaled between 0 and 1, with higher values indicating better model fit. This statistic should be used with caution.</td>
</tr>
<tr>
<td>AGFI</td>
<td>Values greater than 0.95</td>
<td>Adjusts the GFI based on the number of parameters in the model. Values can fall outside the 0-1.0 range.</td>
</tr>
<tr>
<td>RMR</td>
<td>Good models have small RMR (Tabachnik and Fidell, 2007)</td>
<td>Residual based. The average squared differences between the residuals of the sample covariances and the residuals of the estimated covariances. Unstandardised.</td>
</tr>
<tr>
<td>SRMR</td>
<td>SRMR less than 0.08 (Hu and Bentler, 1999)</td>
<td>Standardised version of the RMR. Easier to interpret due to its standardised nature.</td>
</tr>
<tr>
<td><strong>Incremental Fit Indices</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NFI</td>
<td>Values greater than 0.95</td>
<td>Assesses fit relative to a baseline model which assumes no covariances between the observed variables. Has a tendency to overestimate fit in small samples.</td>
</tr>
<tr>
<td>NNFI (TLI)</td>
<td>Values greater than 0.95</td>
<td>Non-normed, values can fall outside the 0-1 range. Favours parsimony. Performs well in simulation studies (Sharma et al, 2005; McDonald and Marsh, 1990)</td>
</tr>
<tr>
<td>CFI</td>
<td>Values greater than 0.95</td>
<td>Normed, 0-1 range.</td>
</tr>
</tbody>
</table>
Appendix B

Table 2: Hu and Bentler’s Two-Index Presentation Strategy (1999)

<table>
<thead>
<tr>
<th>Fit Index Combination</th>
<th>Combinational Rules</th>
</tr>
</thead>
<tbody>
<tr>
<td>NNFI (TLI) and SRMR</td>
<td>NNFI of 0.96 or higher and an SRMR of .09 or lower</td>
</tr>
<tr>
<td>RMSEA and SRMR</td>
<td>RMSEA of 0.06 or lower and a SRMR of 0.09 or lower</td>
</tr>
<tr>
<td>CFI and SRMR</td>
<td>CFI of .96 or higher and a SRMR of 0.09 or lower</td>
</tr>
</tbody>
</table>
University Academics’ Psychological Contracts in Australia: A Mixed Method Research Approach

Branka Krivokapic-Skoko, and Grant O’Neill,
School of Marketing and Management, Charles Sturt University, NSW, Australia
bkrivokapic@csu.edu.au
goneill@csu.edu.au

Abstract: It has been argued that in a workplace environment that is characterised by significant change and uncertainty, the formation and content of the psychological contracts are of increasing importance regarding levels of employee trust, satisfaction, commitment and motivation, and teaching and research outcomes. While research has clearly demonstrated that psychological contracts can have considerable impact upon workplace relations and employee performance, research into the formation, content and effects of psychological contracts between academics and the University has been very limited.

The paper used a sequential multi methods research design to explore the formation and content of psychological contracts established by the academics within an Australian University. The empirical research began with exploratory focus group discussions which were followed by a mail survey. The focus groups were carried first to identify the issues and themes that can subsequently be drawn upon to assist with development of relevant survey questions. Focus groups sought to elicit insights and subjective interpretations of the psychological contracts and the consequences of perceived fulfilment or breach. This, first qualitative phase of research has identified four key foci of academic responsibility that greatly influenced the formation and effects of the psychological contracts that have been formed, and these are: the University, the discipline, society, and students. These four categories were used later on to further develop the questionnaire and carry out exploratory factor analysis (EFA) of a larger survey of the academics. Using exploratory factor analysis of the survey data, eight factors were discovered relating to the University’s obligations to its employees and three underlying factors were found in relation to individual academic’s obligations to the University. In terms of the University’s obligation to the academics, the EFA reinforces the importance of leadership and management, fairness and equity (notably in relation to promotion and provision of opportunities for career development). In terms of the academics’ perceived obligations to the university, the EFA points to the importance of role expectations and commitment to the job and student learning.

Keywords: Mixed methods; psychological contracts; academia

1. Introduction

The past two decades have seen enormous change in the number, funding, and focus of Australian universities, and further changes are inevitable in an increasingly global higher education market. Such changes affect the context and conditions of academic work mainly because of the rise of managerialism, increased external and internal accountability, performance management practices, commercialisation of higher education, and tighter funding (Winter & Sarros, 2000). Many of the changes and problems that Australian universities and academics have experienced have also been experienced by UK universities and academics (Newton, 2002).

It is in the context of such change and uncertainty within Australian universities that the following paper addresses the formation, content and effects of academics’ psychological contracts. We argue that in this era of diminished funding, greater competition, and heightened hierarchy and accountability, the content and effects of psychological contracts are critically important for academics and universities. Understanding and effectively managing the psychological contracts that develop can help organisations succeed and prosper. These contracts can motivate employees to fulfil commitments made to employers when they are confident that employers will reciprocate and fulfil what employees perceive to be their side of the contracts. It has been argued that perceived obligations within the psychological contract are often more important to job-related attitudes and behaviour, than are the formal and explicit elements of contractual agreements (Thomson & Bunderson, 2003).

The paper commences with a discussion of key features of the psychological contract, and briefly discusses past empirical research conducted within academia. The paper then presents empirical findings from applying a sequential mixed method approach to empirical research on the psychological contracts established by the academics employed by the Faculty of Business, Charles Sturt University, NSW, Australia.
2. Psychological contracts: A brief overview

There are two main conceptualisations of the psychological contract that are discussed in literature. The first addresses the perception that there are two parties in the employment relationship who have mutual obligations to each other: the organisation and the employee (Herriot, Manning, & Kidd, 1997). These mutual obligations may have been explicitly communicated through formal contracts or they may be implied through the expectations of organisations and employees. The second conceptualisation addresses the psychological contract as being formulated only in the mind of the employee. The psychological contract is, therefore, about “individual beliefs, shaped by the organisation, regarding the terms of an exchange between individuals and their organisation...A key feature of the psychological contract is that the individual voluntarily asserts to make and accept certain promises as he or she understands them” (Rousseau, 1995; 9-10). For example, the employee may believe that the organisation has agreed to certain commitments, such as providing job security, high pay, promotion, and training in exchange for the employee’s hard work and loyalty (Rousseau, 1990).

Beyond the consideration of who is actually party to a psychological contract the difficulties inherent in accurately defining these contracts arise from the fact that they are a subjective and idiosyncratic phenomenon. To begin with, the perceptual and individual nature of psychological contracts makes them distinct from formal written contracts. Further, these contracts are subjective and grounded in the social and cultural contexts where employers and employees believe they have reciprocal obligations and presumably share a common understanding of the nature of these obligations. However, the understanding of the expectations and mutual obligations may not be consistent because the two parties have different and changing perceptions.

As a result of the complex nature of psychological contracts, a diverse range of contract elements have been listed and measured in the literature (Thomas & Anderson, 1998; Kickul & Lester, 2001; Guest & Conway, 2002; Thompson & Bunderson, 2003). A comprehensive review of the various elements listed in the literature (Krivokapic-Skoko, Ivers & O’Neill, 2006) tried to differentiate the contractual elements into different types or sub groups. Thus, employee responsibilities can be categorised into four groups: (a) organisational citizenship behaviours; (b) basic obligations; (c) work environment; and, (d) loyalty. These four categories specified the behaviours and responsibilities that employees were prepared to be accountable for in return for the employer upholding what their employees believe to be their obligations. Employers’ responsibilities can be classified into six categories: (a) payment/ benefits; (b) management; (c) work environment; (d) fairness; (e) empowerment; and, (f) personal needs. These six categories covered the payments and benefits that employers were obligated to provide to their employees, the way in which the organisation was managed, and again the day-to-day work environment within the organisation. Further, employees considered that employers were obligated to ensure that their employees were empowered, treated fairly, and that their employee’s personal needs were addressed.

A violation of the psychological contract occurs when an employee experiences a discrepancy between the actual fulfilment of obligations by the organisation and the promise perceived regarding these obligations (Rousseau, 1995). Morrison and Robinson (1997) suggest that for contract violation to occur, the individual must perceive an imbalance in the exchange relationship and have an affective response to the perceived breach. It has been argued that violation of the contract leads to feelings of anger, betrayal, and resentment which in turn lead to decreased employee motivation, organisational commitment, loyalty and trust, as well as increased staff turnover (Rousseau, 1995).

The empirical research on psychological contracts has documented the direct effects of contract violations on employee behaviour. More specifically, studies have indicated that violation of employee’s beliefs and perceptions of the elements of psychological contracts may influence work outcomes, including job satisfaction, participation in development activities, and intention to remain with the current employer (Cavanaugh & Noe 1999; Freese & Schalk 1996; Dabos & Rousseau, 2004; Coyle-Shapiro & Conway, 2005; Turnley & Feldman, 1999; Del Campo, 2007; Nadin & Cassell, 2007). While empirical research on psychological contracts has developed significantly during the past decade, empirical research on psychological contracts within academia has been very limited. It is represented by the studies of Dabos and Rousseau (2004), Newton (2002), and the work at a New Zealand university initiated in the middle 1990s (Tipples and Krivokapic-Skoko, 1997). Research on the psychological contracts established by scientists/knowledge workers (O’Donohue et al., 2007) can be also discussed within a relatively broadly defined subject area of academia.
Dabos and Rousseau (2004) examined mutuality and reciprocity in psychological contracts by surveying the academics employed by a leading research-oriented school of bioscience in Latin America. Employees and their employers demonstrated convergence in their perception regarding the terms of their psychological contracts. This mutual understanding of the obligations resulted in positive outcomes for both researchers (career advancement and promotion) and the employers (increased research productivity). This is only one of the few empirical studies on psychological contracts which tried to expand beyond the research focusing on the downside of psychological contracts (such as violations, low morale, high turnover) to investigate the positive side of mutually beneficial contracts. In a slightly different context Newton (2002) used the concept of psychological contracts to discuss collegiality, professional accountability, reciprocity and mutual trust at a UK college of higher education. Based on the in-depth empirical research, the author argued that a lack of reward and recognition for academic work as perceived by academics can be explained also by not taking into account the existence of psychological contracts. At the same time, knowledge about the contents and dynamics of the academics’ psychological contracts may be very instrumental in maintaining staff morale and commitment.

Similarly, the empirical research done by Tipples and Krivokapic-Skoko (1997) at Lincoln University, New Zealand, indicated that the psychological contracts at that institution were in a very poor state. In terms of the empirical assessment the authors used a number of different approaches to explore the individual psychological contracts at the University. Besides qualitative interviews and the use of documentary sources, the authors conducted a questionnaire survey of academic colleagues to explore the staff members’ beliefs and expectations about their relations with the University. While analyzing only the employee’s side - academics at the University – the research identified low morale and disappointment amongst the academics. Generally, Lincoln academics were not satisfied with the extent to which the University had met what were perceived as its promised obligations. That dissatisfaction was consequently associated with a low level of job satisfaction.

O’Donohue et al. (2007) examined whether or not psychological contracts adequately reflect the knowledge worker’s contracts. Their findings indicated that scientists and knowledge workers were concerned more about ideological/societal concepts (scientific contributions and knowledge accumulation within the organisation) within their work than the transactional or relational psychological contracts established with their organisation. Also, the need for the knowledge workers to contribute a piece of work/knowledge is inherent along with the expectation that the organisation will reciprocate equally. Thus, continuous contribution to knowledge along with public access and the furthering of Australia’s knowledge base is vitally important to these professionals, forming core elements of their psychological contracts.

3. A mixed method research design

According to Morse (2003, p. 190) mixed method design can be defined as ”the incorporation of various qualitative and quantitative strategies within a single project”. Similarly, Creswell et al (2003, p. 212) define a mixed method research design at its simplest level as mixing both qualitative and quantitative methods of data collection and analysis in a single study. They further suggest that a more elaborate definition would include that a mixed method study involves “the collection or analysis of both quantitative and/or qualitative date in a single study in which the data are collected concurrently or sequentially, are given a priority, and involve the integration of the data at one or more stages in the process of research” (Ibidem, p. 212).

