Triangulation between case study and survey methods in management accounting research: An assessment of validity implications

Sven Modell *

Stockholm University, School of Business, S-106 91 Stockholm, Sweden
Accepted 8 February 2005

Abstract

Repeated calls for validating empirical research by combining qualitative and quantitative methods have recently been made in management accounting research conducted within the positivist and functionalist paradigms. This paper provides a comparative review of management accounting research relying on triangulation between case study and survey methods with specific reference to the emerging validity implications. This research is classified into studies primarily relying on a theory testing and development logic, respectively, and works characterised by a more balanced emphasis on these logics. Our review suggests that theory-testing approaches are primarily geared towards external and construct validation and internal validation through corroboration of converging findings in line with some replication logic. By contrast, a theory development approach is typically based on internal validation efforts stemming from further probing of unexpected or inconclusive findings as a means of theory extension. An argument is advanced for combining elements of these approaches in a more balanced fashion within coherent research programmes based on multiple iterations between case study and survey methods to enable a broader range of potential validity threats to be addressed. Some general areas for further development of such a triangulation approach are also identified.

© 2005 Elsevier Ltd. All rights reserved.

Keywords: Case study; Management accounting; Survey; Triangulation; Validity
1. Introduction

In recent years, management accounting research conducted within the positivist and functionalist paradigms has shown increasing recognition of the need to complement established quantitative methods with a greater or lesser element of qualitative, case study-based research. A number of calls for such a complementary approach, relying on method triangulation, have been made (e.g., Ferreira and Merchant, 1992; Ittner and Larcker, 2001; Shields, 1997) and a growing number of advances in this direction can be found in empirical management accounting research.

The key justification for method triangulation is typically enhancement of the validity of research findings (Brewer and Hunter, 1989; Bryman, 1992; Denzin, 1978; Jick, 1979). However, a more systematic and in-depth discussion of validity implications associated with method triangulation has largely eluded the evolving methodological debate in the management accounting literature. A few prior attempts to contrast validity aspects associated with laboratory, survey and case study research in management accounting have been made (Abernethy et al., 1999; Birnberg et al., 1990; Brownell, 1995). A limitation of these relatively broadly based reviews is that they primarily draw on multiple studies, each dominated by one or the other method, whilst little attention has been paid to validity issues arising from different triangulation approaches in single, or closely related studies conducted by the same researcher(s) (cf. Creswell, 1994; Tashakkori and Teddlie, 1998). Further, the rationale for method triangulation as a means of avoiding potential validity threats has typically been formulated in relatively general terms based on the potential strengths and weaknesses of different methods (see e.g., Abernethy et al., 1999; Birnberg et al., 1990). As noted by Brewer and Hunter (1989) such a priori assumptions are insufficient for making informed judgements of how potential validity threats may be reduced in specific studies.

Our understanding of how triangulation between qualitative and quantitative methods may be used as a validation technique in management accounting research thus requires further development. Specifically, a systematic attempt to take stock of prior experiences with different triangulation approaches would seem warranted. The purpose of this paper is to explicate these issues, based on a comparative review of relevant management accounting research. We concentrate our review to research combining elements of (qualitative) case study and (quantitative) survey methods, which constitutes the dominant mode of triangulating qualitative and quantitative methods in the management accounting literature.

The remainder of the paper is structured as follows. In Section 2, we describe the notion of triangulation, especially method and theory triangulation, in some detail. This leads us to propose a few criteria for classifying and assessing empirical research. In Section 3, the use of triangulation is related to the three key issues of external, internal and construct validity, commonly used to assess research within the positivist tradition, whether of a predominantly quantitative or qualitative nature (see e.g., Abernethy et al., 1999; Cook and Campbell, 1979; Yin, 1984). In Section 4, we review a number of attempts to combine case study and survey methods in the examination of specific management accounting issues with an eye to the triangulation approach adopted and the subsequent implications for the validity issues previously raised. We conclude the paper with a discussion of the ramifications of our observations in Section 5.
2. Triangulation in social science research

In the social sciences, the notion of triangulation can be traced back to Campbell and Fiske’s (1959) use of multiple quantitative methods for assessing convergent and divergent validity. Denzin (1978), however, distinguished such within-method triangulation, or the application of multiple data collection and/or analysis techniques within an overarching research method (e.g., the survey), from *between-method triangulation*, representing a more genuine synthesis between dissimilar methods (e.g., case study and survey methods). The emphasis of the present paper falls within the latter, arguably more rewarding but also complex and challenging form of method triangulation (Denzin, 1978; Jick, 1979). The rationale for between-method triangulation is that it offers some advantages in dealing with validity threats stemming from the biases inherent in any single method. By combining relatively diverging methods, such that the relative strengths of one counter-balance the weaknesses of the other(s), researchers may enhance the credibility of their results whilst reducing the risk of observations reflecting some unique method artefact (Denzin, 1978). Triangulation between case study and survey methods arguably provides a relatively potent means of assessing the degree of convergence as well as elaborating on divergences between results obtained (Brewer and Hunter, 1989; Jick, 1979; Sieber, 1973). For example, surveys may improve our understanding of the incidence of a particular phenomenon and/or the form and strength of conceptual relationships observed in case studies. On the other hand, case study methods may add to a more holistic and richer contextual understanding of survey results and help to explain apparent anomalies or puzzles emerging from the latter.

Method triangulation should be distinguished from, but is not necessarily independent of the three other forms of triangulation found in the literature: data, investigator and theory triangulation (Denzin, 1978; Ryan et al., 2002). Whilst data and investigator triangulation are of limited concern in the present paper, greater attention will be paid to the combination of between-method and theory triangulation. Theory triangulation implies that hypotheses or researcher interpretations are informed by more than one theoretical perspective. This may sensitize researchers to rival or complementary explanations evident in empirical data, but eschewed by the adoption of a single or dominant theoretical perspective and stimulate the combination of different methods in the examination of specific substantive phenomena (Brewer and Hunter, 1989). However, this should not be confused with more eclectic “layering” of theories (see, e.g., Berry et al., 1991) without much synthesis or integration between these to derive explanations of a particular empirical phenomenon (Denzin, 1978). Though theory triangulation may prove useful in enhancing our understanding of inductively inferred explanations. Such means of furthering the integration between theory and empirical research would seem particularly useful where less conclusive findings emerge (Denzin, 1978; Erzberger and Prein, 1997) and may help in overcoming some pertinent criticisms of management accounting research. Much case study research in the area has arguably paid insufficient attention to the integration of theory and empirics and cumulative theorizing due to its strongly inductive and exploratory character (Humphrey and Scapens, 1996; Otley and Berry, 1994; Spicer, 1992). The more general value of integrating complementary theories and methods has recently been emphasised in response to similar criticisms of predominantly quantitative management accounting research (Luft and Shields, 2002).

