Design and Control Issues in Qualitative Case Study Research

Gaynor Lloyd-Jones

Gaynor Lloyd-Jones, Ph.D., CSPPD Coordinator, Medical Teaching Organization, University of Edinburgh, Scotland

Abstract: *Some methodologists have pointed to similarities between experimental method and case study research in terms of design, theory testing and development. However, little is known about how these debates inform qualitative research rationales. The use of a sequential dual case study provided an opportunity to examine these issues and their impact on the unfolding research process. The interplay of inductive and deductive approaches was evident throughout in decisions determining the nature of the research enquiry.*

Keywords: *Qualitative case study research; experimental method; qualitative research design*

Introduction

Much is sometimes made of the distinctions between qualitative and quantitative research design and the development of the enquiry process. Quantitative research honours the logic of experimental or correlational method in adhering to agreed rules and predetermined sequences, irrespective of emerging data and analysis. The role of researcher is detached from the field of enquiry. By contrast, qualitative design displays an interactive, dynamic, and emergent character in which the aims, strategies, data, analysis, and validity are woven together in the process of the study (Hammersley & Atkinson, 1995; Maxwell, 1996; Becker, 1996). The qualitative researcher is the key instrument in the design process, continually deploying reflexivity and evaluative skills to data analysis and to the decisions concerning the direction of the next step in the study. The design of each qualitative research study might therefore be considered unique.

Yet dissenting opinions point to greater complexity beneath the surface of this polarized perspective. Hammersley (1992) argues that the qualitative-quantitative divide is artificially polarized, disguising both methodological similarity and diversity in consequence. In an analysis that emphasizes the trade-offs and overlaps between experimental, survey, and qualitative research, he concludes that no single approach is necessarily ideal and that selection inevitably involves loss as well as gain. This raises the possibility that the underlying logic between approaches may be shared, and also that different approaches may be strategically deployed to offset their particular disadvantages and advantages. Hammersley is not alone in seeing relationships between qualitative research and experimental method. Yin (1994) draws attention to similarities at the design level, seeing the single case study as analogous to a single

experiment in terms of theory testing and development. What is shared between qualitative and quantitative research may therefore be more real than apparent (Becker, 1996).

Although these arguments are well delineated in the literature, their impact on methodological decisions in research practice is less well documented. Whether they can be used in qualitative case study research, what implications ensue from their adoption, and how they might influence the ongoing design, conduct, and study outcomes remain largely unanswered questions.

Qualitative research and experimental method

In order to place the discussion in context, three interrelated features of research methodology will be used to compare qualitative research and experimental method: theory testing, the nature and use of control, and the induction-deduction dimension. This is not intended as a comprehensive review but serves to highlight those aspects of research design that have practical implications for the present study.

The ultimate aim of experimental method is to develop theory through repeated testing of related hypotheses. However, the outcomes of experiments are not generalizable per se but by reference to the related theory. In practice, much replication is required under similar and differing conditions before a theory may be upheld, rejected, or modified, and the point at which this occurs requires the exercise of judgement (Campbell & Stanley, 1963). Yin (1994) employs similar replication logic in his description of case study method, notably in his treatment of multiple case study design. He stresses the careful selection of cases that will either replicate (literal replication) or produce contrasting findings (theoretical replication) in line with the

prevailing theory. Similar findings uphold the theory whereas contradictory findings either demand theory modification or refutation thus mimicking experimental method. Though not all qualitative studies are intended, by any means, either to develop or test theory, theoretical inference has been proposed as a means of combating the problem of local to global generalization inherent in qualitative case study research (Gomm, Hammersley, & Foster, 2000; Yin, 1994).

There is a small, but significant, point to be made concerning theory testing and threats to validity in experimental method. In general, an experiment tests only a single theory, although it may be designed to differentiate between two rival hypotheses. Therefore, valid inference from experimental method is prey to threats from unforeseen, conflicting theories or hypotheses despite the use of control measures (Cook & Campbell, 1979). Bem's (1967) alternative reading of Festinger's (1957) cognitive dissonance theory is a case in point. Qualitative researchers could argue that inductive approaches might be less prone to the problem but this is true in terms of potential only. As in all research, omissions pose potential threats to validity similar to misinterpretations, treatment artefacts, or measurement error (Popper, 1972).