Generally speaking, motivation for combining qualitative and quantitative is to seek ‘elaboration, enhancement, illustration, clarification of the results from one method with the results from the other method’ (Green, Caracelli & Graham, 1989: 258-259). Following Creswell (1994) there are three ways of combining qualitative and quantitative research methods: dominant-less dominant design, two-phase design and mixed method design which would signify the highest extent of integrating methods. A sequential exploratory strategy is characterised by an initial phase of qualitative data collection and analysis, which is followed by a phase of quantitative data collection and the findings are then integrated during the interpretation phase. As Creswell et al. (2003, p. 216) noted this strategy is useful to explore the phenomenon, and particularly to expand on the qualitative findings. It is also useful when the researcher has to develop a new instrument such as in this case a survey of the psychological contracts established by academics.

Although mixed method designs can leverage the strengths of the both major streams and enrich the fields of organisational behaviour, strategic and human resource management, they were somewhat neglected within these disciplines. As Scandura and Williams (2000) and Currall and Towler (2003) noted increased triangulation might result in a more robust set of findings and a higher external validity of the management
research. The empirical research on the psychological contract is dominated by one type of study- the cross-sectional questionnaire survey (Conway & Briner, 2005). Even more, Taylor and Teklab (2004: 279) argued that because of the dominance of the surveys ‘psychological contract research has fallen into a methodological rut’. As a result, some experts on the psychological contract literature have recently strongly indicated the need to use a variety of research techniques (Conway & Briner, 2005), a holistic approach (Pete, 2006) and a triangulation of research methods in order to provide more convincing and reliable results of empirical research (Tipples & Verry, 2006).

4. Empirical results

Using Morgan’s (2006) Priority - Sequence Model this research may be categorised as a research design in which a preliminary qualitative study provides (in this case focus groups) the basis for developing the content of the questionnaire (Figure 1). The qualitative methods provide some exploratory work to ensure that the survey covers the important topics and also to provide valuable insights in how to address the issues of the violation of the psychological contracts in this particular context. However, in a slightly different way in this research both QUAL and QUAN methods receive relatively equal priority.

Figure 1: A sequential exploratory design used to analyse psychological contracts established by academics

The empirical data presented in this paper were generated through applying a sequential mixed method research design within the academics employed by the Faculty of Business, Charles Sturt University, NSW. Charles Sturt University is one of the largest non-metropolitan universities in Australia and it is also recognised as a leader in the provision of distance education. The CSU delivers nearly 500 courses to round 35,000 students. Faculty of Business employs more than 200 full time academics teaching and researching in the field of management, economics, finance, marketing, accounting, computing and IT.

Focus group methodology with a semi-structured format was chosen as it is known to be useful in the identification of issues and themes that can subsequently be drawn upon to assist with development of relevant survey questions (Saunders, Lewis & Thornhill, 2003; Wolff, Knodel & Sittitrai, 1993).

4.1 Focus groups

By their nature, psychological contracts vary significantly across organisations and even across different sections or units of the same organisation. Hence, the literature suggests (Turnley & Feldman 1999; Freese & Schalk 1996; Guzzo & Noonan 1994) that greater use of idiographic methods to assess individuals’ psychological contracts would be appropriate in order to access and understand the varied individual experience of the psychological contract. The focus group technique is a method through which dense subjective experience and interpretations can be addressed and discussed. The group interaction provides safety and generates synergistic effects, so that responses can be far more revealing than those obtained from individual interviews.

Focus groups are often used for concept screening and refinement, as well as to inform and develop the questionnaire. The main feature of focus groups is that they provide for relatively free-flowing and spontaneous discussions which can yield information and insights that would otherwise be unattainable. The focus group technique may be very appropriate for use in organisations as a means to deepen understanding of how employees interpret human resource practices. It may provide a way of accessing dense subjective interpretations in which the views of the individual academics are intertwined with the shared perceptions of the group.

The three focus groups were conducted in a conference room familiar to the academics on their home campus. A semi-structured format was developed to guide the focus group discussion. Potential participants were contacted by email and were provided with an information sheet explaining the aims of the study and the function of the focus group in the broader research design. The focus group discussions were moderated by a research assistant experienced with group facilitation so as to allow for an informal atmosphere and to minimise the power imbalance between the facilitator and other participants. While a limitation of focus groups can be the tendency for participants to deviate from their usual thinking and behaviour in order to “fit-
in’ with focus group norms (Kenyon 2004), the researchers were fortunate to be working with a relatively homogenous group of participants who regularly work with each other thereby minimising the effects of this tendency. The size of the groups was small enough to allow for in-depth discussion of key items yet offered diversity and possibilities for the development of ideas. Twenty six academics (excluding the researchers) participated across the three focus groups. Confidentiality can also be an issue affecting the willingness of participants to speak openly in focus group discussions (Bloor et al., 2001). To minimise the risk of loss of privacy, and perceptions of fear associated with focus group participation first names or pseudonyms were used during the discussion as, with the signed consent of the participants, each of the focus group sessions was audio taped and transcribed verbatim.

Within each focus group, the facilitator sought to elicit insights and subjective interpretations of the psychological contracts and the consequences of perceived breach. The focus group questions encouraged the academics to discuss: (a) what they feel they bring to their work that is not explicitly stated in their employment contract; (b) what they believe the university owes them in return; (c) how the University has fulfilled or exceeded expectations; (d) how the University has failed to fulfil expectations; and, (e) responses to perceived psychological contract violation.

The academics spoke to a range of personal qualities as elements of what they bring to their work and the university, frequently noting that their work involves much more than time and effort, including their creativity, integrity, values and experience. Contractual elements presented in earlier studies (see, for example: Thomas and Anderson, 1998; Kickul and Lester, 2001; Guest and Conway, 2002; Thompson and Bunderson, 2003) were exceeded by the elements named by the academics. The breadth of perceived responsibilities and obligations saw them speak to elements that extend well beyond the university.

The empirical research identified four key foci of academic responsibility that greatly influenced the formation and effects of their psychological contracts and these were:

a) the university
b) the discipline
c) society
d) students

Academics join the university with a strong work ethic and this is evidenced by many comments relating to a willingness to work outside ‘normal’ working hours, to be flexible in taking on various roles and to engage emotionally with their work:

‘...a willingness to work beyond the stated hours and a willingness to take on Faculty and University roles that are not sustained in one’s duty statement and that aren’t remunerated.’

The academics feel that the breadth of knowledge they bring to their work is an important contribution to the University. It was consistently stated that disciplinary knowledge, teaching and industry knowledge and experience, and industry contacts and networks, are highly valuable, but are not equally recognised by management. Conscience, personal ethics, integrity and a desire to make society a better place were strong motivators for staff and represented commonly discussed aspects of personal qualities that staff felt they were bringing to their academic work. Motivation and enthusiasm were frequently discussed in terms of ‘making a difference’, ‘making society a better place’, and generally expressing a desire to advance social justice and ethics.

These responses demonstrate that it would be limiting to attempt to understand the formation of the psychological contract only in terms of what the academic feels they owe the university. The commitment and concerns of academics are often directed more toward the students and society with the institution providing a means of serving those higher goals. If they are frustrated with unmet expectations and promises, it is likely that these frustrations will occur in areas that impinge upon their ability to fulfil their personal mission of attaining these higher goals.

Building upon perceived promises of mutual exchange the academics spoke at length regarding what they were expecting of the University in return for what they bring to their job. A common theme that emerged from the statements is that academics want to be recognised and treated as professionals. Much of the discussion centred on the expectations of leadership, fairness and transparency in promotion and recognition of one’s personal commitment to the profession, the university and the students.

www.ejbrm.com 65
Beyond the more tangible benefits that would normally be associated with employer responsibilities, employees expect good leadership and sound management skills. Issues related to leadership such as trust, clear and honest communication, transparency, advocacy, individual consideration and respect were prominent throughout the conversations. Generally, there was a realistic acceptance of the constraints within which management must make decisions, and that such constraints can lead to broken promises and failure to meet expectations from staff. What was not accepted, and this raised considerable emotion, was failure to address such situations in an honest manner and communicate outcomes effectively:

'Part of the transparency is the explanation for decisions that are made, clear justification and reasons why the decision was made rather than 'this is the decision' and nothing else.'

Commitment to teaching and the desire to contribute to society provide powerful motivators for academic staff and the need for academic freedom and job discretion were linked to these motivations. Staff expressed a strong expectation of autonomy, job discretion and inclusion in decision making and this was related to their professional identity:

'There's an expectation that our professionalism will be respected, that we're not going to be treated as if we've got nothing to add and that we're just automatons in the machine'

'Fairness in all things' was an expectation consistently expressed by the academics, which included: equitable pay, impartiality, fairness in promotion, consistency in applying rules, acceptance of union involvement, reciprocity, and an expectation that family and outside commitments should not cause disadvantage.

The academics also perceived their role as being equally involved in teaching, researching and administration and expect to be rewarded accordingly. This expectation of recognition for effort and achievement goes beyond the desire for a fair promotion and remuneration system, and addresses a basic need to be affirmed, appreciated and acknowledged by others:

'Recognition and acknowledgement particularly when you go beyond …the normal call of duty which I think we do frequently'.

Key areas where the University was considered to have fulfilled or exceeded its implicit promises of employment included support in such areas as research, outside activities, training and development and with regard to personal and emotional issues. While the support was appreciated, staff recognised that it was a reciprocal relationship:

'I think it’s a recognition that they are willing to do something for you to help you out, that you will pay them back [agreement from group] tenfold down the track . . . it makes it sound like an exchange relationship but still I think it is more than just that '

Although many examples of where the university had fulfilled or exceeded expectations were reported, it was obvious that this was not the complete picture. Even the groups who spoke more positively about their psychological contracts had much that they wanted to speak about with regard to when these contracts had been violated. The most striking consistency across the three focus groups of academics carried out for this research was the unprompted repetition of the phrase ‘changing the goalposts’ at each of the focus groups. There were also many references to dysfunctional aspects of the organisational culture such as: competitiveness, bureaucratic centralised control, short-term focus, and lack of customer (i.e. student) focus.

'We have talked about who are our customers and who we are building relationships with. I have seen [the university] do this and once again I expect it happens at other institutions that the student are not the main focus and I think it’s a pity.'

Administrative rules and regulations constituted one of the two key issues that were at the heart of most of the reports of psychological contract violation. Many academics perceived an encroachment of administrative systems stressing compliance, conformity, rationality and efficiency upon their practice as academic professionals who require flexibility, personal discretion and autonomy.

More broadly, the negative effects of the psychological contract violation were shown to be mediated by the nature of the academic work that involved a commitment to the students even when frustration with the institution was high:

‘there is that third dimension which plays a huge part in [the] psychological contract with the students. . . our responsibility and caring for the students that locks us into that contract …’

The most frequently cited responses to psychological contract violation were loss of loyalty and neglect behaviours. Some said that the decreased loyalty was resulting in their ‘giving up’ and feeling helpless. Others referred to behaviour that saw them less likely to engage in extra-role behaviour:
‘You concentrate more on your own interests instead of the broader interests than you have in the past.’

Increased neglect, particularly decreased attention to teaching quality, was a prominent topic of discussion:

‘It goes back to equity theory of motivation . . . You’ll do one of two things. You’ll either withdraw your labour totally . . . or you will slow down …’

However, for some academics the violation event gave them impetus to adapt to the new system and even enjoyment of the opportunities it offered. These adaptations to the new priorities and demands of the University showed that internal and external catalysts during an organisational restructure lead to renegotiations in which the contract evolves. The adaptation response was also related to the professionalism of the academic in that when loyalty to the institution was slipping, loyalty to the discipline and the commitment to students seems to take effect:

‘. . . very few academics slacken off because of their commitment to the students and because of their professionalism [agreement from group] so it doesn’t matter how badly they’re treated, they will still perform close to their optimal level and if they can’t do this they then leave.’

4.2 Survey

In the second phase of this research we administrated a survey of the full time academics employed by the Faculty of Business, Charles Sturt University. A standard Dillman approach was used to contact the target population and in the end we achieved a response rate of 50.0% with the final sample size of 117 academics. As this is a pilot investigative study, no non-response bias testing was completed. The items used in the factor analysis were generated from the psychological contract literature and the focus groups. Existing items were adopted from Janssens, Sels and Van den Barnde (2003), de Vos, Buyens and Schalk (2003) and somewhat altered to reflect the university context of the research.