The use of triangulation for addressing different validity issues should be assessed in relation to the types of research questions posed and the subsequent mode of theorizing (Brewer and Hunter, 1989; Tashakkori and Teddlie, 1998). Survey research in management accounting, generally carried out within
a positivist vein, has primarily aimed at theory testing relying on mainly deductively derived hypotheses (Ryan et al., 2002). Whilst case study methods have typically been confined to a relatively limited role as vehicles for theory (or hypothesis) development, more recent advances within this research tradition recognize their usefulness for broader purposes (e.g., Atkinson and Shaffir, 1998; Keating, 1995). Case study-based theory development generally aims at inductively dominated theory building (Eisenhardt, 1989), but may also encapsulate some refinements or modifications of existing theories (Keating, 1995). However, the use of case studies for testing theories by submitting them to critical attempts at refutation is relatively rare in management accounting research (Keating, 1995). As illustrated in the following section, the incorporation of qualitative methods in triangulated research may enhance their role in the process of theory testing. Research relying on between-method triangulation rarely applies “pure” theory testing or development logics (Tashakkori and Teddlie, 1998). However, we argue that there may be some variation in the relative emphasis on these respective roles, which has important ramifications for how validity issues are addressed.

3. Validity criteria

In this section we briefly consider the three key issues of external, internal and construct validity. Following the distinctions presented in the preceding section, we discuss how each of these validity issues may be addressed in triangulated research primarily relying on a theory testing and development logic, respectively. This discussion is then extended in our review of the management accounting literature.

3.1. External validity

The issue of external validity has traditionally been conceived of as the extent to which the findings of a particular study can be generalized across populations, contexts and time (Birnberg et al., 1990). The differential treatment of this issue in the social sciences has received ample attention in the literature and shall not be repeated in detail here. Whilst the positivist research tradition has emphasised statistical inference as a basis for generalizations, an increasingly accepted alternative criterion is that of analytical generalization based on close iterations between existing and emerging theory and empirical findings in accordance with some replication or extension logic (Eisenhardt, 1989; Lindsay, 1995; Yin, 1981, 1984). Such an approach sits more comfortably with triangulation between case study and survey methods (see Tashakkori and Teddlie, 1998). As explicated below, it also offers some opportunities to address important criticisms levelled at both types of research in management accounting.

The application of some replication/extension logic in case study research is typically manifested by theorizing based on a systematic process of searching for cross-case patterns or regularities whilst

---

3 Another alternative to the positivist concern with statistical generalizability for assessing external validity, occasionally emphasized in the management accounting literature, is that of practical and managerial relevance of findings (Bruns and Kaplan, 1987; Kaplan, 1986). Whilst we do not negate the potential for method triangulation to contribute to this end, such validity criteria are of lesser concern in our discussion as research primarily guided by these has tended to place relatively one-sided emphasis on exploratory case studies at the expense of the wider role that qualitative methods may assume in the process of theory development (see Humphrey and Scapens, 1996; Spicer, 1992).
probing unexpected or divergent relationships across cases (Eisenhardt, 1989; Yin, 1981). This should ideally entail some analytical generalizations through comparisons with existing theory, such that the external validity of the resulting theoretical explanations, or the likelihood of these holding up across a wider array of cases, is enhanced (Yin, 1984). This allows qualitative research to transcend the inductive and exploratory function to incorporate a greater explanatory element (Keating, 1995; Lukka and Kasanen, 1995; Scapens, 1990, 1992; Yin, 1981).

Interestingly, similar concerns with the need for solid theoretical grounding of generalizations based on quantitative methods have recently emerged in the accounting literature. Lukka and Kasanen (1995) questioned the generalizing rhetoric often invoked in quantitative research, typically manifested by heavy emphasis on statistical significance tests, arguing that such rhetoric can never substitute for thorough theoretical knowledge and framing of the research issues at hand. Similarly, Lindsay (1995, p. 39) warned against “the mechanization of inference”, arguing that statistical significance tests do not eliminate the need for theoretically informed judgement. According to Lindsay (1995), a useful alternative approach to generalization is to adopt a more pronounced replication logic than that prevailing in much survey-based management accounting research. This should preferably take the form of coherent research programmes, enabling researchers to generalize across studies in different contexts whilst remaining open to theoretical insights offered by alternative research approaches such as case studies. Exact replications using the same method enhances the risk of spurious convergence hiding shared method biases, whilst replications coupled with some extensions based on different methods may guard against this (Brewer and Hunter, 1989).

Whilst occasionally applied in combination, the relative emphasis on theoretically informed replications and extensions would seem to be closely related to the distinction between theory testing and development. This, in turn, has important implications for the type of external validation efforts involved in triangulated research. Where theory testing is the primary goal, we may expect a quantitatively dominated replication logic and extensive reliance on prior theorizing for developing testable hypotheses to dominate (Creswell, 1994). Whilst such a replication logic is rare in management accounting research (Abernethy et al., 1999; Lindsay, 1995; Lindsay and Ehrenberg, 1993), it may be reinforced by the use of qualitative methods as a complement to theory for developing a priori hypotheses (Eisenhardt, 1989). External validity is enhanced where patterns or hypotheses emerging from case studies are corroborated by survey-based tests across larger samples. However, the use of qualitative methods for extending theory by refining explanations of inconclusive or unexpected survey findings would be comparatively limited where a theory testing logic dominates.4

Conversely, a stronger emphasis on theory development often stems from greater reliance on the qualitative element to stimulate the generation of novel theoretical insights based on the explanation of inconclusive findings (Tashakkori and Teddlie, 1998). Theory extension through formulation of alternative or complementary theoretical predictions (rather than strict replication) is typically the end result of such efforts (Brewer and Hunter, 1989). Whilst this may be particularly fruitful where prior theory is less well-developed, external validity may be enhanced if such extensions are informed by careful theory triangulation (Denzin, 1978; Erzberger and Prein, 1997).

4 Whilst replications should typically entail some extension across different empirical settings, their successfullness rests on the premise that a particular theoretical model remains robust despite such contextual variations (Lindsay and Ehrenberg, 1993). Although alternative methods may be used for corroborating results in accordance with a triangulation logic, such convergence typically implies that the use of this technique for theory extension is limited (cf. Jick, 1979).
3.2. Internal validity

The internal validity of a specific study refers to the credibility of the causal relationships between independent and dependent variables inferred from data. Case study as well as survey methods have typically been considered inferior to controlled laboratory experiments in this respect (see, e.g., Birnberg et al., 1990; Brownell, 1995). However, triangulation between case study and survey methods may mitigate the difficulties in advancing at least plausible causal explanations (cf. Abernethy et al., 1999). As noted in the foregoing, much of the triangulation literature suggests that the combination of complementary methods enhances the opportunities of corroborating causal relationships by revealing converging patterns as well as expanding the quest for alternative causal relationships where findings diverge or are inconsistent with a priori hypotheses.

Achieving high internal validity in any one study is a pre-requisite for external validity (Birnberg et al., 1990). However, the manner in which this is accomplished is likely to vary somewhat depending on whether a theory testing or development logic dominates research. Where theory testing is the primary goal, the credibility of causal relationships is enhanced if survey results are both consistent with a priori hypotheses and converge with qualitative findings within a specific empirical setting. To ascertain that such convergences are non-spurious researchers may need to rely extensively on qualitative methods (whether used as a basis for hypothesis development or as part of follow-up procedures) to exhaustively explicate the form and direction of causal relationships (Brewer and Hunter, 1989; Sieber, 1973). Whilst well-established theoretical frameworks may facilitate this, they may also induce researcher selectivity regarding the detection and further examination of rival causal explanations (cf. Atkinson and Shaffir, 1998; Brewer and Hunter, 1989). An internal validity threat may thus emerge as a result of the causal model being under- or mis-specified. More broadly based replications should only be initiated once such risks have been minimized to enhance the probability of high external validity.

