RESEARCH NOTES AND COMMENTARIES

WHAT PASSES AS A RIGOROUS CASE STUDY?

MICHAEL GIBBERT,[1]* WINFRIED RUIGROK,[2] and BARBARA WICKI[3]
[1] *Business Management Department, Bocconi University, Milan, Italy*
[2] *University of St. Gallen, Switzerland*
[3] *Credit Suisse Business School, Zurich, Switzerland*

This article investigates the methodological sophistication of case studies as a tool for generating and testing theory by analyzing case studies published during the period 1995–2000 in 10 influential management journals. We find that case studies emphasized external validity at the expense of the two more fundamental quality measures, internal and construct validity. Comparing case studies published in the three highest-ranking journals with the other seven, we reveal strategies that may be useful for authors wishing to conduct methodologically rigorous case study research. Copyright © 2008 John Wiley & Sons, Ltd.

INTRODUCTION

Case studies as tools for generating and testing theory have provided the strategic management field with ground-breaking insights (e.g., Penrose, 1960; Chandler, 1962; Pettigrew, 1973; Burgelman, 1983). Despite this, the case study method has been prone to concerns regarding methodological rigor in terms of validity and reliability (e.g., Campbell, 1975; Miles, 1979; Daft and Lewin, 1990; March, Sproull, and Tamuz, 1991; Yin, 1981). While deficiencies in any methodology are problematic (Bergh, Perry, and Hanke, 2006), lacking rigor in case studies is particularly

problematic for at least two reasons. First, case studies are considered most appropriate as tools in the critical, early phases of a new management theory, when key variables and their relationships are being explored (Yin, 1994; Eisenhardt, 1989). A rigor problem in the early stages of theory development would therefore have ripple effects throughout later stages when relationships between variables are elaborated and tested (Eisenhardt and Graebner, 2007). Second, case studies are typically carried out in close interaction with practitioners, and they deal with real management situations. Case studies therefore represent a methodology that is ideally suited to creating managerially relevant knowledge (Amabile *et al*., 2001; Leonard-Barton, 1990). However, as Scandura and Williams (2000) have reminded us, 'without rigor, relevance in management research cannot be claimed' (Scandura and Williams, 2000: 1263).

The purpose of this research note is to (a) establish if there have been any systematic tendencies in case study research published in leading management journals that may negatively affect case study rigor, and (b) to propose research strategies to enhance rigor. Drawing on the framework developed by Cook and Campbell (1976, 1979), and later adapted to the case study methodology by Eisenhardt (1989), and Yin (1994), we reveal how case study researchers discussed the relevant procedures associated with internal validity, construct validity, external validity, and reliability of all case studies published over six years (1995–2000) in 10 leading journals.

VALIDITY AND RELIABILITY CRITERIA IN CASE STUDY RESEARCH

In management and strategy research, interest in the case study as a method for generating and testing theory has recently seen a revival (e.g., Eisenhardt and Graebner, 2007; Siggelkow, 2007; Weick, 2007), after it has lain in limbo for more than 20 years (e.g., Campbell, 1975; Miles, 1979; Daft and Lewin, 1990; March *et al.*, 1991; Yin, 1981). Case studies have been defined as 'research situations where the number of variables of interest far outstrips the number of datapoints' (Yin, 1994: 13). While case studies may, and often do, use quantitative data, a key difference with other research methods is that case studies seek to study phenomena in their contexts, rather than independent of context (e.g., Pettigrew, 1973).

There are numerous criteria to assess the rigor of field research, including case studies. These criteria hinge on what authors subscribe to as the preferred model of science. The model used in this article lies within the positivist tradition (e.g., Behling, 1980; Cook and Campbell, 1979).[1] In the positivist tradition, four criteria are commonly used to assess the rigor of field research: internal validity, construct validity, external validity, and reliability (Campbell and Stanley, 1963; Campbell, 1975).

These criteria have been adapted for use in case studies by Yin (1994), and others (e.g., Campbell, 1963; Eisenhardt, 1989). Table 1 provides an overview of the four validity and reliability criteria (the 'primary reports' the horizontal axis, drawing on Cook and Campbell, 1979), and also gives the research measures or actions that case study researchers may take for each criterion (the 'secondary reports,' along the vertical axis, drawing on Yin, 1994).