In total, 31 items were included to measure University obligations, while 13 were included to measure the obligations of the academic to the university. Exploratory factor analysis (EFA) and principle component analysis (PCA) were used to analyse the data (Hair et al., 2006; Malhotra et al., 2002). Two factor analysis tests were conducted. The first was on the universities obligations and the second on the staff member’s obligation to the university. Existing theoretical frameworks, cumulative explanation, screen plot and eigenvalues all indicated that an eight factor solution for the University’s obligations and a three factor solution for the academics were appropriate.

Hair et al. (2006) argued that for the item to be significant in a sample size of 60 that the factor loadings need to be over 0.70. In some cases in both the University obligations and the academics’ obligations there were cases where factor loadings were marginally below 0.70. However, due to the exploratory nature of this research the slightly lower loadings are not considered an issue.

The first factor analysis was conducted to determine underlying constructs regarding the University’s obligation to the academics. Eight underlying factors were found and are displayed in Table 1.

The first factor analysis was deemed appropriate. The KMO was 0.627 above the general accepted minimum of 0.60 (Hair et al., 2006). The Bartlett’s test was also significant at 0.00, considering that there was only 1.9 cases per variable in the Bartlett’s test was appropriate. There were also some correlations in the correlation matrix that were greater than 0.30. As can be seen in Table 1 there is one variable which did not load significantly into a particular factor.

4.2.1 Factor 1: Fair treatment in promotion

This factor summarised the universities obligations to promotion. The two highest loading variables explicitly noted promotion, while the other two variables focused on the workplaces treatment and acknowledgement of the academics.

4.2.2 Factor 2: Staff development and support

The items in this factor were similar to that of factor one. However, the emphasis of this factor is on the work environment, specifically, career and professional development.
4.2.3 Factor 3: Good management and leadership

This factor directly refers to leadership and management with the new concern of growing bureaucracy.

4.2.4 Factor 4: Academic life

The fourth factor contained the traditional areas of academic research life. With the exception of the item that regards communication, the other items all indicate research and lifestyle balance.

Table 1: University’s obligations to the academics: Exploratory factor analysis

<table>
<thead>
<tr>
<th>Items</th>
<th>Loading</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 2 3 4 5 6 7 8</td>
<td></td>
</tr>
<tr>
<td>Provide clear and consistent requirements for promotion</td>
<td>0.85</td>
</tr>
<tr>
<td>Treat you fairly and equitably with regards to promotion</td>
<td>0.82</td>
</tr>
<tr>
<td>Be fair and equitable in its treatment of academics</td>
<td>0.74</td>
</tr>
<tr>
<td>Include you in decision-making that directly affects you</td>
<td></td>
</tr>
<tr>
<td>Provide opportunities for career development</td>
<td>0.76</td>
</tr>
<tr>
<td>Support ongoing professional development</td>
<td>0.73</td>
</tr>
<tr>
<td>Provide opportunities promotion</td>
<td>0.69</td>
</tr>
<tr>
<td>Provide remuneration that is comparable to other universities</td>
<td>0.61</td>
</tr>
<tr>
<td>Provide a safe and comfortable work environment</td>
<td>0.52</td>
</tr>
<tr>
<td>Ensure that staff act collegially</td>
<td>0.51</td>
</tr>
<tr>
<td>Provide good management</td>
<td>0.76</td>
</tr>
<tr>
<td>Provide good leadership</td>
<td>0.75</td>
</tr>
<tr>
<td>Minimise the impact of red tape</td>
<td>0.72</td>
</tr>
<tr>
<td>Provide security of ongoing employment</td>
<td>0.68</td>
</tr>
<tr>
<td>Allow you autonomy to act as a professional academic</td>
<td>0.66</td>
</tr>
<tr>
<td>Maintain academic freedom</td>
<td>0.62</td>
</tr>
<tr>
<td>Respect the demands of family/personal relationships</td>
<td>0.62</td>
</tr>
<tr>
<td>Communicate important information to you</td>
<td>0.56</td>
</tr>
<tr>
<td>Acknowledge the long hours you devote to work</td>
<td>0.83</td>
</tr>
<tr>
<td>Act ethically</td>
<td>0.79</td>
</tr>
<tr>
<td>Manage the pace of change so that it does not adversely affect you</td>
<td>0.57</td>
</tr>
<tr>
<td>Provide remuneration that is similar to the private sector</td>
<td>0.84</td>
</tr>
<tr>
<td>Provide remuneration that is similar to the public sector</td>
<td>0.81</td>
</tr>
<tr>
<td>Recognise your non-university experience</td>
<td>0.56</td>
</tr>
<tr>
<td>Reward excellence in teaching through the promotion system</td>
<td>0.79</td>
</tr>
<tr>
<td>Reward excellence in research through the promotion system</td>
<td>0.62</td>
</tr>
<tr>
<td>Reward excellence in admin/management through the promotion system</td>
<td>0.53</td>
</tr>
<tr>
<td>Be honest in its communications with you</td>
<td></td>
</tr>
<tr>
<td>Offer flexibility regarding working from home</td>
<td>0.86</td>
</tr>
<tr>
<td>Respect the role of academic unions in the workplace</td>
<td>0.62</td>
</tr>
</tbody>
</table>

4.2.5 Factor 5: Fairness and equity

In factor five the common theme is fair treatment. Workplace ethics and respect for the work that academics complete are the key areas.

4.2.6 Factor 6: Appropriate remuneration

This factor group mainly refers to comparable levels of pay to private and the public sector.
4.2.7 Factor 7: Reward performance
This factor groups the themes of rewarding excellence in teaching, researching and administration.

4.2.8 Factor 8: Good workplace relations
The final factor contained the items related to workplace relations, in particularly the role of unions and the right to work at home.

The second factor analysis was conducted to determine underlying constructs regarding an individual's obligation to the university (Table 2). Three underlying factors were found and are referred to as meet 'academic expectations', 'commitment' and finally 'above and beyond'. Again the factor analysis was satisfactory, with a KMO of 0.765 and a Bartlett’s of 0.00. The eigenvalues indicated that a three factor model was present. Equally, the variance explained was 58%, whilst generally 60% is required, which means the figure is near enough to suggest a three factor model. Some of the colorations in the coloration matrix were also greater than 0.30. Hence, factor analysis is deemed appropriate for use.

Table 2: Academics’ obligations to the university: Exploratory factor analysis

<table>
<thead>
<tr>
<th>Item</th>
<th>Loading</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comply with university rules and regulations</td>
<td>0.74</td>
</tr>
<tr>
<td>Act ethically at work</td>
<td>0.65</td>
</tr>
<tr>
<td>Advance your discipline</td>
<td>0.62</td>
</tr>
<tr>
<td>Publish scholarly research</td>
<td>0.58</td>
</tr>
<tr>
<td>Work effectively and efficiently</td>
<td>0.57</td>
</tr>
<tr>
<td>Stay employed by the university for the next 2 years</td>
<td>0.77</td>
</tr>
<tr>
<td>Travel for work</td>
<td>0.73</td>
</tr>
<tr>
<td>Act collegially</td>
<td>0.61</td>
</tr>
<tr>
<td>Work long hours to complete tasks</td>
<td>0.52</td>
</tr>
<tr>
<td>Complete tasks that are not strictly part of your job</td>
<td>0.78</td>
</tr>
<tr>
<td>Complete tasks that are asked of you</td>
<td>0.66</td>
</tr>
<tr>
<td>Provide teaching quality</td>
<td>0.52</td>
</tr>
<tr>
<td>Enhance student development</td>
<td>0.58</td>
</tr>
</tbody>
</table>

One issue with the second factor analysis was the number of cross loading variables. Again, this may be attributed to the small sample size. Another issues was that some of the commonalities were below the desired 0.60 (Hair et al., 2006).

4.2.9 Factor 1: Meet Academic Expectations
The first factor includes aspects that are regarded as the typical view of an academics job description.

4.2.10 Factor 2: Commitment
The second factor involves more obligations that are part of an academic job. Basically, this factor looked more at commitment to the job. The key items included a two year commitment to the job, travel and a collegial attitude.

4.2.11 Factor 3: Above and beyond
The third factor focused on the students, with teaching quality and student development. This area is beyond the normal expectations covered particularly in factor one.

5. Concluding comments
This research focuses on exploring the formation and content of psychological contracts established by the academics within an Australian University. The psychological contracts are considered here as the perceived exchange relationship that exists between employee and employer. The qualitative phase of the research provided in-depth and rich information about academics’ views and understanding of the psychological contracts. The focus groups provided a way of accessing dense subjective interpretations in which the views of the individual academics were intertwined with the shared perceptions of the group. The survey and factor
analysis were used sequentially to unpack and cluster different types of the academics’ perception of the psychological contracts and the responses to the violation of the contracts.

Our empirical research has shown that the professional aspects of commitment to making a contribution to society, their discipline, and student learning frequently play a prominent part in the development, and moderation of the academics’ psychological contracts. The academics very strongly indicated that they have a professional responsibility and spoke to a significant social role which effectively extends beyond the boundaries of the psychological contracts they establish with the university. It is critical for the University and the academics to be sensitive to possible differences in expectations, since unrealised expectations may result in de-motivation, decreased commitment, increased turnover, and loss of trust in the organisation. These contracts motivate employees to fulfill commitments made to employers when they are confident that employers will reciprocate and fulfill their side of the contracts.

Using exploratory factor analysis (EFA) of the survey data, eight factors were identified in relation to the University’s obligations to the academics and three underlying factors were found in relation to individual academic’s obligations to the University. In terms of the University’s obligation to the academics the EFA reinforce the concepts of leadership and management, fairness and equity, notably in relation to promotion and provision of opportunities for career development. The three underlying factors referred to as meet ‘academic expectations’, ‘commitment’ and finally ‘above and beyond’ were primary constructs explaining academics’ obligations to the University. Since this is still research in progress, the development and understanding of these underlying factors will serve as a basis for further investigation analysis using cluster analysis and structural equation modelling. Obviously, some of the weaknesses of the questionnaire surveys impact the validity of the findings. The survey was based on the respondents from a single organisation and used self-reporting questionnaires to assess the variables which were framed in terms of promises and obligations. Since this research has focused only on employees’ perspectives it was not able to capture the exchange process involved in the psychological contract. Further research should take into account both the University’s and academics’ perceptions in framing psychological contracts.

References


Millennial Students and Technology Choices for Information Searching

Martin Rich
Cass Business School, London, UK
M.G.Rich@city.ac.uk

Abstract: This paper draws together ideas about different generations of students, notably the 'millennial generation' (born from around 1982 to 2000) which encompasses a high proportion of current students in higher education, and ideas about the different types of technology available when searching for information. In the context of higher education, this is particularly relevant when students are encouraged to find out information for themselves, typically to relate this to taught material. This is connected with information literacy, as it reflects students' abilities to carry out simple or complex research. This paper focuses on why students choose particular technologies to support their research and the effect of these choices on their learning and on their written work.

A particular current issue is the emergence of the generation of Internet resources collectively known as 'web 2.0' - notably Blogs and Wikis - and the relationship of these to the way that students presently in universities favour structuring their work. These resources also introduce issues of authoritativeness. It is tempting to dismiss wikipedia as the work of amateurs, but where a blog has been created by a notable expert, author, or journalist, questions arise as to whether the blog should be regarded as of different value from the same person's written work. Therefore some consideration will be given to how students can be encouraged to recognise and draw on intelligent exploitation of these new resources.

Underlying data for the paper is drawn from discussions with current students, both individually and in groups.

Keywords: Millennial students, Web 2.0, information literacy

1. Introduction

In recent years there have been tangible changes in students' favoured techniques for searching for information. One driver for this is ready access to the Internet, but the nature of material available on the Internet is evolving, and with it the sort of strategies that students can use to navigate this material. The work here particularly focuses on the preferences associated with those born since 1982, characterised by Strauss and Howe (1997) as 'millennials' but it should be noted that other classifications of ages are also used. For example the British Library and University College London have identified the 'Google generation' as of particular interest – being those born since 1993 who have no recollection of life without the Internet (CIBER, 2008). At present this generation is too young to attend university – in the British educational system the oldest of these will currently (2008) be studying for their GCSE (general certificate of secondary education) examinations. However some of these will be starting on undergraduate courses in another three years, which is very much within the planning horizon for universities.