In essence, experimental method relies on the logic of comparison, contrasting the outcomes of two conditions that vary by a single assigned variable. Validity threats from rival hypotheses are excluded by controlling extraneous, but not necessarily identified, influences on the outcome. It is important to differentiate between the use of comparison and control measures for the former is no stranger to qualitative researchers. It is central to the analysis of grounded theory, implicit in natural experimental settings and widely used in historical and political research where

multiple cases are available for analysis (Ragin, 1994; Gomm et al., 2000). In the latter instances, the comparisons relate to naturally occurring examples, avoiding the artificial contrivance of control. Experimental control is essentially a means to an end. It is the technique that allows the creation of different manufactured conditions that permit comparison and valid inference.

The approach to control, and the degree to which it is pursued, differentiate between case study and experimental method (Hammersley, 1992). Experimental control measures act as a filter against generic validity threats, relieving the researcher of the need to identify rival hypotheses (Maxwell, 1996) although occasionally threats are specified (Campbell & Stanley, 1963). Experimental control encompasses a range of strategies such as pre- or post-testing, matched groups and contexts, randomization, and treatment equivalence all devoted to close matching of the control and treatment groups. The absence of rigorous control measures is one reason for the pluralistic, diverse character of qualitative research design, as the researcher must seek out alternative explanations and hypotheses throughout the research process design, data collection, and analysis (Maxwell, 1996). The success with which such threats are kept at bay will depend on the researcher's analytical and imaginative insights that inform methodological rationales and decisions. Control, as conceived in quantitative research terms, conflicts with the respect paid to context and naturalism in qualitative research. In practice however, qualitative researchers do operate with lesser degrees of control (Hammersley, 1992). Data collection methods such as structured interviewing and focus groups inhabit a 'no man's land' between naturalism and control. In such situations, the choice of greater control may influence the data in ways that compromise the representativity of the subsequent analysis.

A similar, and often cited, distinction between the inductive approach of case study and the deduction of experimental method also blurs under scrutiny. Though hypothesis raising and testing are claimed as definitive of deductive reasoning they are neither the exclusive preserve of quantitative research nor the only ways that deductive reasoning can be employed (Hammersley, 1992). Qualitative researchers employ both cognitive processes informally in the development of the enquiry process (Becker, 1993). Yet the deductive principles of experimental research do lead to a restrictive focus when control is relied on to exclude invalid inference. By contrast, qualitative researchers must remain open and alert to possible alternatives and it is this quality that marks qualitative research as primarily inductive. It seems likely that the qualitative researcher will encounter choices that reflect some of these considerations during the research enquiry.

Few qualitative researchers have utilized sequential case studies to develop and test theory in a manner similar to experimental method. Though practical considerations of time, scale, and feasibility discourage such endeavours, the assumption that the complex pluralistic nature of social life does not rest on universal laws constitutes the greater obstacle (Lincoln & Guba, 1985). Yet the ethnographic studies of Hargreaves (1967), Lacey (1970), and Ball (1981) offer a striking example from the qualitative case study literature of the development of differentiation-polarization theory. The theory proposes that streaming secondary school pupils, according to academic ability, results in polarization of their attitudes toward the school. The streaming practices varied in the different schools studied and the relationship between these practices and the degree of polarization found represented a test of theory through the medium of qualitative research (Hammersley, 1992). In this article, a similar strategy of theory development and testing

is described, but within a single research study. The methodological rationales that directed the enquiry process are highlighted to explore the conceptual territory between qualitative design and experimental method.

Background to the study

The background to the study was a wave of curricular revision in medical undergraduate education in the United Kingdom, instigated by the General Medical Council (1993). The medical school at the University of Liverpool responded by developing a subject integrated, problem-based learning (PBL) course. Small group, enquiry-led learning replaced the former heavy diet of lectures and discrete, discipline-based courses were combined in a curriculum that integrated natural, clinical, and social sciences throughout. The course was launched in 1996 with a pioneer intake of 200 students, while senior students continued on the traditional course that was gradually phased out. PBL had already established a foothold in US/Canadian medical education but, despite this, there were few successful examples of institutions negotiating curriculum change of the scale contemplated by Liverpool.