Internal validity

'Internal validity' is also called 'logical validity' (e.g., by Cook and Campbell, 1979; Yin, 1994) and refers to the causal relationships between variables and results. Here, the issue is whether the researcher provides a plausible causal argument, logical reasoning that is powerful and compelling enough to defend the research conclusions. Internal validity refers to the data analysis phase (Yin, 1994: 105). Three measures have been proposed to enhance internal validity. First, case study researchers should formulate a *clear research framework*, which demonstrates that variable x leads to outcome y, and that y was not caused spuriously by a third variable z. Second, through *pattern matching,* researchers should compare empirically observed patterns with either predicted ones or patterns established in previous studies and in different contexts (Denzin and Lincoln, 1994; Eisenhardt, 1989). Third, *theory triangulation* enables a researcher to verify findings by adopting multiple perspectives (Yin, 1994).

Construct validity

The 'construct validity' of a procedure refers to the quality of the conceptualization or operationalization of the relevant concept. Construct validity needs to be considered during the data collection phase. As such, construct validity refers to the extent to which a study investigates what it claims to investigate, that is, to the extent to which a procedure leads to an accurate observation of reality (Denzin and Lincoln, 1994). It has been said that case study researchers sometimes do not develop a well-considered set of measures and that 'subjective' judgments are used instead (Yin, 1994: 41). In order to enhance construct

---

[1] We adopted the positivist model for the purpose of the present study because it underlies the most respected and extensively cited text for conducting case study research, Yin (1994). As such, the present survey may be appreciated by those subscribing to the positivist tradition. In addition to this, scholars who are unfamiliar with, or critical of, the positivist tradition may also appreciate a review of the level of rigor in case-based research conducted according to this model.

Table 1. Framework for an investigation of the methodological rigor of case studies

Internal validity	Construct validity	External validity (Cook and Campbell, 1979)	Reliability
Research framework explicitly derived from literature (diagram or explicit description of causal relationships between variables and outcomes)	Data triangulation • Archival data (internal reports, minutes or archives, annual reports, press or other secondary articles)	Cross case analysis • Multiple case studies (case studies of different organizations)	Case study protocol (report of there being a protocol, report of how the entire case study was conducted)
Pattern matching (matching patterns identified to those reported by other authors)	• Interview data (original interviews carried out by researchers)	• Nested approach (different case studies within one organization)	Case study database (database with all available documents, interview transcripts, archival data, etc.)
Theory triangulation (different theoretical lenses and bodies of literature used, either as research framework, or as means to interpret findings)	• Participatory observation derived data (participatory observation by researchers)	Rationale for case study selection (explanation why this case study was appropriate in view of research question)	Organization's actual name to be mentioned explicitly – as opposed to anonymous)
	• Direct observation derived data (direct observation by researchers)	Details on case study context (explanation of e.g., industry context, business cycle, P/M combinations, financial data)	
	Review of transcripts and draft by peers (peers are academics not co-authoring the paper) Review of transcripts and draft by key informants (key informants are or have been working at organization investigated) Clear chain of evidence Indication of data collection circumstances (explanation how access to data has been achieved) Check for circumstances of data collection vs. actual procedure (reflection of how actual course of research affected data collection process) Explanation of data analysis (clarification of data analysis procedure)		

(Yin, 1994)

validity in case studies, two measures have crystallized. First, researchers have been encouraged to *establish a clear chain of evidence* to allow readers to reconstruct how the researcher went from the initial research questions to the final conclusions (Yin, 1994: 102). Second, researchers have sought to *triangulate*, that is, adopt different angles from which to look at the same phenomenon, by using different data collection strategies and different data sources (Denzin and Lincoln, 1994; Yin, 1994).

External validity

'External validity,' or 'generalizability,' is grounded in the intuitive belief that theories must be shown to account for phenomena not only in the setting in which they are studied, but also in other settings (e.g., Calder, Phillips, and Tybout, 1982; McGrath and Brinberg, 1983). Neither single nor multiple case studies allow for statistical generalization, for example, inferring conclusions about a population (Yin, 1994: 31; Numagami 1998: 3). This does not mean, however, that case studies are devoid of generalization. Methodologists differentiate between statistical generalization and analytical generalization. Analytical generalization is a process separate from statistical generalization in that it refers to the generalization from empirical observations to theory, rather than a population (e.g., Yin, 1994, 1999). How, then, can case studies allow for analytical generalization? In her widely cited paper, Eisenhardt (1989) argues that case studies can be a starting point for theory development and suggests that a *cross-case analysis* involving four to 10 case studies may provide a good basis for analytical generalization. Instead of conducting and analyzing multiple case studies of different organizations, researchers may also conduct different case studies *within* one organization (a *nested approach*, e.g., Yin, 1994). Finally, researchers should provide a clear *rationale for the case study selection*, and ample details on the *case study context* to allow the reader to appreciate the researchers' sampling choices (Cook and Campbell 1979: 83).