Not all commentators view the emergence of this generation in a positive light. Notably Keen (2007) takes an apocalyptic view in which the Internet is seen as a destroyer of more traditional literacies. However the growth of the Internet has been accompanied by the emergence of new ways of organising and classifying data, notably the use of 'tags' and 'folksonomies' as ways to impose a structure on information held on the Internet. These are typically associated with web resources that include a significant volume of content generated by users, and provide an approach to classifying and navigating that information which contrasts with that associated with more traditional taxonomies (Lambe, 2007; Golder and Huberman, 2006). Arguably it is a misnomer to associate this trend with the 'Google generation', as this approach has emerged more recently than the widespread use of Google. But members of the millennial generation are acquiring the sort of adult information navigation skills in an environment where folksonomies are widely used, and this potentially raises challenges for educators in universities as to how to promote information literacy. A particular aspect of these challenges is the channels, or technologies, that students choose in searching for information.

2. Participation and user-generated content

The term web 2.0 typically refers to a range of Internet applications, characterised by contributions from a large number of participants. Its purpose has been described as ‘harnessing collective intelligence’ (O'Reilly, 2006), and is most familiar to many Internet users through Blogs and through wikis – particularly Wikipedia. It is very easy for somebody to set up a blog or to contribute to Wikipedia, and Blogs in particular are
associated with tools that allow anybody with the time and interest to generate content to produce something that looks very professional. Like any written material, web 2.0 content may be good, or may be inaccurate, or it may present a view that is extremely partisan. As web 2.0 has become a commonplace component in many users’ experiences of the Internet, it has attracted interest among educators and in business. Tredinnick (2006) draws attention to the potential for this approach to be used within an organisation to build an Intranet, used for sharing information and knowledge between various stakeholders: in a university this would cover knowledge being built collaboratively by staff and students. Alexander (2006) discusses a range of interesting ways in which web 2.0 concepts can be applied in higher education, illustrated by some useful specific examples. Some of these go beyond the everyday uses of web 2.0 resources and require active participation and sharing by students – for instance the use of shared searches which both students and faculty could follow over a number of weeks (Alexander, 2006: 40).

This participative model also has the potential to encourage learner-centred and more informal approaches to learning – for example the extent to which students become ‘non-formal’ learners where they assimilate knowledge informally in parallel with attending classes (Eraut, 2000). Much teaching and learning within universities is influenced by the concept of reflection-in-action (Schön, 1983) and participative approaches to Internet use have the potential to stimulate this, for example through one contributor raising an original, and possibly contentious, point that can prompt a reflective discussion among others. Furthermore, web 2.0 offers an environment where people can be seen to engage in reflection and to produce useful, and evolving, content as a result. This has the benefit of transparency, of it being possible to see how knowledge is built up – a point addressed by Clark (2008) who robustly criticises academics who forbid students from using Wikipedia, and instead suggests that students be encouraged to use the ‘discussion’ tab on Wikipedia pages. Smith (2001) notes, among criticisms of Schön’s concepts, the observation that they do not necessarily contribute to praxis – to informed, committed, action. Web 2.0 tools do offer the potential for groups of people working together to work towards action. Incidentally it should be noted that Smith’s work cited here is an example of a thoughtful and scholarly resource that is freely available on the Internet, and that is continually revised and expanded.

The very ubiquity of Wikipedia offers evidence that students are prepared to use, and trust, web 2.0 content. However the indications so far are that many of them – even those of the millennial generation who have grown up with the Internet and who are used to participating in social networking sites such as Myspace and Facebook – have a limited awareness of the origins of material that they use in their studies. In other words Wikipedia is just another web site carrying information – and perhaps because it is comprehensive, easy to search, and easy to find it is one that is tempting to use. Students are very ready to use Wikipedia, but much less willing to engage with it. This lack of engagement is not necessarily a bad thing; notably the concept of communities of practice (groups of people united by a very specific common interest, often working in different places but linked by the Internet) depends on valuing peripheral participants as well as active contributors (Wenger et al, 2002). Web 2.0 supplements user-generated content with a series of tools, such as RSS feeds, that facilitate navigation of knowledge and, for instance, to track contributions that an individual might make to a discussion. At present many students remain reluctant to use these tools effectively, and will only draw on material that is presented in a simple form on a web page.

In theory, RSS feeds, by alerting users of the Internet to changes in pages which might be particularly relevant to their interest, match one particular characteristic that is frequently associated with the millennial generation. This is the effect termed by Stone (2008) as ‘continuous partial attention’ – the ability to work on several activities at once and to distribute attention between them. Davenport and Beck (2001) identify attention as an important commodity within organisations, and one of the challenges in using web 2.0 in learning is to understand how material can be structured to attract students’ attention.

3. A classification of web 2.0 competences

This paper is intended to be exploratory in nature, and uses student experiences observed by, and reported to, the author to determine a classification of different students’ abilities to use Web 2.0 resources. The intention is to use this classification as a framework for conducting further inquiry in more depth in the future. It became apparent at an early stage of discussion that web 2.0 competencies could be divided into two broad categories:

- **Necessary** competences are essential for students to use web 2.0 resources, and students who lack these competences run the risk of producing work which is ill-informed or positively misleading.
Supplementary competences allow students to explore web 2.0 resources in more depth and potentially to contribute to such resources and to the construction of a body of knowledge.

The students whose work informed these concepts were predominantly undergraduates, and were all studying business or management related subjects. The subject area is relevant because of the range of topics that it encompasses – a reason that it is tempting for students to set off and search the entire Internet for sources, rather than to look at specific journals, or specialised websites.

Two key necessary competences are proposed:

- **Accurate searching** implies the ability to use a range of search tools on the Internet, to formulate searches accurately (notably to understand how and when to exclude terms from a search), how to extract the most useful and relevant items from a long list of search results. The skills implied here are relevant when using specialised databases, for example of articles from business journals, as well as when using Google or other popular search engines. They include judging search results by context and type as well as by keywords, so an article which originated in the *Harvard Business Review* carries more weight than an unpublished piece on a student website.

- **Judging authoritativeness** is much more important in the web 2.0 context than when using the traditional web, or when using printed materials. At a simple level, students can be asked to say why a particular source is authoritative if they are to cite it in their work. But there are also several dimensions to material being authoritative. If it does deal with a contentious subject, then students need to acknowledge any bias, and possibly to draw on two or more sources with contrasting views. They need to distinguish between cases where the author of an item declares that they are authoritative (do they take the reasons why at face value, or can they investigate them in some way) and cases where somebody independent vouches for the material being authoritative. They need to recognise that a range of types of publication exists on the web (magazine, scholarly journal, and so on) and that some, but by no means all, of these mirror types also exist on paper. In the context of the web they need to recognise that the number of other sites that link to a page is some measure of the destination page’s value: there is a parallel here with the use of the number of citations as a measure of the value of an scholarly paper. All these facets of this competence depend on an ability to take a critical viewpoint, and a lack of criticality among students can lead to material being used inappropriately.

Drawing on the issues raised above, supplementary competences allow students to draw on resources in more depth, and also to engage actively in building knowledge. Three supplementary competences are proposed here, but it is recognised that they may usefully be subdivided into others in the light of further inquiry, and also that there are potential overlaps. These focus on understanding the complexity of knowledge.

- **Use of a variety of channels** implies being able to use (for example) RSS to monitor information that is of particular interest, and being able to synthesise information from different resources. A prerequisite for this is understanding the different types of content and knowing how this can be build up to provide something useful for a particular student’s needs. The emphasis here should be on understanding the relationship between different resources, and in seeing the potential to generate new and useful material by combining different sources. This addresses a concern that has been observed with undergraduate students in particular: when searching for relevant literature they often look in vain for one book or paper which covers the exact subject about which they want to find out, but they are reluctant to look for several different sources which cover overlapping issues which together are relevant to their work.

- **Structural understanding** implies an interest and understanding in the intricacies of how web 2.0 content is formed. Potentially there is some overlap with the previous competence, but here the emphasis is on how the knowledge has been built up over time, and on the contribution that different people have made to it. In the case of resources structured as wikis (not just Wikipedia itself), an example of this would be a readiness to look behind the ‘page history’ and ‘discussion’ tabs on a page and to draw inferences about the page content.

- **Positive engagement** implies participation in the construction of knowledge through web 2.0. This could be done though students creating their own resources (but these need to be ones that actively invite discussion, such as Blogs which welcome comments), or more typically through students becoming active participants in a discussion. Prerequisites to this include understanding the tacit and explicit rules of behaviour that prevail in an electronic forum: Wikipedia is a particularly instructive example of this because it has evolved a complex set of
rules and conventions for contributors, and being able to identify whether other contributors are knowledgeable, willing to collaborate, and so on.

4. Encouraging these competences

These competences can be viewed as an emerging set of research skills which students will need to learn in the future. One characteristic of web 2.0 is that the tools available can be expected to change over time, so students need to learn skills which are not restricted to a particular generation of software applications. It is difficult to be prescriptive about how educators can best encourage these competences, especially the supplementary ones. The necessary competences can be incorporated into the assessment criteria for students, through for instance an instruction only to use Internet resources where students can make a strong case for the resource being relevant and authoritative. They can be strengthened by educators themselves encouraging and monitoring good practice, and also by inviting students to check their own work against certain standards.

Encouraging the supplementary competences is a greater challenge. Preliminary evidence suggests that students are extremely reluctant to reach the stage of positive engagement, and even if they are active participants in web discussions in connection with their hobbies or social life, they are reluctant to become similarly active in contributing to knowledge. It is not unrealistic to expect every student to be comfortable with all of these competences, which would be the equivalent of expecting every student in a large formal lecture to be comfortable with posing questions from the floor. Also, just as there is a concern in classes where participation is encouraged, that the most vocal students dominate the interaction and therefore gain the most from it, the use of web 2.0 raises a concern that it might privilege a small group of students with very highly developed competences. It is reasonable to expect that positive engagement, in particular, will be easier for students when a critical mass has been reached of students willing to participate in this way. One possibility for educators is to focus on particular areas where students can usefully build up a critical mass of useful knowledge that includes user-generated content, and where educators can comment on the validity or otherwise of the knowledge to guide student contributors in a useful direction. Finally it should be noted (returning to generational issues) that this is a very different approach to building knowledge from that practised by most academics, and it is essential that academics are prepared to learn, if necessary, from the experience of others who have contributed effectively to web 2.0 resources.

References

Keen A (2007): The cult of the amateur. London, Nicholas Brealey
White Researcher-Black Subjects: Exploring the Challenges of Researching the Marginalised and ‘Invisible’

Gisela Schulte Agyeman
City Psychological Service Kingston upon Hull UK
Gisela.Agyeman@hullcc.gov.uk

Abstract: How to access the life world of the Other without mis-representing it, has long been a concern in a number of research traditions, including disability, feminist and black research traditions. Feminist debates have addressed issues such as whether researchers can only speak on behalf of their own social groups or whether there is legitimacy in speaking on behalf of the Other. Justification for this often arises from the fact that dis-empowered groups may not have access to a public forum where they can be heard, so that social researchers see their role as ‘giving a voice’ to those who would otherwise not be heard. By contrast, other researchers see commonalities with their respondents as a particular strength.

The paper focuses on theoretical perspectives which provided the framework in a study of the lived experiences and expressed views of second-generation female young people of African and Caribbean heritage in predominantly white community and school settings. It examines some of the theoretical as well as personal considerations of a white female researcher and educational practitioner whose personal position and interests had to be negotiated within the research process. The paper highlights some of the difficulties, dilemmas and challenges of a white researcher attempting to access the lived realities of ethnic minority children whose lives are ‘invisible’ in dominant discourses. Drawing on theoretical perspectives from the feminist, black and disability literature, the paper explores the implications for researching ‘sensitive’ topics in organisational contexts from the perspective of an outsider ‘looking in’ and argues for a conscious ‘positioning’ of the researcher who may or may not be part of the lifeworld of individuals or groups being researched.