Design of the study

The double case design employed in the research (see Table 1) was not predetermined at the start of data collection but emerged gradually in response to literature reviews, collected data, and analysis. This section describes the rationales and possibilities entertained at the start of the study.

As a medically qualified curriculum facilitator working in the medical school, the researcher was familiar with experimental research method and was attracted to the possibility of theory

development and testing within Yin's (1994) model of multiple case study design. As a two-year

period of qualitative data collection was planned, it was feasible to look at two separate academic

sessions, each representing a single case study. The choice of cohorts and course years was not

made at the outset, other than a decision to start data collection with the launch of the course. By

default, the pioneer year thus became the first case study.

First case study	Second case study
1996/7 academic session (pioneer entry)	1997/8 academic session
Peer social reference group – senior medical students	Peer social reference group – senior medical students
No peer educational reference group	Peer educational reference group – pioneer entry group

Table 1. Multiple case study design of the Liverpool study

The choice of the second case study was left open, although three possibilities were considered.

The one ultimately adopted was to study the next student entry on the PBL course, in effect

comparing the experiences of two first-year cohorts in two case studies (Table 1). Alternatively

the pioneer cohort studied in the first case study could be followed into the second course year,

thereby comparing the experience of different course years from the viewpoint of a single cohort.

A third plan combined the two, researching the second first-year student entry plus the pioneer

cohort experience in the second year of the course. The latter idea proved too ambitious for the

resources of a single part-time researcher and was discarded early on.

Reviewing the literature revealed three areas of relevance for study design development. First,

there was a series of ethnographic case studies of the medical student experience (Fox, 1957;

Becker, Geer, Hughes, & Strauss, 1977; Haas & Shaffir, 1991), which could provide comparison

with the present study. In two of these student experience was characterized by insecurity

attributed to impending professional responsibilities (Fox, 1957; Haas & Shaffir, 1991) but in the third, educational cueing[1] to faculty, tutors, and assessment was prominent and uncertainty was not in evidence (Becker et al., 1977).

A separate review of the PBL literature showed that little was known about the nature of PBL student experience. This was highly significant because PBL proponents claim that the actions of the typical PBL student differ from more conventional students. PBL students are self-directed learners who recognize their own individual knowledge deficiencies and take responsibility for satisfying their learning needs (Barrows & Tamblyn, 1980; Schmidt, 1983). Neither cueing to teachers, nor to peers is compatible with PBL student action, for PBL is an emphatically individualistic educational approach. The outcomes of both literature reviews therefore justified a focus on the student perspective, which in essence became the case at the heart of each case study.

A study of reference groups at a women's college in the United States during the depression years also fed into design development (Newcomb, Koenig, Flacks, & Warwick, 1967). In that study, Bennington College (now co-ed) students adopted contrasting political allegiances to those of family (Democrat) and background (Republican), attitudes that were retained 20 years later. Yet it was not the findings of the American study that were of interest to the Liverpool setting, but the application of the concept of reference groups. On reflection, it was apparent that incoming students could employ senior students as reference groups for advice about social and educational matters. A closer analysis revealed potential distinctions between the experiences of the pioneer entry students and the following year entry in this respect. Clearly, both cohorts

would have access to advice about social affairs but this would not be true for education. The pioneer entry would lack a senior student reference group, as they were experiencing a quite different curriculum. The following year would revert to normative circumstances, as the pioneer entry became the senior educational reference group. At this stage, the analysis was speculative and clearly did not preclude other contextual differences between pioneer and later cohorts but it tended to favour the study of different cohorts. However, the final decision was not taken until completion of the first case study analysis.

Methods

Details of the rationales underlying the unfolding enquiry process in relation to methods of data collection are described in full elsewhere (Lloyd-Jones, 2002; Lloyd-Jones & Hak, in press), but are only outlined in this article concentrate on the aspect of theory development within the design. Participant observation and direct observational methods formed the mainstay of data collection methods and were supplemented with interviews, focus groups, analysis of course documents and resources, and a survey derived from a nominal group analysis. Consent for the study was gained at the start from the Dean and all course directors. Every student received a copy of the research guidelines assuring them of their rights to confidentiality and of their right to decline participation. Staff tutors were approached individually on a similar basis.