Where considerable problems of establishing convergence across different methods prevail, it is reasonable to assume that the emphasis of theorizing will shift towards theory development by more fully utilising the richness of qualitative data (Jick, 1979). Where inconclusive or unexpected survey findings emerge, qualitative methods may be useful for advancing a plausible theoretical rationale for the limited internal validity of the survey method based on in-depth examination of qualitative evidence (Brewer and Hunter, 1989; Jick, 1979). This does not affect the internal validity of the survey results per se. However, qualitative inquiries in the same empirical setting, preferably complemented with some theory triangulation (Denzin, 1978), may reveal potentially moderating or mediating variables omitted in original quantitative work. These may subsequently be incorporated into refined explanatory models clarifying the causal direction of relationships and reducing the risk of spuriousness in future studies (Tashakkori and Teddlie, 1998). Inconclusive survey findings may thus stimulate theory extension by energizing the quest for refined or alternative theoretical explanations (Jick, 1979). However, to ensure that the modified theoretical model is indeed internally valid probably requires corroboration through some small-scale, survey-based test in the same empirical setting before it is used for replication to assess its broader external validity (cf. Lee, 1991).

---

5 I am grateful to one of the anonymous reviewers for emphasising this.
3.3. Construct validity

Construct validity refers to whether theoretical concepts are adequately reflected by the operational definitions and measures of empirical phenomena. Whilst case study as well as survey research may vary from low to high construct validity depending on the meticulousness of the research design (Birnberg et al., 1990), it has typically been considered more problematic in the latter due to the need for abstraction inherent in such research. The greater closeness to the empirical research object often characterising case study research implies that investigators may be better positioned to achieve good fit between their conceptual apparatus and empirical data (Atkinson and Shaffir, 1998; Brownell, 1995). Survey-based management accounting research has arguably been fraught with unresolved construct validity problems (Brownell, 1995; Lindsay and Ehrenberg, 1993; Young, 1996). Recent recommendations for overcoming such problems stress inter alia the importance of complementing the development of measurement instruments with qualitative elements, such as feedback from practitioners as well as peers familiar with the research issues at hand (e.g., Kwok and Sharp, 1998; Roberts, 1999).

The typical means of addressing construct validity issues in triangulated research following a theory testing logic is to rely on some qualitative inquiry prior to the distribution of surveys, occasionally as an integral part of empirical development of hypotheses to be tested in subsequent phases (Eisenhardt, 1989; Jick, 1979; Sieber, 1973). This is especially useful where prior research provides few or inadequate measurement instruments to operationalise key constructs. However, prior advances in empirical management accounting research caution against premature and exaggerated creativity in this respect (Lindsay and Ehrenberg, 1993). Theory testing based on a pronounced replication logic requires some stability in the measurement instruments used such that unnecessary variance in results due to measurement differences is avoided (see, e.g., Otley and Fakiolas, 2000). On the other hand, a danger inherent in close replications is that the construct validity of measurement instruments is jeopardized if attributes of the measured empirical phenomenon are unstable across contexts and time (Brownell, 1995; Young, 1996). This would, in turn, have serious ramifications for the internal as well as external validity of replicated studies (cf. Birnberg et al., 1990).

Empirical detection of construct validity threats occasionally necessitates further qualitative probing, which may amplify the theory development aspects of triangulated research. Where potential construct validity threats manifest themselves in the form of inconclusive survey findings, there may be a need to disentangle these threats from internal validity issues, such as whether results are due to mis-specified causal relationships or omitted variables (cf. Brewer and Hunter, 1989). If successful, this may lead to important refinements of constructs and measurement instruments based on previously neglected dimensions and/or context-specific meanings contributing to the gradual specification of theoretical models (cf. Lee, 1991; Sieber, 1973). However, such validation efforts, often taking the form of interview-based follow-up procedures, are rare and generally of limited scope due to the procedural complexity and considerable resources required (Brewer and Hunter, 1989).

4. Review of empirical management accounting research

This section provides a review of existing management accounting research relying on between-method triangulation incorporating elements of case study and survey methods. The review primarily includes studies identified by scanning all volumes of leading academic accounting journals, which
might be expected to publish triangulated research of the kind concerned here between 1970 and 2002.\(^6\)

An overriding aim of the review is to enable readers to learn from prior experiences of applying between-method triangulation in a more “holistic” manner concomitantly addressing a broader range of validity issues (cf. Brewer and Hunter, 1989; Jick, 1979). The studies were thus selected on the premises that these (1) use case study and survey methods for extensively addressing more than one validity criterion and (2) apply such methods in an integrative manner for theorizing causal explanations.\(^7\) Given these constraints, the review should be considered illustrative and representative of reasonably extensive triangulation efforts rather than exhaustive.

Two further caveats pertaining to our review should be noted. First, possible constraints on the display of details pertaining to research methods and findings may bias validity assessments, especially when studies are subject to the space limitations inherent in journal articles (Brannen, 1992; Jick, 1979). Whilst there may be a risk of such constraints influencing the conclusions drawn, we did not experience any major classification problems emanating from the availability of information in the studies reviewed. In a few cases, the assessment of validity issues was facilitated by separate methodological reflections by the investigators (Covaleski and Dirsmith, 1990; Otley and Berry, 1994) and the author’s own involvement in one study.

Second, our discussion of the extent to which between-method triangulation has been used for enhancing external, internal and construct validity does not necessarily imply that specific studies score high or low on these criteria per se. Other means of validation (e.g., within-method triangulation) may have been applied concomitantly to between-method triangulation. The main concern of our assessment is the specific validity implications stemming from between-method triangulation in terms of how different validity issues are addressed and what strengths and weaknesses are associated with different triangulation approaches.

Table 1 contrasts the studies reviewed. Using the distinction between theory testing and development guiding our prior discussion of validity issues we categorize the studies into three groups. The first two groups consist of studies dominated by one of these logics, respectively. The third group comprises a number of studies applying both logics in a more balanced fashion. The ensuing discussion focuses on how specific validity issues are addressed within the three groups of studies.

4.1. Group 1: primarily theory testing


Merchant (1985, 1990) explored the influence of a range of contextual variables (e.g., environmental uncertainty, leadership style, business strategy, recent economic performance) on a number of financial as

\(^6\) This includes the following journals: Accounting, Auditing and Accountability Journal, Accounting and Business Research, Accounting, Organizations and Society, the Accounting Review, Contemporary Accounting Research, the European Accounting Review, Financial Accountability and Management, Journal of Accounting Research, Journal of Management Accounting Research and Management Accounting Research. A smaller number of research monographs were also reviewed.