Reliability

'Reliability' refers to the absence of random error, enabling subsequent researchers to arrive at the same insights if they conduct the study along the same steps again (Denzin and Lincoln, 1994). The key words here are transparency and replication. Transparency can be enhanced through measures such as careful documentation and clarification of the research procedures, for example, by producing a *case study protocol* — a report that specifies how the entire case study has been conducted. Replication may be accomplished by putting together a *case study database*, which should include the case study notes, the case study documents, and the narratives collected during the study, organized in such a way as to facilitate retrieval for later investigators (Yin, 1994), that is, to facilitate the replication of the case study (e.g., Leonard-Barton, 1990).

Thus, the methodology literature has proposed a number of research actions, or 'measures' to address validity and reliability concerns. Since methodological rigor is not a random attribute, we postulate that researchers who apply measures to deal with reliability, that is, the absence of random error, are more likely also to take measures to deal with the other three validity concerns, which deal with the absence of nonrandom or systemic error.

Importantly, the three validity types are not independent of each other. Without a clear theoretical and causal logic (internal validity), and without a careful link between the theoretical conjecture and the empirical observations (construct validity), there can be no external validity in the first place (Cook and Campbell 1979: 83; Mook, 1983: 379; Scandura and Williams, 2000: 1261). Thus, there is a hierarchical relationship of validity types, with construct and internal validity acting as a *conditio sine qua non* for external validity.

One might expect that review and selection practices at the highest-ranked journals are more likely to publish case studies in which the authors demonstrate awareness of the hierarchical relationship of validity and reliability types. Therefore, we hypothesize that case studies published in the highest-ranked journals will be more likely than those published in the other journals in the sample to make explicit mention of validity and reliability types ('primary reports') as well as outlining the actual research measures taken to enhance each validity and reliability type ('secondary reports') on the two most fundamental validity issues (internal and construct validity).

Hypothesis 1: Case studies that provide secondary reports on reliability are more likely also to provide secondary reports on construct validity, internal validity, and external validity.

Hypothesis 2a: Case studies published in the highest-ranked journals are more likely to provide primary reports on internal validity than papers published in lower-ranked journals.

Hypothesis 2b: Case studies published in the highest-ranked journals are more likely to provide secondary reports on internal validity than papers published in lower-ranked journals.

Hypothesis 3a: Case studies published in the highest-ranked journals are more likely to provide primary reports on construct validity than papers published in lower-ranked journals.

Hypothesis 3b: Case studies published in the highest-ranked journals are more likely to provide secondary reports on construct validity than papers published in lower-ranked journals.

Although the three types of validity build on each other, several forces may actually lead case study researchers to pay more attention to external validity than to the other three concerns. In particular, case study researchers are exhorted in the literature to pay special attention to this criterion. Eisenhardt's (1989) influential paper argued how (multiple) case studies could be used to build theories and enhance external validity. March, Sproull, and Tamuz have provided guidelines for 'learning from samples of one or fewer' (March *et al.*, 1991: 1). More recently, Rouse and Daellenbach (2002) argued that in some cases only 'intrusive' research can help to 'enhance our understanding of both the generalizability and managerial implications' (2002: 965) of results previously obtained through secondary data research. Other authors cautioned that generalization cannot and should not be emphasized in all research, because it may draw researchers' attention away from understanding the case itself and lead to overgeneralization (Mir and Watson, 2000). However, these cautionary comments further underline the bias toward external validity. Put together, such forces may lead case study researchers to emphasize external validity over other validity or reliability measures.

Hypothesis 4a: Case studies are more likely to provide primary reports on external validity than on construct validity, internal validity and reliability.

Hypothesis 4b: Case studies are more likely to provide secondary reports on external validity than on construct validity, internal validity, and reliability.

METHODOLOGY

Sampling

We selected case studies that were published over a recent six-year period (1995–2000) in 10 leading management journals: *Academy of Management Journal, Administrative Science Quarterly, California Management Review, Journal of International Business Studies, Journal of Management Studies, Long Range Planning, Organization Science, Organization Studies, Sloan Management Review*, and *Strategic Management Journal*.