First case study

Insecurity characterized the student experience during the first term. Study practices contradicted the PBL model as students were observed cueing to staff, the given resources, and references. Interview data triangulated the observational data, and inspection of student notes also confirmed

students' reliance on course resources. The insecurity proved problematical for research

development as its causes could not be attributed with certainty. Though the students blamed

uncertainty on the lack of an explicit, shared syllabus, the unfamiliarity and organizational

teething problems of a novel course could not be excluded as contributory. Nor was it possible to

assess how the absence of an educational reference group affected the pioneer students. During

the spring term, insecurity gradually declined as peer interaction and social comparison increased

among the cohort. Students compared workloads, notes, and PBL study, and modified private

study accordingly if they deviated from the consensus. However, insecurity resurfaced in the

summer with approaching assessment as students worried that the exclusive pursuit of PBL

learning might risk examination failure by omitting unseen elements of the faculty agenda. For

the first time, students articulated concern about their pioneer status, speculating about their

knowledge and competence in comparison with students at other medical schools.

The main points to emerge from an analysis of the first case study were insecurity, students'

rejection of self-directed study practices in favour of faculty-directed learning and the gradual

development of collective identity and action among the cohort. These findings favoured the

design in Table 1 for three reasons. First, the study practice findings were novel and unexpected.

The assumption that PBL students acted according to PBL claims had not previously been

challenged, although there was little supporting evidence for it. The first case study findings

could therefore be dismissed as atypical because of the peculiarities of context unless additional

evidence could be found. Evidence of similar practices in a different student cohort in a more

normative context would give greater robustness to the first case study findings. Second,

studying a new cohort experience under more normative circumstances might test the speculation

relating to reference groups and, third, help to illuminate the cause of insecurity in the first case study. Based on the logic of the study design and the first case study analysis, hypotheses were raised for the second case study, which predicted recurrences of all three phenomena in the second case study. However, the presence of the pioneer group as a source of educational advice was anticipated to reduce, but not eradicate, uncertainty. How this effect might be mediated was regarded as an emergent research question to be answered by data collection and analysis.

Second case study

The hypothesis on study practices was strongly upheld from the start of term as students were observed openly cueing to resources, tutors, and faculty, similar to their predecessors. Collective identity and action developed rapidly but, on the other hand, insecurity was striking by its absence. The refutation of the related hypothesis required explanation and the answer was sought, and found, in interactions between the pioneer student entry and the new group first year students. Acting backstage of formal educational settings, the pioneer year had conveyed advice on a specific textbook, which had been adopted, almost without exception. The effects of this were not only to standardize content knowledge and reduce insecurity, but also to undermine the individualistic aims of PBL learning. However, uncertainty reappeared in focus group data in the spring term, prompted by a recent changeover of PBL tutorial groups and tutors. Students' former cues of tutor direction were rendered unreliable as they encountered unexpected variation in tutor and group practice.

In neither case study had students' study practices conformed to those claimed by advocates for PBL. Though the timing and degree of insecurity in the two case studies was strikingly different,

in both groups a consistent relationship was found between the existence of collective social groupings and the diminution of uncertainty. Group consensus and conformity appeared to allay insecurity in both case studies. Conversely, uncertainty was maximal when group consensus was either undeveloped, as at the start of the first case study, or when challenged by group and tutor changes in the second case study.

Multiple case study design and the natural experiment

The dual case study design capitalized on the natural experimental features of the setting by comparing the experience of two successive student cohorts in naturally occurring, but different, contexts. The various contextual similarities and differences thus constituted a naturally occurring control structure. The institution, the curriculum and medical student social culture could be regarded as consistent features across both case studies. On the other hand, initial problems of implementation in the first case study and growing staff familiarity with the course in the second were potential differences. However, the most striking distinction that emerged fully during data collection arose from the different relationships between first year and senior students as a consequence of curriculum. The initial speculation on senior student role and action in the early stages of the study proved justified in the findings of the second case study, where their advice on textbook use and workload at registration was universally adopted, immediately promoting educational conformity and undermining PBL process.