\(^7\) The first premise led to the elimination of studies nearly exclusively using between-method triangulation for construct validation (Albrethsen and Browne, 1999; Albrethsen and Lillo, 1995; Brewer, 1995; Mia, 2000; Mia and Cheshall, 1994; Mia and Clarke, 1999). The second premise implies that we exclude a number of studies either failing to use triangulation for inferring some clear causal reasoning within a specific empirical setting (Hitner and Larcker, 1997; O’Connor, 1995) or making only limited attempts to advance theoretically generalizable explanations of causal relationships (Berry et al., 1991; Hoppner et al., 1992; Imposi, 1989; Joseph et al., 1996; Selbo et al., 1995).
Table 1
Overview of management accounting research triangulating case study and survey methods

<table>
<thead>
<tr>
<th>Studies</th>
<th>Substantive research issue</th>
<th>Main use of between-method triangulation</th>
<th>External validation</th>
<th>Internal validation</th>
<th>Construct validation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1: primarily theory testing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group 2: primarily theory development</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group 3: balanced emphasis on theory testing and development</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hopwood (1973)</td>
<td>Attitudinal and behavioural effects of budgetary control within one firm.</td>
<td>Theory refinement and extension.</td>
<td>Corroboration and enrichment of hypothesised relationships. Resolution of divergences through examination of moderating variable.</td>
<td>Instrument development and pilot testing.</td>
<td></td>
</tr>
</tbody>
</table>
well as non-financial management controls in an US electronics firm using a combination of case study and survey methods. In Merchant (1990), the effects of financial controls on various dysfunctional behaviours were also examined. The survey portion of this part of the study was later extended to a Japanese firm in the same industry to test for the influence of cultural differences between the two settings on the relationship between management controls and dysfunctional behaviour (Chow et al., 1996).

In the initial study, qualitative methods were relatively extensively used prior to the survey for validating existing measurement instruments as well as developing new, more context-specific ones. To the latter end, particular efforts were made to operationalise the different management controls used by the firm. However, great pains were also taken to demonstrate the consistency of these operational definitions with existing control frameworks (see Merchant, 1985). Hence, the new measures do not constitute novel theoretical constructs as such.

Qualitative methods were also used to a more limited extent in combination with relevant literature for generating hypotheses to be tested in the survey. Less extensive, qualitative follow-up procedures were then used to elicit reactions on the first survey. These corroborated some key survey findings but left other, less conclusive results (e.g., concerning the influence of strategy, recent economic performance and leadership style) only partially explained (see Merchant, 1985, 1990). As far as the influence of strategy is concerned, Merchant (1985, 1990) also hinted at possible construct validity threats (i.e., problems of capturing relevant aspects of strategy at the profit centre level) as a cause of inconclusive findings but failed to use the qualitative follow-up procedures to fully resolve these and disentangle them from internal validity issues.

Despite the relatively exploratory character of the first study (see especially Merchant, 1985), it primarily followed a theory testing logic, with qualitative methods mainly being used in the earlier phases and the relevance of key variables being corroborated through some case study-based replication in a second firm (see also Brownell, 1995). The theory testing logic was reinforced in the study by Chow et al. (1996). This study may be considered a partial replication of that of Merchant (1985, 1990) in that identical measurement instruments to those used in the initial study were applied and validated in a different cultural context. However, very different theoretical explanations for the influence of various controls on dysfunctional behaviour grounded in differences in national culture between Japan and the US were sought. Whilst most of the deductively derived a priori hypotheses were supported Chow et al. (1996) made some concessions to the limited opportunities to derive richer explanations of findings contradicting one of these. Hence, similar to the initial study a triangulation approach based on a pronounced theory-testing logic proved less adequate for explicating internal validity issues stemming from the emergence of unexpected findings.


Hoque and Hopper (1994, 1997) study of budgeting in the nationalised jute industry in Bangladesh is a good example of an attempt to address all three validity issues in a systematic manner through between-method triangulation. triangulation was here based on the distribution of a survey to all mills in the industry, covering a number of budgetary control issues emerging from a relatively extensive pilot study in three jute mills as well as higher managerial levels within the industry. Parallel to the survey a more in-depth case study was also conducted in one of the organizations included in the pilot phase.

* However, the reporting of findings from the initial phases of the research was largely confined to the focal case study firm.
An initial exploratory analysis based on a pattern matching approach (cf. Eisenhardt, 1989) revealed considerable consistencies between quantitative and qualitative findings and allowed the authors to draw somewhat more generalizing inferences concerning causal patterns pertaining to the influence of regulatory and financial pressures as well as industrial relations on budgeting (see Hoque and Hopper, 1994). In further refining the hypotheses to be tested, the study also benefited from the combination of a number of well-established bodies of literature, representing technical-rational as well as political-institutional perspectives on the substantive research issue. Apart from allowing the authors to systematically consider the explanatory power of multiple theories at an early stage (see Hoque and Hopper, 1994), this made the largely empirically generated hypotheses, subsequently subjected to more formal tests by Hoque and Hopper (1997) more theoretically informed.

This use of some theoretically informed replication logic contributed to external validation of the findings from the entities studied in greater depth at an early stage. However, Hoque and Hopper (1994, 1997) also relied extensively on between-method triangulation for internal and construct validation. In addition to the procedures described above, internal validity was furthered through follow-up interviews with informants who largely corroborated survey findings. Further, considerable adjustments of the survey instrument were undertaken on the basis of qualitatively inferred insights, including the development of a novel measurement instrument capturing context-specific sources of environmental uncertainty. Whilst some of the factors included in this (e.g., market factors, governmental constraints) conform relatively closely to aspects of uncertainty examined in prior contingency theory research and do not constitute novel constructs per se Hoque and Hopper (1997) admitted that this constrained the possibilities of analytical generalization. Some ex post consideration, informed by insights from the case studies, was also given to potential construct validity threats (pertaining to the complex influence of aid agencies) in the explication of less conclusive survey findings (see Hoque and Hopper, 1997). However, the authors were unable to probe further into the underlying reasons for the lack of support for the relevant hypothesis as a result of not having examined the role of aid agencies in sufficient depth in the earlier phases (see Hoque and Hopper, 1994).

4.1.3. Alam (1997)

Alam (1997) set out by elaborating a theoretical framework integrating contingency and institutional theory explanations of budgetary processes. Two complementary hypotheses pertaining to the relationship between uncertainty and the differing nature of budgeting (i.e., as a functional control mechanism and for dealing with institutional pressures) were thus deduced. These hypotheses were then tested on data drawn from two nationalised industries in Bangladesh where the level of uncertainty was expected to vary significantly. Interview and documentary data were collected at several levels within the two industries to develop and validate context-specific measurement instruments for all key constructs and derive richer information on potential causal relationships. The survey was then administered to a larger sample of operating-level managers within the two industries.