To select these journals, we used the Tahai and Meyer ranking (Tahai and Meyer, 1999). To identify the highest-ranked journals in this sample, we also followed Tahai and Meyer (1999). These authors reported a substantial influence gap between *Administrative Science Quarterly* and *Organization Science*, although these would appear quite close in the ranking (number 6 and 8, respectively, Tahai and Meyer, 1999: 291). Thus, we chose *Academy of Management Journal* (number 2), *Administrative Science Quarterly* (number 6), and *Strategic Management Journal* (number 1) as the 'top three.' To identify relevant articles in these journals, we used the definition of case studies provided earlier, that is, empirical papers that reported results based on primary fieldwork in one or more for-profit organizations, in which no experimental controls or manipulation were involved, and which used multiple sources of data. This exercise produced a total of 159 case studies published in the aforementioned journals over the period 1995–2000.[2]

---

[2] Full references of these case studies are available upon request from the corresponding author. Initially we had also selected *Journal of Management* and *Management Science*, ranked number 7 and 13 respectively (Tahai and Meyer, 1999: 291). While we identified one study, which the authors themselves labeled

Coding

Case studies' methodological rigor was analyzed by 23 coding dimensions involving both primary reports (i.e., explicit mention of the measure), as well as secondary reports (i.e., an explication of the concrete research measures taken, see Table 1). Furthermore, we also coded for the use of measures not included in the 23 coding dimensions, especially measures used by authors with an interpretivist or constructionist stance (e.g., Lincoln and Guba, 1985, Glaser and Strauss, 1967), yet we did not find any such reports. A 'multiple-rater' approach was used (following Larsson, 1993; Scandura and Williams, 2000). Pre-consensus coding interrater reliability, measured as average pairwise percent agreement of coding across raters (Larsson 1993) was 0.97. We also carried out tests using a composite variable for construct validity, internal validity, external validity, and reliability. These composite variables refer to case studies that, according to our consensus coding, reported at least six of the nine secondary reports stated in Table 1 for enhancing construct validity; at least two of the three secondary reports stated for increasing internal validity; at least two of three secondary reports for enhancing external validity; and at least two of the three secondary reports stated for increasing reliability. The hypotheses were tested using a difference of means test (and we report the resulting t-test), or by using a limited dependent variable regression.

RESULTS

Descriptive statistics reveal some interesting tendencies. As Table 2 shows, only six percent of all articles (159 out of 2,643 papers) published in the 10 journals over the 1995–2000 period were explicitly based on case studies on for-profit organizations. The percentage of case study-based papers published varied considerably per journal, ranging from less than two percent to more than 15 percent of the total number of papers published in the period. The *Strategic Management Journal* published by far the most case studies in the top-tier bracket.

'case study' in the *Journal of Management* and three 'case studies' in *Management Science*, these did not report primary field work results in for-profit organizations, so these three journals were excluded from our final sample.

The sheer numbers of validity and reliability reports are also noteworthy. Table 3 shows that in the 159 case study papers, there were only two primary reports on construct validity (both counts were in 2000 and both were in the top-tier journals), and only five on internal validity (two counts in 1996, both in the three top-tier journals, and one count in 1998, 1999, and 2000, each one in a different journal of the second group). Table 3 exhibits the frequencies of the secondary reports on validity and reliability measures (we show composite measures, that is, studies reporting at least two-thirds of the measures for enhancing validity and reliability from Table 1). The picture is similar to the primary reports in that the three top-tier journals, despite only publishing 14 percent of all case studies, show more than twice as many internal and construct validity counts than do all other seven journals taken together.

A different picture, however, emerges when looking into external validity counts (whether primary or secondary reports). Here, authors in *California Management Review, Journal of International Business Studies, Journal of Management Studies, Long Range Planning, Organization Science, Organization Studies,* and *Sloan Management Review* seem to pay significantly more attention to external validity (and not much else) than do authors in *Academy of Management Journal, Administrative Science Quarterly,* or the *Strategic Management Journal.*