Keywords: Lifeworld, marginalised individuals, ‘invisible’, dominant discourses, black/white perspectives

1. Introduction

Only little attention has been paid to the experiences of minority ethnic pupils in geographical areas in which there may be no peer reference group or access to community cultural experiences, other than the dominant white culture. The almost exclusive research focus in the literature on minority ethnic pupils in multiethnic urban areas (Mac an Ghaill, 1988, 1991; Verma and Pumfrey, 1988;) has tended to highlight issues around racism, teacher expectations, academic underachievement and identity at the expense of examining the experiences of ethnic minority children at the margins, whose lives and experiences remain ‘invisible’. The paucity of research in this area relates to the fact that such pupils’ experiences do not lend themselves easily to be quantified. Within the context of their particular school, their exam results are not reflected statistically and their educational performance cannot be measured reliably in terms of their group. (Cline, de Abreu, Fihosy, Gray, Lambert and Neale, 2002). The experience of being on the margins and not ‘worthy of study’, represents a strange paradox for ethnic minority children in this context. On the one hand, their physical characteristics such as skin colour, hair and features make them highly visible within their predominantly white social settings, yet in terms of representation and importance they are often ‘invisible’.

The present paper explores issues in relation to researching second-generation teenage girls of African and Caribbean descent, several of whom shared the commonality of mixed parentage (having a white and a black parent). Whilst a wealth of literature in the field has focused on the experiences and underachievements of ‘African-Caribbean’ boys in the British education system (e.g. Coard, 1971; Mac an Ghaill, 1988; Sewell, 1998;) or on the experiences of South Asian children (e.g. Ahmed, 2000; Bhatti, 1999), and some studies have focused on the experiences of black girls (Mirza, 1992); there are few studies of girls of mixed African/white heritage and who live in predominantly white communities.

Furthermore, much of the literature has focused on ethnic minority children as falling short of the norms for their white counterparts, e.g. in relation to schooling. This perspective has ‘pathologised’ the experience of ‘black’ childhood through Western models and practices of psychology and child development, social work and education (Owusu-Bempah and Howitt, 2000; 2002).

Only a small number of studies have focused on the positive aspects of being different and how positive role models can be a source of inspiration (e.g. Hoyles and Hoyles, 1999).
This paper explores the writer’s theoretical considerations and personal reflections in a study of the ‘Voices’ of young black females (aged between 15 and 21 years) who were in full time education at the time of interviews, with regards to their experiences of education and life within predominantly white communities. Themes which initially emerged from a taped group discussion, were followed up by individual conversations with young ethnic minority females and discussions with first generation, black African parents whose children attend schools and further education in predominantly white area.

Accessing the Life Worlds of under-represented and marginalised groups presents a significant challenge for researchers who are not part of these life worlds and who want to be credible in their representation of the Other. As a white female researcher of European descent with a strong interest in the lived experiences of young ethnic minority females in predominantly white communities, this has meant addressing my own role in the research process. It has involved a search to find a research framework within which it is credible to re-present individuals and groups who are on the margins or invisible.

The first section of this paper examines wider contextual challenges to dominant discourses which have contributed to the marginalisation of under-represented groups. The second section examines the scope of research and the role of the researcher as an outsider. Finally, the implications of researching sensitive topics in the context of organisations will be addressed.

2. Challenging dominant discourses

A number of black academic writers and researchers have expressed dissatisfaction with white dominant discourses of black pupils, e.g. in relation to educational underachievement (e.g. Callender, 1997; Channer, 1995; Mirza, 1992).

The ‘deficit model’ of Western thinking in relation to black and other ethnic minority groups and their children is also highlighted by Owusu-Bempah and Howitt (2000) who argue that this has found legitimisation in disciplines such as Psychology, Psychiatry and Social Work. The authors argue that simplistic and Eurocentric paradigms have been applied to other ethnic groups without taking into account social and cultural factors.

Owusu-Bempah and Howitt (2002) distinguish between levels of understanding at the user level, service level, organization level, culture level and value level. Their criticism of Psychology as a discipline which encourages agents of society to change individuals to fit the dominant culture needs to be taken seriously.

The issue of professionals pathologising culturally different behaviours or declaring it as deviant affects teachers, psychologists, social workers, psychiatrists and many more. The outcomes are experienced at many levels by those affected by institutionally racist thinking, practices and policies.

Ahmed (2000) also points to the failings of mainstream Psychology in engaging with issues of race and racism, often falling into the trap of dualistic concepts (e.g. individual vs. society ) and or ignoring wider socio-economic and cultural factors, thus reducing the racism to the level of social exchanges and interactions at the micro-level.

Hoyles and Hoyles (1999) stress the achievements of black/mixed race people throughout history which are not usually known and seldom find their way into the written history books or school lessons (e.g. the Jamaican born Mary Seacole who never made it into the history lessons whilst Florence Nightingale is known to everyone in Britain). Not bringing black and other ethnic histories into the school curriculum, is not simply an omission but, I would argue, another way of making ‘invisible’ the culture and contributions of black people. At the same time dominant discourses highlight the ‘problematic’ (and often implied ‘inherent nature’) of black people.

Ahmed (2000) also points to the failings of mainstream Psychology in engaging with issues of race and racism, often falling into the trap of dualistic concepts (e.g. individual vs. society ) and or ignoring wider socio-economic and cultural factors, thus reducing the racism to the level of social exchanges and interactions at the micro-level.

For Asante (1998) all discourse is culturally centred, including dominant theories, such as phenomenology. He argues, writing from an African American perspective, that discourses of blacks to a large extent remain “essentially white or Eurocentric discourse by black people.” (p.185).

My argument is that by making people, cultures and other traditions ‘invisible’, we are trivialising, marginalising their experiences and contributions. At the individual level this means that lived experience is negated, whilst for communities this means being at the margins of society. My intended research led me to examine the work of Ahmed (2000) who looked at a variety of discourses based on research interviews with second generation Bangladeshi living in Britain, arguing that discourses perpetuate oppressive social
relations of the dominant culture. For me, this further highlights the need to make the voice of the minority group central to the research.

Livia (1996) shows how the absence in literature and thought of certain groups of women (e.g. black lesbians or Jewish lesbians) has led to a position that makes even the start of any conversation or discussion difficult. She states:

*When I started writing I was aware of the feminist tenet that the writer is accountable for what she creates, but it did not occur to me that I might also be accountable for what I left out.* (Ibid, p.35)

In this context, challenging oppression therefore has to include challenging omissions, which refers to my argument of having to challenge the dominant discourse which makes ‘invisible’ individuals, communities or ideas.

Reinharz (1992) states that contemporary feminist ethnography has three goals which influence the feminist researcher:

- Documenting the lives and activities of women
- Understanding the experiences of women from their point of view
- Conceptualising women’s behaviour as an expression of social contexts.

In this context, listening to the ‘Voice’ of the Other plays an important role, as demonstrated in the work of Gilligan (1982) who shares the concern with other feminist researchers that psychology and dominant discourses have frequently used male behaviour as a yardstick against which to measure female behaviour. This has in the past led to female behaviour being seen as ‘deviant’ from standards of psychological expectations based on male ‘norms’. The critiques of dominant ideologies and discourses voiced by feminist researchers and writers share similarities with views expressed by black researchers and academics discussed earlier (e.g. Owusu-Bempah and Howitt, 2000) regarding the ‘pathologising’ of black experience in a white-dominated society. Feminist debates have frequently addressed issues such as whether researchers can only speak on behalf of their own social group or whether there is legitimacy in speaking on behalf of the Other. Justification for this often arises from the fact that disempowered groups often do not have access to a public forum where they can be heard, so that social researchers see their role in ‘giving a voice’ to those who would otherwise not be heard (e.g. Bell, 1996; Russell, 1996), a view which I share and drew on in my own research. There are, however, difficult issues involved when researchers take it upon themselves to represent Others, not least around potential dangers of mis-representation, mis-use of power and abuse. A useful distinction here is pointed out by Wilkinson and Kitzinger (1996) which concerns the notions of ‘re-presenting’ and ‘representing’, a distinction originally made by Marx (*Darstellung*/*Vertretung*). The notion of *re-presenting* is about giving a voice to marginalised individuals and groups who are often under-represented, whilst *representing* has a patronising quality of standing in the place of and speaking on behalf of the Other.

Some writers (e.g. Aziz, 1997) have cautioned against white feminism, meaning from a white perspective (not based on colour or ethnicity) which upholds existing power structures. For Aziz (1997) racism often obscures the fact that whiteness and blackness are intertwined in the experience of oppression. However, whilst white women experience oppression as patriarchal, black women experience it as racist and patriarchal.

Mama (1995) stresses the importance of understanding discourses on the Other (black women) as being embedded in historical conditions of slavery, colonialism and racism as well as the discourses and practices which these bring about. She argues that individual consciousness arises “out of the resonances between collective history and personal experience” (p.163).

Contextualising individual experience therefore means, going beyond current day economic, social and cultural contexts and extending this into political and historical contexts. Psychology has in the past provided insufficient theoretical grounding for looking at individuals beyond their current contexts and taking into account collective historical positions. Even interest in the relatively ‘new’ areas of cultural and cross-cultural perspectives in Psychology (e.g. Gardiner and Kosmitzki, 2002; Squire, 2000) do not offer a framework of analysis that takes political and collective historical positions into account. Whilst theories Psychology have concerned the behaviour and development of the individual, these studies have traditionally been based on relative small numbers of individuals, often not representative of non-white or Western populations. Dissatisfaction with the discipline from marginalised groups is therefore understandable.
Another theoretical perspective which influenced the conceptualisation of my research came from the disability literature which emphasises that individuals with disabilities belong to a group of disempowered people within society (Oliver, 1990, 1996; Gleeson, 1997). With this they share a common perspective with other marginalised groups, such as women, blacks and other ethnic minorities. The question arises whether a ‘non-disabled’ person can ever research issues relating to disability without being patronising and whether they have the right to speak on behalf of a group that has been silenced by the dominant political paradigm (Drake, 1997). In arguing that the political dimension is about challenging oppression and exclusion (Goodley, 1997; Oliver, 1990) and by accepting the social model of disability (Oliver, 1990, 1996), ‘non-disabled’ writers can join in highlighting oppression. If the experience of the Other cannot be ever experienced, this poses a difficulty in representing any such experience without distorting it. In the disability research literature this has found expression in studying the Voices (which express experiences) of those concerned (e.g. Leicester and Lovell, 1997).

3. Exploring and positioning my own role in the research process

My interest in researching the Life worlds of young ethnic minority females in predominantly white areas was a personal as well as professional one. On the one hand, I am a white female who is married to a black West African and we have lived in a predominantly white community in North East Lincolnshire. Our three children have attended local primary schools where they have, in some years, been the only non-white child in their year group, if not the entire school. They have experienced, at times, open hostility and racism such as name calling or being excluded from games because ‘your skin is brown’. However, as long as racism and hostility are out in the open, they can be addressed. As parents we have been able to talk about how prejudice works, about strategies of dealing with difficulties on an individual level and about talking to teachers and schools. At the institutional level, schools are required to log racist incidents and although many under-represent incidents because ‘minor’ incidents such as name-calling are not reported or recorded, there is a system through which data can be collected and compiled. In contrast to this, most of the life experiences our children have had cannot be quantified in terms of numbers and statistics and to do so, I would argue, would be to diminish the reality of their lived experiences. One of the realities I have touched on earlier, is that of being ‘invisible’ in terms of representation and impact.

On a professional level, my work as an Educational Psychologist in the City of Hull in the North of England, brings me into contact with many vulnerable pupils such as those with special educational needs and disabilities. Over centuries, Kingston upon Hull has been made up of a mainly white population with only few black or ethnic minority people living in the area who were not part of clearly defined ‘communities’ such as those in other cities.