Whilst not testing the direct relationships between uncertainty and the two roles of budgeting, the survey data were used to confirm the existence of significant differences in uncertainty and the use of budgets between the two industries. Qualitative data were then used for corroborating but also deriving richer

It is unclear to what extent the development of measurement instruments was also informed by existing ones. However, the items presented by Alam (1997) correspond relatively closely to aspects of uncertainty and budgeting discussed in the contingency and institutional theory literatures. Hence, these should not be viewed as “new” constructs per se.
explanations of these differences, which were consistent with a priori hypotheses. Hence, the validation of causal explanations in each industry was primarily based on qualitatively inferred insights. However, the systematic comparisons of patterns across the two industries in accordance with some replication logic (cf. Eisenhardt, 1989) also enhanced the external validity of the study.

4.1.4. Kalagnanam and Lindsay (1999)

In contrast to the previously reviewed studies Kalagnanam and Lindsay (1999) adopted a somewhat simpler and more “traditional” theory testing approach relying on three case studies in combination with prior literature as a basis for developing hypotheses pertaining to the contingent relationships between the adoption of just-in-time (JIT) production, organizational structure and performance. These hypotheses were then tested across a larger sample of production plants drawn from a variety of manufacturing industries. The use of between-method triangulation for construct validation mainly took the form of comparisons of the theoretically derived measure of organizational structure with observations from the case studies (coupled with largely converging statistical tests for construct validity). Case study findings were thus used for corroborating the validity of these measures rather than adjusting them to make them more context-specific. Given the adoption of a replication logic and the sampling from a variety of industries such an approach is conducive to deriving theoretically generalisable insights (Lindsay, 1995; Lindsay and Ehrenberg, 1993) whilst minimising the risk of bias due to measurement variability. The use of between-method triangulation for internal validation was mainly confined to inferring causal relationships from the case studies, especially concerning the influence of the technology employed (JIT or mass production) on organizational structure. Although the consistent support for the hypotheses across a broader range of industries suggests that these inferences have some external validity, it is difficult to ascertain to what extent the survey enhanced the internal validity of the overall study given the expanded range of industry-specific factors that might impinge on the causal relationships examined. Kalagnanam and Lindsay (1999) conceded to the limited possibilities of controlling for such potentially confounding factors.

4.1.5. Davila (2000)

Similar to Kalagnanam and Lindsay (1999), Davila (2000) adopted a relatively conventional theory testing approach relying on a number of exploratory case studies in combination with prior literature for hypotheses development followed by a more broadly based survey within the targeted industry (pharmaceutical devices). The main contribution of the case studies lay in the identification of different types of uncertainty associated with product development as “drivers” of the intensity of management control systems use and formulation of hypotheses in this respect. Whilst these findings suggest a causal relationship other hypotheses pertaining to the contingent influence of different product strategies on the intensity of management control systems use and subsequent effects on performance were primarily derived through generalisations from prior contingency theory research. Although the empirically derived hypotheses were largely corroborated by survey findings, the use of between-method triangulation for internal validation was therefore limited. However, the concentration of the study to only one industry and the relatively rich case study descriptions preceding the survey may reduce the risk of spuriousness due to omitted variables causing variations in control systems use.10

10 It is not clear whether the firms and operating units targeted in the case studies were also included in the survey.
Qualitative data (derived from academics as well as project managers) were more extensively used for developing and pilot testing the survey instrument. The insights from the initial case studies appear to have been especially important for elaborating measures reflecting the types of information used for management control and different sources of uncertainty. Although this did not result in entirely “new” constructs, an important aspect of these construct validation efforts was the de-composition of measures of uncertainty into more specific categories to fit the industry-specific context and research issues at hand.

4.1.6. Moores and Yuen (2001)

In contrast to much prior contingency theory research on management accounting, Moores and Yuen (2001) adopted an organizational life-cycle perspective grounded in a systems-based, configurational approach. The design of management accounting systems was examined across different life-cycle stages, denoted by factors such as the age, size, ownership structure, strategy and organizational structure of firms. The random collection of survey data11 from firms within one industry was deepened by simultaneously conducting in-depth interviews with survey respondents in a representative sub-sample of firms in different life-cycle stages selected on the basis of two theoretically relevant parameters (age and size of firms). Survey data were then used for testing hypotheses derived from prior literature (using cluster analysis), whilst qualitative data provided a greater amount of enriching evidence. The latter were used for both corroborating survey findings supporting a priori hypotheses and for refining causal explanations of unexpected relationships.

The triangulation approach described above led to the refinement of some theoretical predictions pertaining to the formalization of management accounting following the transition between different life-cycle stages. However, the study primarily relied on a theory testing logic, as qualitative findings were used for more clearly explicating the rationale for different causal explanations rather than extending and developing these based on competing or complementary theories (cf. Keating, 1995). Overall, the findings were interpreted as consistent with the contingency-based life-cycle approach to the development of management accounting. The primary merits of the triangulation approach adopted by Moores and Yuen (2001) thus seem to lie in internal rather than external validation, as it deepened the quest for causal explanations within representative sub-groups of the survey sample. This may be especially fruitful in this particular study given the complex task of interpreting causal relationships in research relying on a configurational approach coupled with cluster analysis (Gerdin and Greve, 2004).

4.1.7. Summary

Consistent with our prior discussion of different validity criteria, triangulation approaches primarily relying on a theory testing logic mainly use qualitative methods for hypotheses development and elaboration and pilot testing of measurement instruments before conducting large-scale survey-based tests. This places an emphasis on external and construct validation but also implies that the replication logic adopted is more akin to the pattern matching approach advocated by Eisenhardt (1989) than generalisation based on strict replications (cf. Lindsay and Ehrenberg, 1993). The occasional incorporation of a quantitative element into such pattern matching for ascertaining the more general plausibility of causal explanations within a well-defined empirical setting at an early stage (Alam, 1997; Hoque and Hopper, 1994) constitutes an interesting development of this approach. It effectively implies more integrative use

---

11 This is the only study of those reviewed using a more strictly randomised sample. However, a low response rate (14.5%) was reported.
of between-method triangulation for internal and external validation, sometimes before conducting more formalized hypothesis tests (see Hoque and Hopper, 1994, 1997). However, internal validation efforts are generally limited to the corroboration of converging findings within reasonably similar empirical settings (generally the same industry). More extensive use of qualitative methods for also enriching and refining causal explanations is mainly found where surveys are combined with more systematic follow-up procedures (i.e., Alam, 1997; Moores and Yuen, 2001). As demonstrated next, more far-reaching theory extensions and elaboration of theoretical constructs are found where a theory development logic dominates triangulation.

4.2. Group 2: primarily theory development

4.2.1. Covaleski and Dirsmith (1983)

Covaleski and Dirsmith (1983) set out with a fairly well-established conceptual framework, deductively derived from predominantly functionalist and instrumental theories of budgeting as a tool for managerial control, to assess how nurse managers use budgeting. A modified version of an existing survey instrument to this end was used across a relatively broadly based sample (representatives of 41 hospitals). However, initial survey findings largely failed to correspond to such a traditional view of budgeting. This prompted the authors to decompose the measurement instrument into sub-items reflecting the use of budgeting for internally orientated control purposes and political advocacy, respectively. Re-testing the revised model across a smaller sub-sample of nurse managers and conducting in-depth interviews with each of these, some alternative explanations of the failure to corroborate the initial model emerged. Consistent with emergent theories stressing the more symbolic roles of budgeting (e.g., institutional theory), the hierarchical position of nurse managers and the formalization of their work roles proved important in explaining the use of budgets as means of loose coupling in addition to more traditional control purposes. This resulted in a set of revised hypotheses informed by such alternative theories. The emphasis of the research thus shifted from a theory testing to a theory development logic with important insights being inferred from qualitative data (see also Covaleski and Dirsmith, 1990).