The proportion of construct and internal validity reports (composite) on external validity reports is particularly noteworthy. The methodology literature suggests that relatively more emphasis be put on construct and internal validity than on external validity. For example, the literature suggests four measures that can be used to address external validity and a total of 12 measures for construct and internal validity (see Table 1 for details). This points to a relationship of construct and internal validity measures to external validity measures of 3 : 1. The three top journals exhibit a relationship of 1 : 1, whereas the other seven journals report more than twice as many external validity measures than internal and construct validity taken together. Thus, while the practices in the three top journals are clearly not 'perfect,' they at least balance attention between validity types. For the other

Table 2. Percentage of case study-based papers published and coded (1995–2000)

Journal	1995	1996	1997	1998	1999	2000	1995–2000
Academy of Management Journal	8.3	0	0	0	0	1.3	1.9
Administrative Science Quarterly	0	0	0	0	0	8.3	1.3
California Management Review	3.6	11.1	10.7	5.1	0	7.4	6.1
Journal of International Business Studies	2.6	0	3.2	0	2.4	0	1.2
Journal of Management Studies	25.7	2.9	29.7	7.7	15.8	10.4	15.1
Long Range Planning	7.6	5.1	6.6	9	15.7	9.7	8.2
Organization Science	0	0	0	4.3	17	4.3	4.5
Organization Studies	5.9	4.9	8.3	9.5	10.5	6.7	7.6
Sloan Management Review	13.8	13.3	33.3	10.7	8.8	21.4	9.5
Strategic Management Journal	2.1	6.5	0	1.5	1.6	10.1	3.6
Total	6.9	3.9	6.1	5	7.5	7.1	6

Table 3. Frequencies per journal of primary and secondary reports

Journal	Case study papers 1995–2000	Primary reports				Secondary reports (composite)			
		Internal validity	Construct validity	External validity	Reliability	Internal validity	Construct Validity	External validity	Reliability
Academy of Management Journal	7	0	1	1	2	4	1	3	1
Administrative Science Quarterly	2	0	1	1	1	1	2	2	1
Strategic Management Journal	13	2	0	1	2	4	5	11	4
Subtotal top-tier journals	**22**	**2**	**2**	**3**	**5**	**9**	**8**	**16**	**6**
California Management Review	11	0	0	1	1	0	2	5	2
Journal of International Business Studies	3	1	0	2	1	0	0	3	1
Journal of Management Studies	35	0	0	11	3	6	4	24	7
Long Range Planning	35	1	0	2	0	1	2	5	2
Organization Science	12	1	0	4	4	3	3	8	3
Organization Studies	18	0	0	8	2	2	4	17	3
Sloan Management Review	23	0	0	0	0	0	0	4	3
Subtotal other seven journals	**137**	**3**	**0**	**28**	**11**	**12**	**15**	**66**	**21**
Total	*159*	*5*	*2*	*31*	*16*	*21*	*23*	*82*	*27*

Note: Composite measures refer to case studies that reported at least two-thirds of the research procedures for enhancing construct validity, internal validity, external validity, and reliability.

group of journals, there is a strong imbalance in favor of external validity.

Finally, there is the question of whether journals that publish significantly more case studies are more 'lax' in their requirements than others. Despite the small numbers, the data seem to suggest that journals in the top bracket that publish more case studies (such as the *Strategic Management Journal*) are clearly not more lax in their rigor requirements. Table 3 suggests that this cannot be said of case study friendly outlets in the other seven journals.

Our first hypothesis argued that the rigor of case study research is not a random feature, so case studies that provide secondary reports on reliability are more likely also to provide secondary reports on construct validity, internal validity, and external validity. This hypothesis is only partly confirmed by our data (Table 4). Case studies that provide secondary reports on reliability are more likely also to provide secondary reports on construct validity and external validity (at a 10% level). Table 4 also includes a composite measure called 'overall validity,' defined as case studies that reported at least 11 of the 16 research procedures stated in Table 1 for increasing validity. This new composite variable is only significant at a 10 percent level.

Table 5 shows that Hypotheses 2a (p < 0.1), 2b, 3a, and 3b are supported. Case studies published in *Academy of Management Journal, Administrative Science Quarterly*, and *Strategic Management Journal* are more likely to provide primary and secondary reports on both internal validity and

construct validity. Furthermore, Table 5 shows that while case studies in the highest-ranked journals are not more likely to provide primary reports on external validity, they are also not more likely to provide secondary reports on external validity.

Hypothesis 4a stated that case studies are more likely to provide primary reports on external validity than on construct validity, internal validity, and reliability. Table 6 shows that this hypothesis is strongly supported by our data. However, Hypothesis 4b (which stated that case studies are more likely to provide secondary reports on external validity than on construct validity, internal validity, or reliability) is supported only as far as case studies' tendency to provide secondary reports on reliability is concerned.