Although the City has seen a strong influx of a range of minority groups in recent years, such as Congolese refugee families and Eastern European economic migrant families, the pupils from these groups remain concentrated in some schools, whilst other schools still remain predominantly white. Several schools I work with have very few ethnic minority pupils on their roll. Even in schools with relatively high representations of minority ethnic pupils, the teaching and support staff teams remain almost exclusively white. Furthermore, outside agencies working with schools, are almost entirely white in staffing. On a professional level I am aware that many of the issues affecting young ethnic minority people in predominantly white settings, such as the lack of a peer reference group, sense of isolation and not being reflected in dominant discourses, are similar to those affecting my own children. As my eldest daughter grew into a young adult, I was increasingly able to have open and critical discussions about what her experiences of growing up as a young black female (my daughter refers to herself as ‘black’ in a political sense) in a predominantly white area have been. Many of the discussions were initiated by my daughter and I often wished I could have taped or recorded her views to share with other young people in a similar position and also professionals working in schools and educational establishments.

The decision to research the topic area further did not come easy as I had to grapple with many issues, such as my own emotions and involvement as well as whether I could ever do justice to the topic since I am not black myself or am part of an under-represented group. Being white represents being part of the dominant discourse and in the words of my own daughter on a day she experienced a racist physical attack: “Mum, you have no idea and you can never know what it is like to be black or different!” Furthermore, being a psychologist is being part of a profession that has historically contributed to prejudice and oppression by applying a white Western (supposedly objective) yardstick to measure all human ability and performance. This is a reductionist approach, at best neglectful at worst a shameful example of how professions have abused power and ‘abnormalised’ the experiences of others. There were many questions I asked myself:
How could I possibly be ‘objective’ when I was emotionally involved? How could I possibly hope to represent a group of young people with whom I only shared in common our gender? How could I be credible if I had never experienced or even come close to experiencing their life worlds?

Blair (1998) queries whether ‘neutrality’ in educational research is ever possible, as all research is based on prior positions which in turn influence knowledge construction. Pretending otherwise, would in fact be misleading. Blair (ibid) states that whilst feminist research has to some extent succeeded in representing the marginalised through theoretical perspectives and more inclusive research methodologies, research on ‘race’ and education has been more of an uphill struggle.

Feminists have for a long time argued that the personal is political and that the purpose of ‘giving voice’ is essentially about effecting social change (e.g. Wilkinson and Kitzinger, 1996).

Bell (1996) writes about her own research experience as a white, female researcher which highlighted incidences of rape in Aboriginal communities. Russell (1996) examines her personal research within a white feminist paradigm, whilst conducting research on black women in South Africa; whilst Jeffreys (1996) examines whether and how feminists can theorise prostitution if they have never worked as prostitutes. Other researchers see commonalities with their respondents as a particular strength. Mirza (1992) points out that like her interviewees, she had originated from the West Indies, was young and therefore enjoyed a “unique insight into the experiences of the young black women in the study.” Mirza points out, that any subjective bias should be acknowledged, so that interest positions are openly known.

Callender (1997) also states her conviction that being a black female teacher and researcher gave her greater insight into the topic area of black pupils within the education system. However, Callender also acknowledges that having a similar ethnic background can have disadvantages. Whilst it can “open doors to meanings normally reserved for ingroup members” (p.11), it can also close doors as “the researcher becomes too immersed in the situation to see clearly what is happening.” (p.11). Mac an Ghaill (1991), reflecting on his own ethnographic research in a school which contributed to Young Gifted and Black, states that as a white male researcher investigating black female and male students, methodological and political considerations were of central importance in conceptualising his research. The importance of knowing the life world of participants as well as understanding the wider social and political contexts within which they are located is also stressed by Channer (1995). My reflections also drew on Barn (1994) who stresses the need for an anti-discriminatory research model in which issues of gender, race, ethnicity are acknowledged and consciously confronted. Research also builds on those who have ‘gone before’ and entered new territories in a sensitive way. In this context Edwards (1993) was useful as she showed that writing from a feminist perspective as a white academic who interviewed black women was possible. She states that factors such as race, class sex differences and similarities of researcher have implications for the research process. They enable both researcher and researched to ‘place’ each other within the social structure and can increase or lessen the sensitivity of research topics.

In the context of my own research, I had intended to use a methodology which enabled me to explore the ‘lived experience’ of individuals in a way which accepted as valid their experiences, thoughts and interpretations and meaning making. Interpretative phenomenological analysis (IPA) initially seemed to fulfil my requirements as it is set within an inductive research paradigm which builds on the accounts given by participants (e.g. Reid et al 2005; Smith 1996, 1999; Smith et al 1999; Willig, 2001). However, the approach also requires researchers to reduce the complexity of experiential data through processes of rigorous and systematic analysis. This means that greater levels of abstraction are achieved through increasing levels of analysis but it also means that each level of analysis is further removed from the original data (the accounts and words given by the participants themselves). My methodological concern here is that, with each level of analysis there is increasing scope for omissions, distortions and mis-representations of the kind critiqued by researchers concerned with re-presenting marginalised groups, as discussed above. Taking a methodologically ‘purist’ stance has the potential of being distorting of the depth of experiences of the participants and the complexities of their social realities. IPA opens up the potential to access lived realities but it is also an attempt to reduce and quantify ‘data’. It could be argued that it represents an attempt to justify an ideographic approach in the light of potential criticism from the positivist quarters. Given these considerations, I eventually adopted an approach to analysis used by Channer (1995) in her research of black people’s life histories. It is based on themes which emerge from the descriptions and recounted experiences of the participants. They are not imposed themes but reflect the emphasis given by participants to areas of their lives.
Channer (1995) describes her rationale for the less formal and structured type of data collection through conversations below:

In conversations, ideas flow from one notion to another. This fluid nature of gathering data inevitably results in the overlapping of the themes. The need for rigorous analysis of the data is not disputed; however, restricting the evaluation of the material by using an unsuitable mechanistic approach would fail to identify significant issues fully. (Channer, 1995; p. 66)

This also reflects a methodological approach taken by Mama (1995) in her research on black women’s identities. Her discussions often took the form of conversations which were initiated by Mama by explaining her interest in the topic area in general terms. Often a small tape recorder was left running, which discussants rapidly became less conscious of as in the case of my group discussion.

4. Conclusion

Sensitive research should be a journey of discovery in which the researcher becomes ‘sensitised’ to the potential challenges and dilemmas that their chosen topic may hold. When researching the Other in the role of an outsider, this also means addressing the role of self in research and engaging in critical questioning of one’s own role and scope. Using ethnographic methods to access the life worlds of others and enabling their voices to be heard, is one way of trying to address the issue of power imbalance in the research process. The issue therefore is not whether the researcher or writer is disabled, black or female themselves but rather whether they are writing from that perspective. This dimension also takes into account the intentions and purpose of research and acknowledges that all research of Others is essentially a political process which the researcher needs to be critically aware of. From this perspective it is a requirement that the researcher consciously positions herself/himself before the practical research commences and makes this position clear to those they wish to research.

Furthermore researchers in organisations need to be sensitive not to exclude those already on the margins by ‘omission’ because of a perceived need to use purist research methodologies which may not always be appropriate to the subject matter.

References


Channer (1995) describes her rationale for the less formal and structured type of data collection through conversations below:

In conversations, ideas flow from one notion to another. This fluid nature of gathering data inevitably results in the overlapping of the themes. The need for rigorous analysis of the data is not disputed; however, restricting the evaluation of the material by using an unsuitable mechanistic approach would fail to identify significant issues fully. (Channer, 1995; p. 66)

This also reflects a methodological approach taken by Mama (1995) in her research on black women’s identities. Her discussions often took the form of conversations which were initiated by Mama by explaining her interest in the topic area in general terms. Often a small tape recorder was left running, which discussants rapidly became less conscious of as in the case of my group discussion.

4. Conclusion

Sensitive research should be a journey of discovery in which the researcher becomes ‘sensitised’ to the potential challenges and dilemmas that their chosen topic may hold. When researching the Other in the role of an outsider, this also means addressing the role of self in research and engaging in critical questioning of one’s own role and scope. Using ethnographic methods to access the life worlds of others and enabling their voices to be heard, is one way of trying to address the issue of power imbalance in the research process. The issue therefore is not whether the researcher or writer is disabled, black or female themselves but rather whether they are writing from that perspective. This dimension also takes into account the intentions and purpose of research and acknowledges that all research of Others is essentially a political process which the researcher needs to be critically aware of. From this perspective it is a requirement that the researcher consciously positions herself/himself before the practical research commences and makes this position clear to those they wish to research.

Furthermore researchers in organisations need to be sensitive not to exclude those already on the margins by ‘omission’ because of a perceived need to use purist research methodologies which may not always be appropriate to the subject matter.

References


Issues and Challenges in the Use of Template Analysis: Two Comparative Case Studies from the Field.

Teresa Waring¹ and David Wainwright²
¹Newcastle Business School, UK
²Northumbria University, Newcastle-upon-Tyne, UK
Teresa.waring@northumbria.ac.uk
David.wainwright@northumbria.ac.uk

Abstract: One of the most problematic issues for researchers who conduct qualitative research using semi-structured, unstructured interviews or story telling data collection methods is the analysis of large quantities of rich data. In the past this has often led to fairly unmethodical approaches to analysis which in turn has led to qualitative business and management research being seen as insubstantial and unworthy of consideration.

A relatively recent development in organisational research has been the application of Template Analysis to rich unstructured qualitative data following the primary data collection phase. Template Analysis appears to have emerged from the USA during the 1990s and academics familiar with the Grounded Theory approach to data analysis may see similarities in the techniques used. Nevertheless, it has gained credibility in the UK through the work of Nigel King and other colleagues researching in health and sociology related fields.

This paper provides an overview of the origins of Template Analysis and discusses how it has been used to structure qualitative data. It then goes on to examine through the two case studies how Template Analysis has been extended and used by the authors in two different research projects. In the first case study the research team worked within a Primary Care Trust in the North East of England on a project that explored the Diffusion of Innovation of clinical and administrative computer systems across General Practice within the Trust. Seventeen Trust members were interviewed for approximately one hour and this led to over 85000 words of rich data. The second project focused on the NHS Secondary Care sector and examined IT project management practice related to the development of integrated pathology computing systems across eight separate laboratories in the North of England. Eight senior managers were interviewed and the resulting 3 years of document collection, also resulted in a large volume of rich textual material. The use of template analysis, combined with a critical success factors methodology, resulted in a novel approach for learning about current IT project management practices.

This paper critically examines these two case studies in terms of their particular research philosophy, epistemological approach and the lessons learnt from the techniques employed. The paper then provides a discussion of the principles and practicalities of template analysis and explores the benefits to the business and management research community at large.

Keywords: template analysis, qualitative, NHS, interview, information

1. Introduction

It can be seen increasingly in Business and Management research a growth in the use of qualitative methodologies and data collection methods. This often results in large volumes of textual material that must be analysed and interpreted. Text data may include field notes from participant observation, transcripts from semi-structured interviews, diaries and stories or narratives. Just as the researcher is required to construct an overall research design it is also necessary to develop an analytical strategy within the interpretive process so that the research is seen as substantial and worthy of consideration.

A relatively recent development in organisational research has been the application of Template Analysis to rich unstructured qualitative data following the primary data collection phase. It is the aim of this paper to consider the use of Template Analysis within the context of business and management qualitative research and illustrate it with reference to two Information and Communications Technology (ICT) research case studies. Both of these case studies used a more innovative approach than is suggested by Crabtree and Miller (1999) and King (2004) but still remain within the accepted Template Analysis framework. In the first case study the research team worked within a Primary Care Trust in the North East of England on a project that explored the Diffusion of Innovation of clinical and administrative computer systems across General Practice within the Trust. Seventeen Trust members were interviewed for approximately one hour and this led to over 85000 words of rich data. The second project focused on the NHS Secondary Care sector and examined IT project management practice related to the development of integrated pathology computing systems across eight separate laboratories in the North of England. Eight senior managers were interviewed and this, combined with participant observation and over 3 years of document collection, also resulted in a...
large volume of rich textual material. The use of template analysis, combined with a critical success factors methodology, resulted in a novel approach for learning about current IT project management practices.

This paper critically examines these two case studies in terms of their particular research philosophy, epistemological approach and the lessons learnt from the techniques employed. The paper then provides a discussion of the principles and practicalities of template analysis and explores the benefits to the business and management research community at large.