This study provides an example of how close iterations between quantitative and qualitative methods to resolve (ex post) an emergent construct validity issue pertaining to different meanings of budgeting may be extended to shed light on unexpected causal relationships explaining such differences. The latter corroborated the influence of certain independent variables (i.e., hierarchical position and formalization of work roles) impinging on budgeting. This also resulted in some external validation through extensions of existing theories of budgeting. In addition, this emergent theory was partly corroborated through dialogues with nurse managers in conjunction with training sessions. In presenting their insights, however, Covaleski and Dirsmith (1983) relied more on prior theorizing whilst displaying only limited portions of their seemingly rich qualitative data.

4.2.2. Modell and Lee (2001)

Modell and Lee (2001) examined the influence of decentralization of decision-making on reliance on the controllability principle (RCP) in conjunction with budgetary control at the departmental level of a large hospital. Prior research on RCP, primarily conducted in the private sector, was used for developing an a priori hypothesis. Whilst an alternative perspective relying on institutional theory also informed this discussion, the study assumed a strongly exploratory character making extensive use of interview data collected prior to and parallel to the distribution of the survey to probe into the influence of institutional
factors. Interview data were also used for developing a novel survey instruments (for measuring RCP) and pilot-testing the questionnaire in its entirety.

More importantly, however, qualitative data were used for developing a richer, institutionally informed explanation of the weak direct relationship between decentralization and RCP emerging from the survey. A number of internal validation techniques were applied in this respect. First, interviews and documents were used for advancing some explanations of how the responses of senior hospital management to recent reforms in the institutional field concerned had impinged on RCP within the hospital. Second, a systematic multi-level approach (cf. Sieber, 1973; Tashakkori and Teddlie, 1998) was adopted by comparing responses to an open-ended survey question at the departmental level with interview data from higher echelons to assess the causes of the perceived controllability of specific financial items. This largely corroborated the constraining effect of various institutional factors (e.g., labour market regulations, power struggles) on the possibilities of decentralization and RCP.

These triangulation efforts produced some tentative causal explanations, pivoting around how (previously ignored) institutional factors may mediate the relationship between decentralization and RCP. Some extensions pertaining to the constraining role of institutional factors and the scope for managerial choice were thus added to the literature on RCP. A limitation of the study, however, was that the revised model was neither presented in the form of refined hypotheses nor subjected to further statistical tests aimed at corroboration.

4.2.3. Summary

In comparison with most research in group 1, the studies in group 2 display more integrative use of between-method triangulation for internal and external validation based on extensions of existing theory in explaining causal relationships. In both studies, theory triangulation played a critical role in this respect. However, only Covaleski and Dirsmith (1983) come close to the “full sequence” of not only using between-method triangulation for discovering alternative explanations but also corroborating these by re-testing the revised theoretical model across a smaller sub-sample. This study is also unusual for its relatively far-reaching use of between-method triangulation for ex post resolution of construct validity issues (cf. Brewer and Hunter, 1989).

4.3. Group 3: balanced emphasis on theory testing and development

4.3.1. Hopwood (1973)

Drawing on prior literature on budgetary control, organisational sociology and psychology, Hopwood (1973) initially developed a conceptual framework to examine how different budget-related performance evaluation styles impinged on job-related tension and subordinate managers’ relations with superiors and peers as well as different dysfunctional behaviours. Whilst using existing measurement instruments for most dependent variables, extensive efforts went into developing the key independent variables (different performance evaluation styles) based on a primarily interview-based pilot study. The closeness to the organisation under study and access to quantitative budget data also enabled Hopwood to operationalise a moderating factor (the level of inaccuracy of the accounting system) which was arguably difficult to

12 The choice of an inductive, qualitative approach for further examining the influence of institutional factors was itself theoretically motivated by the observation that the exact properties of institutional environments are difficult to define a priori but need to be identified empirically (DiMaggio and Powell, 1983).
capture with the survey instrument. Similarly, qualitative data were used for testing hypotheses pertaining to topics less susceptible to survey-based data collection (e.g., illicit and dysfunctional budget-related behaviours) in the in-depth case study following the survey.

Qualitative data were not extensively used for developing hypotheses, however, but primarily for corroborating and enriching survey findings (e.g., the consequences of different performance evaluation styles and the existence of contagion effects across hierarchical levels in this respect). Between-method triangulation was also used for deriving alternative causal explanations of unexpected survey findings. For example, the moderating effect of budgetary participation was examined for explaining the unexpected direct relationships between inaccuracies of the accounting system and subordinate tension. First, the analysis of survey data was extended by considering how the effects of such inaccuracies varied across different levels of budgetary participation under different performance evaluation styles. Second, interview data from several managerial levels were used for examining the underlying processes explaining the role of budgetary participation in this respect. This highlighted the role of budgetary participation in exchanging information as a potential cause of reduced tension.

A pronounced multi-level triangulation approach was also utilised in examining the existence of contagion effects through which different performance evaluation styles were transmitted from superiors to subordinates. Survey findings provided strong support for such an effect whilst interviews conducted at multiple managerial levels provided richer explanations of why these occurred (cf. Sieber, 1973; Tashakkori and Teddlie, 1998). Before conducting this analysis, however, quantitative methods were used to address a potential construct validity threat related to disagreements among some subordinate managers over the performance evaluation style used by their joint superior emerging from the survey. This indicated that such variations were due to differences in whether subordinate managers met their budgets and the size of their departments rather than properties of the measurement scale used for assessing performance measurement styles.

Hopwood (1973) thus relied extensively on between-method triangulation for internal and construct validation, setting out with a theory testing approach. However, the rich empirical insights stemming from the use of triangulation for refining as well as extending the rudimentary understanding of the use of budgets for performance evaluation implied that the study also came to encompass a pronounced theory development logic.


The series of studies conducted by Otley provides an illuminating example of the role of between-method triangulation when operating within a broader, gradually evolving research programme. Otley (1978) initially set out to replicate Hopwood’s (1972, 1973) findings regarding the effects of budget-related performance evaluation styles on a range of behavioural variables (e.g., job-related tension, job ambiguity) in a different context. He also extended the study to examine the effects of evaluation styles on managerial performance. Whilst the study was designed as a relatively close replication, some adjustments of the measure of evaluation styles were made to better fit the profit-centre context in which the study was set. Survey findings failed to corroborate those of Hopwood (1972), however, and in attempting to explain these divergences Otley (1978) turned to qualitative data collected parallel to the distribution of questionnaires. In particular, the finding that high reliance on budgetary information for performance evaluation was negatively related to adverse outcomes was thus explained with reference to environmental and economic differences drawing, in part, on emerging contingency theory research. The emphasis thus
shifted from a replication-based theory testing logic to theory development through an attempt to extend causal explanations.