The design shares some features with Yin's (1994) model of multiple case study design, notably in his treatment of each case study as separate entities that permit replication and in the testing of theory in the second case study. These are principles shared in common with experimental

method, although in the natural world replication is too precise a term. The testing in the second case study occurs under some conditions that are shared with the first, but also under circumstances that uniquely differentiate between the case studies. The findings that appeared in both case studies therefore acquire greater robustness as they have withstood the influence of contextual variation. For instance, the relationship between insecurity, educational cueing, and collective identity held good in the second case study despite the educational referent action, the different triggers to uncertainty and the temporal variations in the appearance of phenomena when compared to the first case study (Lloyd-Jones, 2002). The second case study data also served as a source, and means, of refining and developing theory. As an example, the low level of initial student insecurity in the second case study confounded expectations and triggered a search for explanations that ultimately led back to second-year student action.

However the design differs from Yin (1994) in at least two respects. First, the approach taken with the first case study was exploratory and inductive, for the uncertainty of outcomes and the novelty of the situation militated against a deductive strategy so early in the study. This meant that theory emerged from data analysis of the first case study rather than, as Yin suggests, being developed deductively before any data are collected. Here, a more qualitative approach to study design was necessary to run neither the risk of prejudging events nor of being deaf to alternative hypotheses too early in the research. The first case study therefore, served as an exploratory vehicle for theory development from which hypotheses or predictions were raised for testing in the second case study.

The second contrast lies in the coexistence of literal and theoretical replications within the second case study, which again imposes a more open and emergent stance on the researcher. Since the second case study was contingent on the first in terms both of data and design, it was not possible to identify at the start of the research what might constitute literal, and what theoretical, replications in Yin's (1994) terms. That could only be achieved following an analysis of the first case study data, and then only in the form of provisional hypotheses. The hypotheses, founded on literal replications and for testing in the second case study, predicted that educational cueing and the influence of collective action on study practices would recur. These were upheld in the data. Another hypothesis, based on theoretical replication, anticipated diminished initial uncertainty, lessened by the presence and actions of the senior second year students. However, exactly how the educational reference group action might be mediated was less certain and had to await further data collection. Here the research reverted to the emergent, inductive approach of the first case study.

The case study posits an alternative explanation for PBL student action and behaviour to the one outlined by educationalists (Barrows & Tamblyn, 1980; Schmidt, 1983). The model is derived from empirically grounded data that respect those social and contextual influences neglected in the educational theory. In terms of theory testing in experimental method, this constitutes an alternative hypothesis to the existing claims of PBL. That the qualitatively derived model is congruent with PBL research derived from a variety of methodologies suggests the findings of the present study might not be unique (Lloyd-Jones, 2002). Yet caution should be exercised in generalizing to other PBL curricula on two grounds. The first is because the study represents the

equivalent of two experiments only, and second reason is because of the general acceptance of probabilistic rather than deterministic laws governing social action amongst social researchers.

The discussion has been primarily concerned with maximizing research opportunities to deploy theory construction and testing but there are limitations to the approach. Significant methodological implications flow from the design since the logic of the multiple case study rests on comparing like with like. This is problematical for qualitative research where depth is valued more highly than breadth, for there is the danger that the research may concentrate on non-equivalent elements for comparison. Furthermore, without knowledge of the relationship between the whole case and the constituent parts, as represented by the collected data, claims to generalization within the case study may be tenuous. Gomm et al. (2000) have pointed to the lack of attention paid by qualitative researchers to generalization within the case study and of the cautions necessary in making assertions about the case on the basis of partial knowledge. In the present study attempts to combat the problem were made through rigorous attention to sampling of respondents, time and events and the use of a survey instrument (Lloyd-Jones, 2002) but the issue remains a potential validity threat.

Conclusion

Recent interest in qualitative research methodology has created opportunities to explore conceptual relationships between qualitative and quantitative design and methods. Though it is accepted by some researchers that theory development may be achieved by qualitative case study research, these examples are rare. A qualitative case study, which blended notions of experimental method with qualitative research, has been described and the design rationale

examined. By rendering more explicit the arguments and decision making processes that

contribute to qualitative case study design, it is possible to gain a clearer assessment of a study's

strengths and limitations, as well as contribute to our understanding of qualitative case study

design and methodology.