2. Origins of template analysis

Template analysis is a relatively recent development and appears to have emerged from more structured approaches such as Grounded Theory and Interpretative Phenomenological Analysis (IPA). King (2004) argues that although template analysis makes use of codes and coding of data it is not as prescriptive as Grounded Theory and is not wedded to its realist methodology. In fact it can be used within a range of epistemological positions and thus can be useful to a large number of researchers. However when template analysis is used within a broadly phenomenological approach it is very similar to IPA. The main difference between the two approaches are the use of ‘a priori’ codes in template analysis and the balance between within and across case analysis.

3. What is Template Analysis and how is it used?

A well accepted text on the early emergence of template analysis is that of Crabtree and Miller (1999). They go on to say:

“...researchers can develop codes only after some initial exploration of the data has taken place, using an immersion/ crystallisation or editing organising style. A common intermediate approach is when some initial codes are refined and modified during the analysis process.” (Miller and Crabtreee, 1999:167)

A more recent advocate of Template analysis is Nigel King (http://www.hud.ac.uk/hhs/research/templateanalysis/) whose work in this area is mainly based in healthcare similar to Crabtree and Miller (1999). The Template approach involves coding a large volume of text so that segments about an identified topic (the codes) can be assembled in one place to complete the interpretative process.

The complete analysis process of organising, connecting and corroborating/legitimizing involves:

- Creating a code manual/coding scheme
- Hand or computer coding the text
- Sorting segments to get all similar text in one place
- Reading the segments and making the connections that are subsequently corroborated and legitimised.

Another good source of understanding how codes are developed and then applied can be seen in King (2004) and Miles and Huberman (1994). King advocates one of three positions when starting out on the research:

- Have pre-define codes/ a priori codes based on the theoretical position of the research.

OR

- Develop codes after some initial exploration of the data.

OR

- Take a half way position – some initial codes ( possibly from the interview questions?) and refinement after exploration of the data. It may depend on your epistemological position.

Once having established a high level coding scheme King (2004:260) goes on to describe through the use of case material the coding hierarchies e.g.

1. CASE BACKGROUND HISTORY
2. THE CONSULTATION
3. SERVICE CONTACT
4. POSSIBLE AREAS OF IMPROVEMENT
Within “The Consultation” there are various hierarchies/levels

2. THE CONSULTATION
   2.1. Presenting problem
   2.2. Treatment/management offered
      2.2.1. Prescription
      2.2.2. Advice
      2.2.3. Referral
   2.3. Factors influencing treatment
      2.3.1. Patient/GP interpersonal relationship
      2.3.2. The GP role
         2.3.2.1. GP perception of role
         2.3.2.2. GP workload

Each of the higher levels can be developed in a similar manner to “The Consultation”. King (2004) also discusses parallel coding of segments of text whereby segments of text are classified within the same level. This is referred to later in this paper. In order to contextualise the template approach we will consider two different case studies in business and management research located within a healthcare setting. The intention is to provide insight into the application of the analytic approach.

4. Issues and challenges in case study one

During 2004/5 a Primary Care Trust in the North East of England took part in a pilot project to explore the diffusion of ICT innovation within their GP practices. This section of the paper focuses primarily upon the data collection aspect of this project and the analysis of the interviews conducted with stakeholders.

It is not our intention to provide a detailed discussion of the vast literature base that addresses the area of Diffusion of Innovation. Diffusion of innovation (DOI) research and practice originates from many diverse fields of study. These include sociology, anthropology, healthcare, medicine, social policy, psychology, strategic management, economics, marketing, entrepreneurship, organizational behaviour, research and development, and technology management. An innovation is not just an outcome but a process and is the effort to create purposeful, focused change in an enterprise’s economic or social potential (Drucker 1985). It may be viewed as something that is new to an adopting organization but not necessarily new in its own right. Nevertheless, if the reader wishes to explore this subject further they could look at the following authors Rogers (1995); in the area of Information and Communication Technology DOI - Caldeira and Ward 2003; Cragg 2002; Grandon and Pearson 2004; Poon and Swatman 1999; Southern and Tilley 2000; Mustonen-Ollila and Lyytinen (2003)).

Initially a pilot site was selected for the first trial run of the data collection. This enabled the research design to be reviewed, assessed, and refined prior to expanding it across a wider range of respondents and organizations. The study was conducted between January and September 2004 and aided by a small university research grant.

The method of data collection for this study was that of the semi-structured interview. At each of the five research sites, tape recorded interviews were carried out with a core set of staff: a general practitioner (doctor), practice manager, and practice nurse – fifteen interviews in total. The interviews were based around a broad set of topic headings to guide the interviews but we tried as far as possible to allow the participants to lead the discussion. Research consent forms and information leaflets were provided in advance of the study.

Interviews took place at a time and place convenient to the respondent and ideally were limited to a maximum of 1 hour duration. All data has been made anonymous and all confidences respected. Transcripts (verbatim) were presented back for verification by each respondent. Wherever possible, the two researchers were present to enable consistency of the data collection approach and to enable peer reflexivity at a later date in terms of the data analysis and also the research process itself.

Defining the codes and creating the initial template: Template analysis normally starts with some predefined codes intended to help guide analysis. The first issue we faced was how we addressed our extensive NHS and Information Systems Management (ISM) knowledge and how this should be presented in
the template form for NHS staff. Secondly we needed to consider how extensive the initial template should be. King (2004) suggests that if you start with too many predefined codes, then the template might blinker analysis and prevent exploration of more pertinent issues. On the other hand, too few codes may lead to an overwhelming mass of rich and complex data.

Our approach and starting point was the letter sent out to the GP practices introducing the research topic and some of the government and department of health concerns. From these we constructed a template in a graphical format (see Figure 1). This took a period of time and reflection to ensure we agreed on the initial codes. The template was used as a prompt for the respondent and adapted for use if necessary. The key words have been derived from the research literature, NHS documentation related to ICT strategy initiatives, and the personal research experience of the researchers. This approach has a proven record of successful use in qualitative research in healthcare settings specifically investigating "soft" organizational issues (Clarke and Wilcockson 2001, 2002).

**Figure 1:** Template for Key Issues Associated with ICT in English General Medical Practices.

**Revising the template:** It is important to realise that the template must be revised in response to the concerns of the interviewees. Respondents were able to suggest other key words that they thought should be included and enlarged on them if they felt it necessary. They were provided with the template at least 1 week prior to the interview. As part of the research we soon discovered that many of the key issues, words, and phrases that we identified initially were not relevant to all of the interviewees and a number could be deleted. It was also apparent that GPs had issues and perspectives different from nurses and practice managers even when they came from the same organization practice.

**The analysis of the data:** The template used in Figure 1 was developed to support data collection from individual respondents who were not conversant with the language of ICT diffusion. These people were clinical professionals and therefore in order to engage them in a productive dialogue the codes were developed from the researchers’ cumulative knowledge (combined 30+ years) in health and Information Systems research. An objective of the research was to develop a more informed framework of ICT diffusion for Innovations in healthcare and IS. This led us return to the literature to identify a more suitable academic/theoretical framework from which we could identify a priori codes for the data analysis.
Table 1:

<table>
<thead>
<tr>
<th>Innovation</th>
<th>Doctors</th>
<th>Nurses</th>
<th>Practice Managers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mobile technology</td>
<td>Internet</td>
<td>Paper-lite</td>
</tr>
<tr>
<td></td>
<td>New use of EMIS –Population Manager</td>
<td>Results reporting</td>
<td>Manager as IT manager</td>
</tr>
<tr>
<td></td>
<td>Results reporting</td>
<td>EMIS – Population Manager, templates</td>
<td>E-comms</td>
</tr>
<tr>
<td></td>
<td>Internet</td>
<td></td>
<td>EMIS Population Manager</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Results reporting</td>
</tr>
</tbody>
</table>


If we consider ‘Results reporting’ as an example it can be seen that it is an Innovation in all of the three categories of staff. However, within the coding what emerged was:

1. Results reporting
   1.1. Process
      1.1.1. Internal email
      1.1.2. Distribution of printouts
   1.2. Problems
      1.2.1. Technical
      1.2.2. Human/organisational

Results reporting is the process by which diagnostic tests are sent to a hospital laboratory for analysis and the results sent either by electronic or paper-based means back to the GP practice. Staff discussed how it worked for them and the difficulties encountered on a day to day basis. Thus when writing up the research it is much more easy to see how to structure the sections and support categories with illustrative text. It allows comparison within specific groups of staff and across groups.

5. Issues and challenges in case study two

Teespath was formed in the year 2000 in order to establish a ‘collaborative clinical partnership across the community, in order to support the Tees Valley clinical network’ with the main improvement goals focusing on: information systems, common equipment platforms, common approaches to systems of work, expansion of test repertoire and range of services and improved recruitment and retention of key staff (Clayton, 2000). In January 2002 Teespath was allocated £1.82 million from the NHS Phase III Pathology Modernisation funds towards the costs of supporting the development of a single managed clinical network in pathology (DoH, 2004).

The exploratory nature of this research, within the context of a complex suite of pathology related IT projects and a new organisational network design, guided the authors to adopt a qualitative and interpretivist approach to the inquiry. The study is still on-going but the main interviews were conducted in a 2 month period in 2003. Other primary data such as core documents, emails and notes from personal observation within project meetings have been collected over the period 2002 through 2007. One of the authors is employed full time as a Pathology laboratory manager at one of the participating hospital Trusts. Primary data was collected as part of a negotiated research project in collaboration with the Teespath managed pathology network project. This provided legitimate access, whilst conforming to all ethical requirements, to project meetings, minutes, formal and informal documents as well as his role as a participant observer.

The main focus of the primary data collection involved both unstructured and semi-structured interviews. In the first instance unstructured interviews were conducted with the Teespath programme director and IT manager in order to identify possible ‘critical factors’ deemed important by these key members of the group. The information collected was used to restructure the Somers and Nelson (2001) top ten Enterprise Resource Planning (ERP) Critical Success Factors (CSFs) to contextualise it for the pathology modernisation initiative. This formed the basis of a set of questions for the semi-structured interviews directed towards eight IT managers from every Trust within the Teespath pathology organisational network. Interviews were audio-recorded and fully transcribed for later analysis. In addition the interviewee’s were asked to comment on the CSF list and encouraged to comment on the validity of the factors in a NHS environment, and to suggest any factors they felt were absent.
In order to organise, analyse and interpret the text, template analysis was utilised to first code and then re-arrange the text from the semi-structured interviews. Crabtree and Miller (1999:165) state that 'when using a template, the researcher defines a template or codes and applies them to the data before proceeding to the connecting and corroborating/legitimating phases of the analysis process...where the template or codes can be constructed a priori, based on prior research or theoretical perspectives'. King (1998) advocates compiling a list of codes occurring in each transcript. The distribution of the codes within and across transcripts can help draw attention to aspects of the data, which warrant further examination. In this study an a priori list of codes was drawn up based on a key selection of CSF research from the field of ERP adoption (Somers and Nelson, 2001) and implementation research (see Table 1.) Interview transcripts were manually coded and analysed using coloured highlighters. The distribution within and across the template was facilitated by the use of a spreadsheet. The frequency and distribution of codes was used as a means of making ‘connections’ within the text. (Crabtree and Miller 1999:169)

The IT managers interviewed in the Teespath project were asked to re-assess and re-rank the top ten CSF’s (from Somers and Nelson) to reflect their importance to the ‘Teespath’ initiative. This was done in order to create a comparison with the original Somers and Nelson (2001) list of ERP critical success factors. It was envisaged that a list of critical factors relevant to NHS Information Systems implementation would be produced, advancing knowledge in this area. The results (averaged out) can be seen in Table 2.