A similar shift in emphasis is discernible in Otley’s (1990) attempt to replicate the original study in the same organization 15 years later. Initial survey findings indicated that performance evaluation styles had converged considerably over time such that context-specific differences within the organization were less noticeable (Otley, 1990; Otley and Berry, 1994). Using qualitative data and contrasting findings with those of the original study (Otley, 1978), a number of causal explanations for this convergence, pertaining to the influence of changing competitive conditions and the length of managerial tenure, were advanced (Otley, 1990). However, this attempt to extend internal validation efforts by inferring more generalizing conclusions over time was partly hampered by the relatively loose connections between empirical insights and existing theory (see also Otley and Berry, 1994).

Finally, Otley and Pollanen (2000) attempted to resolve some conflicting results of a range of studies extending the works of Hopwood (1972) and Otley (1978). This mainly survey-based study aimed at replicating previous work in a different contextual setting (Canadian universities) as closely as possible. Whilst great pains were taken to maintain some stability in measurement instruments,13 some adjustments were still prompted to preserve some sensitivity to context-specific, environmental factors. Further, the authors made some concessions to how the breadth of their sampling procedures limited their ability to gain a deeper understanding of their relatively mixed results and called for more in-depth, case-based attempts to probe causal explanations of unresolved issues.

This series of studies illustrates the difficulties in conducting close replications in management accounting research but also points to potential advantages of combining such attempts with an element of between-method triangulation. The latter may facilitate theory extensions stimulated by the failure to replicate results (cf. Lindsay and Ehrenberg, 1993) and thus produce a more balanced mix of theory testing and development logics. Whilst such a development is discernible in Otley’s works, the primary intention behind these seems to have been to enhance external validity by conducting replications within an increasingly well-established research programme. This may be an explanation for the constrained use of between-method triangulation for extending the probing of internal validity issues into better developed theoretical insights. Further, whilst great efforts were made to maintain some stability in measurement instruments the studies illustrate the need to adjust these to reduce construct validity threats stemming from context-specific variations which in turn exacerbates close replications (Lindsay and Ehrenberg, 1993).


The research programme established by Simons (1987a, 1987b, 1990, 1991, 1994) presents an interesting contrast to the one stemming from Hopwood’s and Otley’s work in its predominantly qualitative emphasis encompassing a combined replication and theory extension logic (cf. Eisenhardt, 1989). Although between-method triangulation assumed a relatively limited role in this research it is instructive to consider how it may be incorporated into a coherent research programme dominated by qualitative methods.

This research gradually extended and refined theoretical explanations of senior managers’ use of management control systems. Two distinct uses of control systems (interactive versus programmed) were initially identified in a case study in one firm and led to some reconsideration of prior theory on the role of

---

13 This task was further complicated by the lack of consistency in the use of measurement instruments in prior studies (see also Otley and Fakiolas, 2000).
formal control systems in the face of different strategic contingencies (Simons, 1987a). The results of a survey across a broad range of industries were interpreted as being largely consistent with these inductively inferred results (Simons, 1987b). The insights from these early studies were more closely examined in one industry following a comparative case study approach and resulted in a set of testable hypotheses pertaining to managerial use of control systems (Simons, 1990, 1991). The emerging theoretical model was finally further extended through the discovery of two additional controls (boundary and beliefs systems) in an intensive qualitative study of a limited number of top managers (Simons, 1994). However, the resulting theoretical model was not subjected to survey-based tests.

Between-method triangulation was primarily used by Simons (1987b) for addressing construct and internal validity issues. Measurement instruments mostly derived from prior literature were partly validated based on qualitative data. Interviews were also conducted with a representative sub-sample of survey respondents to deepen the analysis of causal explanations of the use of management control systems. These were mainly used for corroborating some survey findings (e.g., concerning the use of external scanning and forecasts and the tailoring of control systems to users’ needs) but only partly for advancing explanations of divergences (e.g., concerning results monitoring). This triangulation approach thus resembles that characterising research primarily following a theory testing logic. However, some of the overriding insights regarding causal relationships from the early studies, especially that the overall emphasis on diagnostic versus interactive control differs with the choice of strategy (Simons, 1987a, 1987b), also informed the pattern matching efforts in the comparative case study phase (see especially Simons, 1990). This latter phase both enriched and extended initial findings and thus resulted in a more balanced mix of theory testing and development enhancing the external validity of emerging causal relationship.


Finally, a number of studies of the implementation of activity based costing (ABC) systems in the US automobile industry deserve recognition. This series started with an exploratory case study (Anderson, 1995) followed by a predominantly survey-based research phase enriched with qualitatively inferred insights into firm-specific factors influencing implementation (Anderson and Young, 1999; Anderson et al., 2002). The research programme as a whole was summarised in a research monograph (Anderson and Young, 2001).

Anderson (1995) initially explored the implementation of ABC in General Motors, whilst augmenting the sparse and largely anecdotal literature on this substantive issue with theoretical insights gleaned from the information systems and organizational change literatures. As empirical research on the topic grew, the extension of this qualitative study to a competitor (Chrysler) was accompanied by a systematic comparison of variables hypothesised to influence ABC implementation in a range of other empirical studies with those inductively derived by Anderson (see Anderson and Young, 1999). Apart from demonstrating the exhaustiveness of the constructs generated by Anderson (1995) and thus increasing the chances of subsequent findings reflecting patterns of more general relevance, this facilitated the achievement of good fit between theoretical constructs and (existing as well as novel) measurement instruments further developed in the later survey- and interview-based study (Anderson and Young, 1999; Anderson et al., 2002).

Whilst close iterations between prior research and qualitative data thus took place for developing and validating constructs, the more specific causal models and hypotheses tested in the later research phase

14 Whilst the authors claim to have observed many converging patterns in the initial qualitative research in Chrysler, the formal elaboration of this framework was primarily based on the General Motors case (see Anderson and Young, 2001).
were largely derived from more firm-specific, inductively inferred insights. This exploratory approach was most pronounced in Anderson and Young (1999), where survey data were used for assessing the validity of a broadly defined model rather than specific hypotheses. Internal validity was also assessed through quantitative analyses of the stability of causal explanations across the two companies, various categories of respondents and older and more recent ABC implementations. However, despite repeated references to extensive interview data collected after the survey, little explicit use was made of these for providing a richer understanding of causal relationships ex post. Hence, the use of between-method triangulation for internal validation, apart from demonstrating the overall significance of qualitatively generated variables in explaining implementation outcomes, was relatively limited. This is especially problematic in the case of Anderson et al. (2002) where several hypotheses pertaining to the influence of implementation teams on ABC outcomes received no or limited support from survey data.

The triangulation approach adopted in this research programme has some resemblance to a theory testing logic primarily relying on qualitative methods for developing measurement instruments and, to a more limited extent, hypotheses. However, the systematic extensions and integration of emerging theoretical perspectives and the relatively exploratory approach to assessing the causal direction of relationships between constructs led to a more balanced approach encompassing an element of theory development (see especially Anderson and Young, 1999). Yet, it is possible that this theory development logic could have been more pronounced, had qualitative data been more extensively used for enriching causal explanations.