Notes

1. The term *cueing* or *educational cueing* refers to the way students seek clues or "cues" from the educational environment as to what the faculty or school wishes them to know and learn. In its most common form students try to find out about assessment, particularly examination content, from their teachers. It was particularly obvious in the study quoted in the article because PBL curricula do not specify curricular content, leaving it up to the student to define what they think they should learn. This was a cause of insecurity to them. The students' problem became acute when they faced examinations as there is no formal curriculum to refer to. Consequently, they try to 'read' tutor's behaviour and search the resources for clues as to what might be included in the assessment. The term 'curriculum hunting' can sometimes be used as a substitute.

References

Ball, S. J. (1981). *Beachside comprehensive: A case study of secondary schooling.* Cambridge, UK: Cambridge University Press.

Barrows, H. S., & Tamblyn, R.N. (1980). *Problem-based learning: An approach to medical education*. New York: Springer.

Becker, H. S., Geer, B., Hughes, E. C. & Strauss, A.L. (1977). *Boys in white: Student culture in medical school*. Chicago: Chicago University Press.

Becker, H.S. (1993). How I learned what a crock was. *Journal of Contemporary Ethnography*, *22*, 28-35.

Becker, H. S. (1996). The epistemology of qualitative research. In R. Jessor, A. Colby, & R. Schweder, (Eds.), *Essays on ethnography and human development* (pp. 53-72). Chicago: University of Chicago Press.

Bem, D. J. (1967). Self-perception: An alternative interpretation of cognitive dissonance phenomena. *Psychological Review*, *74*, 198-200.

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research.* Chicago: Rand McNally College.

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation design and analysis issues for field settings*. Chicago: Rand McNally College.

Festinger, L. (1957). *A theory of cognitive dissonance.* Stanford, CT: Stanford University Press.

Fox, R. (1957). Training for uncertainty. In R.K. Merton, G. G. Reader, & P. L. Kendall, (Eds.), *The student-physician Introductory studies in the sociology of medical education* (pp. 207-241). Cambridge, MA.: Harvard University Press.

General Medical Council. (1993). *Tomorrow's doctors*. London: General Medical Council.

Gomm, R., Hammersley, M., & Foster, P. (2000). *Case study method*. London: Sage.

Haas, J., & Shaffir, W. (1991). *Becoming doctors: The adoption of a cloak of competence.* Greenwich, CT: JAI Press.

Hammersley, M. (1992). *What's wrong with ethnography?* London: Routledge.

Hammersley, M., & Atkinson, P. (1995). *Ethnography Principles in practice* (2nd ed.). London: Routledge.

Hargreaves, D. H. (1967). *Social relations in a secondary school.* London: Routledge.

Lacey, C. (1970). *Hightown grammar.* Manchester, UK: Manchester University Press.

Lincoln, Y. S. & Guba, E. G. (1985). The only generalization is: there is no generalization. In Y. S. Lincoln & E. G. Guba (Eds.), *Naturalistic inquiry* Newbury Park, CA: Sage.

Lloyd-Jones, G. (2002). *A multiple case study of the first year student perspective in a medical undergraduate PBL curriculum.* Unpublished doctoral dissertation, University of Liverpool, England.

Lloyd-Jones, G., & Hak, A. (in press). Self-directed learning and student pragmatism. *Advances in Health Sciences Education.*

Maxwell, J. A. (1996). *Qualitative research design: An interactive approach*. Thousand Oaks, CA: Sage.

Newcomb, T. M., Koenig, K., Flacks, R., & Warwick, D. P. (1967). *Persistence and change: Bennington College and its students after 25 years*. New York: John Wiley & Sons.

Popper, K. R. (1972). Truth, rationality and the growth of scientific knowledge. *Conjectures and refutations: The growth of scientific knowledge* (pp. 215-250). London: Routledge and Kegan Paul.

Ragin, C. C. (1994). *Constructing social research: The unity and diversity of method*. Thousand Oaks, CA: Pine Forge Press.

Schmidt, H. G. (1983). Problem-based learning: Rationale and description. *Medical Education*, *17*, 11-16.

Yin, R. K. (1994). *Case study research: Design and methods* (2nd ed.). Thousand Oaks, CA: Sage.