Table 2: Comparison between Somers & Nelson’s (2001) list of ERP CSF’s compared to those generated for ‘Teespath’ development

<table>
<thead>
<tr>
<th>Critical Success Factor</th>
<th>ERP Ranking</th>
<th>NHS ranking</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top Management Support</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Clear Goals and Objectives</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Project Team Competence</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Project Team Management</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Project Champion</td>
<td>8</td>
<td>5</td>
</tr>
<tr>
<td>Interdepartmental Co-operation</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>Interdepartmental Communication</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>Management Expectations</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>Careful Package Selection</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>Vendor support</td>
<td>9</td>
<td>10</td>
</tr>
</tbody>
</table>

5.1 Template analysis

Each CSF was used as a category to identify further sub-codes within the text generated from the interview data, table 3. The sub-codes were derived by immersion/crystallisation (Crabtree and Miller, 1999) through reading the texts; this was a deductive leading to an inductive research approach with sub-codes emerging from the data and expressed in the particular discourse and language of the respondents. The frequency of statements pertaining to each sub-code and category were then interpreted and recorded by interviewee. The analysis presented in Table 4 demonstrates the density of comments related to each CSF and the particular contextualisation of the issue within the Teespath project.

The most frequent negative comments (most mentions) related to issues concerning project team competence (people leaving and joining at will; projects could be improved by taking wider organisational perspectives); inter-departmental co-operation (identification of cultural differences between pathology and other stakeholders); project team management (requirement to involve users, unstructured project management and no initial success and no delivery). Other strong views were expressed concerning top management support (no active support and lack of direction from the Board) and careful package selection (may conflict with the Care Records System (a national centralised IT health records or electronic patient record system)). Positive comments related to the vendor support (vendors have been supportive), clear goals and objectives (goals and objectives are clear).
Table 3: Template analysis, categories and codes (Mn = number of managers’ response)

<table>
<thead>
<tr>
<th>Codes</th>
<th>M1</th>
<th>M2</th>
<th>M3</th>
<th>M4</th>
<th>M5</th>
<th>M6</th>
<th>M7</th>
<th>M8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top management support</td>
<td>1</td>
<td>1.1</td>
<td></td>
<td></td>
<td>1.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No active support</td>
<td></td>
<td>1.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IT not high on agenda</td>
<td>1.12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nothing past IT manager</td>
<td>1.13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>North Durham opted out</td>
<td>1.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>1.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All Chief Executives are signed up</td>
<td>1.31</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Project Team Competence</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Problems with the initial make-up of teams</td>
<td>2.1</td>
<td>2.1</td>
<td>2.1</td>
<td>2.1</td>
<td>2.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>People leave and join at will</td>
<td>2.2</td>
<td>2.2</td>
<td>2.2</td>
<td>2.2</td>
<td>2.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Projects lack systemic review</td>
<td>2.3</td>
<td>2.3</td>
<td>2.3</td>
<td>2.3</td>
<td>2.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Project teams are competent</td>
<td>2.4</td>
<td>2.4</td>
<td>2.4</td>
<td>2.4</td>
<td>2.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Generally a good mix of staff on projects</td>
<td>2.5</td>
<td>2.5</td>
<td>2.5</td>
<td>2.5</td>
<td>2.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Could be improved by taking a wider organisational perspective</td>
<td>4.4</td>
<td>4.4</td>
<td>4.4</td>
<td>4.4</td>
<td>4.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Co-operation</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cultural differences within pathology</td>
<td>3.1</td>
<td>3.1</td>
<td>3.1</td>
<td>3.1</td>
<td>3.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lack of co-operation between laboratories</td>
<td>3.2</td>
<td>3.2</td>
<td>3.2</td>
<td>3.2</td>
<td>3.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lack of co-operation between pathology and IT</td>
<td>3.3</td>
<td>3.3</td>
<td>3.3</td>
<td>3.3</td>
<td>3.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clear goals and Objectives</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No clear goals and objectives</td>
<td>4.1</td>
<td>4.1</td>
<td>4.1</td>
<td>4.1</td>
<td>4.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Confusion over what the projects are trying to achieve</td>
<td>4.2</td>
<td>4.2</td>
<td>4.2</td>
<td>4.2</td>
<td>4.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Confusion with suppliers</td>
<td>4.3</td>
<td>4.3</td>
<td>4.3</td>
<td>4.3</td>
<td>4.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brief getting wider by the minute</td>
<td>4.4</td>
<td>4.4</td>
<td>4.4</td>
<td>4.4</td>
<td>4.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes the goals and objectives are clear</td>
<td>4.5</td>
<td>4.5</td>
<td>4.5</td>
<td>4.5</td>
<td>4.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Project Team Management</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Initially unstructured project management</td>
<td>5.1</td>
<td>5.1</td>
<td>5.1</td>
<td>5.1</td>
<td>5.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Initial lack of a project manager</td>
<td>5.2</td>
<td>5.2</td>
<td>5.2</td>
<td>5.2</td>
<td>5.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lack of documentation</td>
<td>5.3</td>
<td>5.3</td>
<td>5.3</td>
<td>5.3</td>
<td>5.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not too good: PRINCE 2</td>
<td>5.4</td>
<td>5.4</td>
<td>5.4</td>
<td>5.4</td>
<td>5.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Requirement to involve users</td>
<td>5.5</td>
<td>5.5</td>
<td>5.5</td>
<td>5.5</td>
<td>5.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No initial success not delivered anything</td>
<td>5.6</td>
<td>5.6</td>
<td>5.6</td>
<td>5.6</td>
<td>5.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improved with formal project management</td>
<td>5.7</td>
<td>5.7</td>
<td>5.7</td>
<td>5.7</td>
<td>5.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Well managed</td>
<td>5.8</td>
<td>5.8</td>
<td>5.8</td>
<td>5.8</td>
<td>5.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interdepartmental communication</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Involvement of people who need to know</td>
<td>6.1</td>
<td>6.1</td>
<td>6.1</td>
<td>6.1</td>
<td>6.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lack of communication to lab staff</td>
<td>6.2</td>
<td>6.2</td>
<td>6.2</td>
<td>6.2</td>
<td>6.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not considered a role of the project team</td>
<td>6.3</td>
<td>6.3</td>
<td>6.3</td>
<td>6.3</td>
<td>6.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good communication</td>
<td>6.4</td>
<td>6.4</td>
<td>6.4</td>
<td>6.4</td>
<td>6.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Staff gain interest as projects complete</td>
<td>6.5</td>
<td>6.5</td>
<td>6.5</td>
<td>6.5</td>
<td>6.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Management Expectations</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Modernise pathology</td>
<td>7.1</td>
<td>7.1</td>
<td>7.1</td>
<td>7.1</td>
<td>7.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>For financial benefit</td>
<td>7.2</td>
<td>7.2</td>
<td>7.2</td>
<td>7.2</td>
<td>7.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>As stand-alone projects</td>
<td>7.3</td>
<td>7.3</td>
<td>7.3</td>
<td>7.3</td>
<td>7.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not a money saving exercise</td>
<td>7.4</td>
<td>7.4</td>
<td>7.4</td>
<td>7.4</td>
<td>7.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reaching to government initiatives</td>
<td>7.5</td>
<td>7.5</td>
<td>7.5</td>
<td>7.5</td>
<td>7.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Project Champion</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Initially Director of Teess Health Authority</td>
<td>8.1</td>
<td>8.1</td>
<td>8.1</td>
<td>8.1</td>
<td>8.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Procured</td>
<td>8.1</td>
<td>8.1</td>
<td>8.1</td>
<td>8.1</td>
<td>8.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IT manager</td>
<td>8.1</td>
<td>8.1</td>
<td>8.1</td>
<td>8.1</td>
<td>8.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No one person</td>
<td>8.2</td>
<td>8.2</td>
<td>8.2</td>
<td>8.2</td>
<td>8.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vendor support</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vendors have been supportive</td>
<td>9.1</td>
<td>9.1</td>
<td>9.1</td>
<td>9.1</td>
<td>9.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Careful package selection</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extremely careful</td>
<td>10.1</td>
<td>10.1</td>
<td>10.1</td>
<td>10.1</td>
<td>10.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Required demonstration of operability</td>
<td>10.2</td>
<td>10.2</td>
<td>10.2</td>
<td>10.2</td>
<td>10.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chose the most proactive supplier</td>
<td>10.3</td>
<td>10.3</td>
<td>10.3</td>
<td>10.3</td>
<td>10.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cost was a major factor</td>
<td>10.4</td>
<td>10.4</td>
<td>10.4</td>
<td>10.4</td>
<td>10.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not a project the traditional IM&amp;T suppliers would be interested in</td>
<td>10.5</td>
<td>10.5</td>
<td>10.5</td>
<td>10.5</td>
<td>10.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>May conflict with ICRS</td>
<td>10.6</td>
<td>10.6</td>
<td>10.6</td>
<td>10.6</td>
<td>10.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>There is no one of the bid option</td>
<td>10.7</td>
<td>10.7</td>
<td>10.7</td>
<td>10.7</td>
<td>10.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not actually bought anything</td>
<td>10.8</td>
<td>10.8</td>
<td>10.8</td>
<td>10.8</td>
<td>10.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Doubts over lab to lab project</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Question the whole raison d’etre of the lab to lab project</td>
<td>11.1</td>
<td>11.1</td>
<td>11.1</td>
<td>11.1</td>
<td>11.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lack of resources</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No revenue to back capital</td>
<td>12.1</td>
<td>12.1</td>
<td>12.1</td>
<td>12.1</td>
<td>12.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difficulty getting funds released</td>
<td>12.2</td>
<td>12.2</td>
<td>12.2</td>
<td>12.2</td>
<td>12.2</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
6. Discussion

Template analysis is now well embedded in healthcare qualitative research (King, 2004; Crabtree and Miller, 1999). However, it is not so well established in Business and Management research and thus is innovative yet challenging in itself when applied to this different context. Traditionally Business and Management research has emerged from a positivist paradigm and only more recently has begun to accept that interpretivist qualitative research has theoretical as opposed to anecdotal value. Nevertheless, it has been hampered by poor data analysis techniques and a lack of cumulative theory building. Our approach attempts to develop a more rigorous approach to analysis of large amounts of rich textual data but has been adapted from the methods advocated by King (2004) and Crabtree and Miller (1999). We could be challenged that Template Analysis bears little difference from the use of software packages such as NVivo for analysis of data and in fact the software might allow a more comprehensive approach. We would argue that immersion in the data is an essential part of the interpretive process and use of technology can often act as a substantial barrier.

Figure 2 is a diagrammatic representation of our work to date and demonstrates two alternative routes that can be adopted when undertaking template analysis in this context. Case 1 began with the researchers initially considering a grounded theory approach to this project. It soon became obvious that such an approach was unsuitable due to the existing knowledge, experience and biases of the two researchers within the domain of Healthcare ISM research. The research design was thus modified to retain the inductive component but with elements of structure provided by a template analytic approach. The researchers with reference to relevant literature and practice engaged in a contextual dialogue to identify a set of a priori codes that could be used to engage the specific health professional identified in the research project. Data was then collected via semi-structured interviews using the initial codes (Figure 1).

In contrast Case 2 began with a process of comprehensive literature review with an explicit aim of identifying a list of suitable critical success factors for IT project management. The research was designed to assess these CSFs for relevance within a healthcare ICT management project. The literature review was successful in identifying the top ten CSFs which had a major impact on the success of large scale ERP project implementations. This selected list was therefore used as the a priori codes to structure the interview schedule and inform the data collection process. This was a deductive process leading to a very structured...
data collection and analysis exercise. The template analysis was very informative in identifying frequency and relationships between individual codes – these are often referred to as parallel codes (Crabtree and Miller, 1999). Through an inductive process of immersion in the text a hierarchical list of sub-codes emerged to provide a much richer interpretation of the IT project management process (Table 2).

It may be seen as a critique of this work that deductive and inductive research is incompatible within the same research project. However Business and Management research can be critiqued for its lack of emphasis on theory generation and development. Many researchers who do use a deductive approach very often find that there is still much unknown about certain theoretical aspects. One area of our research that has highlighted this is the emergence of power and politics within the DOI research we conducted yet the frameworks within the literature do not acknowledge this as having any substantive significance. Figure X allowed us to explore issues that can inform new theory building.

7. Conclusion

This work is very challenging and can be explored further. We still have work to do on reconciling the difficulties within our method. However, we do have other research projects underway that will look at different research contexts. One such project is in the area of Internet abuse in the workplace and hopefully the initial pilot project will be ready to report later this year. We would encourage constructive comments on our work to strengthen it and intend to reflect on them over the coming months.

References