4.3.5. Summary

In contrast to most research in groups 1 and 2, triangulated management accounting research displaying a more balanced emphasis on theory testing and development is generally rooted in more coherent research programmes. This typically takes the form of multiple iterations between research phases making extensive use of both qualitative and quantitative methods. This enables a broader range of internal validity issues to be addressed within the same study or series of studies (Bryman, 1992; Lee, 1991). There is relatively common use of between-method triangulation for corroborating and enriching causal explanations as well as explaining divergences related to inconclusive or unexpected results (see especially Hopwood, 1973; Simons, 1987a, 1987b, 1990), although internal validation efforts are occasionally hampered by insufficient use of qualitative methods (Anderson and Young, 1999; Anderson et al., 2002). This is in turn mirrored by the broader range of external validation tactics in group 3 than in the two other groups (see Table 1). In comparison with the studies in group 2, for example, theory extensions are not only based on ex post examination of internal validity issues but are frequently combined with an element of hypothesis development or replication.

5. Conclusions

Our review reveals how variations in triangulation approaches impinge on the validation of empirical findings. This contributes to nuance the rather general and unproblematic claims concerning the merits of triangulated research in the management accounting literature (see e.g., Abernethy et al., 1999; Birnberg...
et al., 1990). More careful consideration should be given to the specific validity threats likely to emerge, given the choice of a particular triangulation approach and how these might be mitigated.

A useful starting point for such considerations is to recognize that the types of validation efforts involved in triangulated research can vary depending on whether theory testing or development is the principal aim. Where a theory testing logic dominates (group 1 above), triangulation efforts tend towards external and construct validation through the development of hypotheses and/or measurement instrument applied in survey-based research. This is a suitable approach where prior theorizing provides reasonably well-established constructs and a solid understanding of causal relationships albeit some modifications of theory may be required depending on context-specific factors. Research may thus entail elements of a replication logic facilitating the process of analytical generalisation (cf. Eisenhardt, 1989; Lindsay, 1995). However, internal validation is often limited to the corroboration of converging findings between different methods. Whilst some examples of qualitative methods also being used for enriching or refining causal explanations were found, researchers relying on a pronounced theory testing logic seem less inclined to rely on these for radically extending theory by elaborating alternative explanations of unexpected or inconclusive survey findings.

By contrast, the studies primarily aiming at theory development (group 2 above) originate from failures to corroborate particular theoretical models or hypotheses and use triangulation as a means of extending established theory based on the examination of alternative or complementary causal relationships. Such internal validation efforts typically take the form of ex post probing of survey results. This may also enable researchers to examine potential construct validity threats emerging after the distribution of surveys (see Covaleski and Dirsmith, 1983). Whilst this gives research a more exploratory flavour, theory extensions based on further examination of inconclusive findings are likely to benefit from extensive reliance on theory triangulation. This contributes to some external validation through analytical generalisation and adds a theoretically informed explanatory element to empirical research (cf. Humphrey and Scapens, 1996; Ryan et al., 2002).

The discussion above is indicative of the critical role of the sequence in which different methods are applied in triangulated research (Brannen, 1992; Creswell, 1994; Tashakkori and Teddlie, 1998). Although all three validity criteria may be addressed to varying extents irrespective of whether a theory testing or development logic dominates, these logics tend to be associated with a specific sequence facilitating certain types of validation efforts. For example, most studies in group 1 concentrate the use of qualitative methods to the early phases of research, whilst the studies in group 2 mainly invoke qualitatively informed explanations after the survey. As noted above the main contribution of the former approach lies in testing and corroborating causal relationships whilst the latter seems more conducive to extending theory by examining inconclusive survey findings (see also Brannen, 1992).

This distinction suggests that triangulation approaches relying on closer interweaving of case study and survey methods in multiple iterations stand a better chance of addressing a broader range of validity issues (cf. Bryman, 1992; Lee, 1991). Indeed, this is the case in research relying on a more balanced emphasis on theory testing and development (group 3 above). Whilst this group of studies includes a broad spectrum of approaches, most of them share a concern with using between-method triangulation for addressing internal validity issues pertaining to both converging and diverging (or unexpected) findings. This, in turn, gives rise to a broader range of external validation efforts combining elements of a replication and theory extension logic. It also underscores the merits of locating multiple iterations between qualitative and quantitative methods within an evolving or established research programme (Birnberg et al., 1990; Lindsay, 1995). Specifically, this may facilitate the process of analytical generalisation back to existing and
emerging theories irrespective of whether empirical findings initially converge or diverge (cf. Abernethy et al., 1999). However, this needs to be weighed against the considerable time and resources required for building up more extensive research programmes.

Whilst some progress in the use of between-method triangulation has thus been made in empirical management accounting research, further development is required in a number of respects. Despite the emergence of some clearly identifiable research programmes, there would still seem to be considerable scope for elaborating triangulation approaches combining replication and theory extension logics in more equal measure (cf. Brewer and Hunter, 1989; Lindsay, 1995). Further, as far as internal validation is concerned, qualitatively inferred explanations of inconclusive or unexpected survey findings are rarely corroborated through additional survey-based tests in the same empirical settings before theory extensions are specified (but see Covaleski and Dirsmith, 1983). Such extensions are mainly limited to analytical generalisations, with researchers primarily relying on qualitative methods in combination with theory triangulation (see also Bryman, 1992). Finally, the use of between-method triangulation for construct validation is mostly confined to instrument development and pilot testing prior to the distribution of surveys, which may testify to the complexity and costs of ex post construct validation (Brewer and Hunter, 1989). However, the latter might be potentially rewarding, especially from a theory development perspective (see Covaleski and Dirsmith, 1983).

A limitation of our review is the lack of attention to methodological issues arising from the combination of between-method triangulation with theoretical perspectives grounded in diverging epistemologies (see Lee, 1991). Whilst most of the studies reviewed are based on a positivist or functionalist paradigm, embedded in more or less explicit applications of contingency theory, some attempts at triangulation with theories with stronger interpretive and social constructivist underpinnings (e.g., institutional theory) can be observed (Alam, 1997; Covaleski and Dirsmith, 1983; Hoque and Hopper, 1994, 1997; Modell and Lee, 2001). This is a defensible position if one accepts that pragmatic method and theory choices may achieve a certain degree of independence from their epistemological roots (Bryman, 1992; Hoque and Hopper, 1997). However, it has stood far from uncontested (see, e.g., Blakie, 1991; Sale et al., 2002). Applying validity criteria originating from the positivist research tradition would nevertheless seem acceptable as the studies cited above are clearly concerned with suggesting plausible causal relationships. Further, the kind of theorizing invoked persistently testifies to concerns with analytical generalisation. Although we recognize the critique that may be levelled at our position debating subsequent epistemological ramifications is beyond the scope of this paper.

Acknowledgements

The author is grateful for helpful comments on earlier versions of the paper at the 26th Annual Congress of the European Accounting Association, Seville, the 7th Interdisciplinary Perspectives on Accounting conference, Madrid, 2003, and workshops held at Jönköping International Business School, the Royal Institute of Technology and Curtin University of Technology.

References