DISCUSSION

While the results reveal some irregularities in applying validity and reliability measures, they also suggest how the rigor problem might be successfully (if not perfectly) addressed. Case study papers published in the highest-ranked journals (*Academy of Management Journal, Administrative Science Quarterly,* and *Strategic Management Journal*) were more likely to provide primary reports on internal validity and construct validity, but not on external validity. Although the small numbers need to be interpreted with some caution (the highest-ranked journals together published a grand total of 22 case studies in the six-year time horizon, see Table 2), the 'best practices' in highest-ranked journals are clearly in line with the theory that a logical prerequisite for external validity is a case study's internal and construct validity (e.g., Cook and Campbell, 1979; Gibbert, 2006). That is, case study authors in highest-ranked journals demonstrate that they not only are aware of the four validity and reliability criteria, they also demonstrate that they are aware of the relationships among them (Campbell and Stanley, 1963; Cook and Campbell, 1979; McGrath, 1981). Put differently, the results suggest that a case study author may emphasize the more fundamental types of validity at the expense of external validity, without diminishing the case study's overall rigor, but not vice versa (see Table 3).

The results suggest important implications for authors, reviewers, and readers of case studies. Table 2 shows that in the journals and time period

Table 4. Regression of composite reliability on composite validity measures

Composite measure		
Construct validity	2.11**	
	(0.72)	
Internal validity	0.80	
	(0.55)	
External validity	0.89*	
	(0.51)	
Overall validity		1.13*
		(0.67)
LR Chi2	16.09	2.59
P > Chi2	0.00	0.11
Observations	159	159

Notes:
1. Standard errors in parentheses
2. * = significant at 10% level, ** = significant at 5% level

Table 5. Testing for differences between journals and validity/reliability concerns

Primary report	Test: (highest-ranked journals − other journals) = 0	Secondary report (composite variables)	Test: (highest-ranked journals − other journals) = 0
Internal validity	$t = 1.73$ (0.09)	Composite internal validity	$t = 3.96$ (0.00)
Construct validity	$t = 3.67$ (0.00)	Composite construct validity	$t = 2.00$ (0.05)
External validity	$t = 0.74$ (0.46)	Composite external validity	$t = 2.16$ (0.03)
Reliability	$t = 2.14$ (0.03)	Composite reliability	$t = 1.38$ (0.17)
Observations	159		159

Table 6. Testing for differences between external validity and other validity/reliability concerns

Primary report	Test: (external validity − variable) = 0	Secondary report	Test: (composite external validity − variable) = 0
Construct validity	$t = 5.71$ (0.00)	Composite construct validity	$t = 1.00$ (0.32)
Internal validity	$t = 5.39$ (0.00)	Composite internal validity	$t = -1.15$ (0.25)
Reliability	$t = 2.32$ (0.02)	Composite reliability	$t = 3.78$ (0.00)
Observations	90		90

investigated, the percentage of case studies published in the 10 journals has varied at around six percent of all papers published. Nevertheless, the results also suggest that there is indeed a market for rigorous case studies, and top-tier journals in our sample, such as the *Strategic Management Journal* do seem to be willing to publish high-quality case study work. As such, we hope that some of the empirical guidelines discussed in this study will make editorial policies more transparent and will encourage authors to prepare more, and more rigorous case studies for submission.

Eisenhardt and Graebner (2007) recently pointed out that case studies are often among the most interesting articles to read. As such, our study may be appreciated also by readers of case studies, who can now be more confident in a case study's findings if they know how to assess validity and reliability. Reviewers are encouraged not only to assess whether authors name 'the usual suspects' of case study rigor, but to probe deeper into individual emphases on validity and reliability measures in case studies, both in terms of primary and secondary reports (see Table 1).

Last but not least, we acknowledge that researchers have very different views on the case study method. Some researchers may see the case study method as an alternative to 'mainstream' or positivist research methods and may be critical of

an attempt to emulate the natural science model in data collection and analysis strategies (e.g., Daft and Lewin, 1990; Behling, 1980; Jick, 1979; Lincoln and Guba, 1985; Miles, 1979; Mir and Watson, 2000). However, as noted above, we failed to identify in our sample a single case study that used, and explicitly reported, rigor criteria other than the validity and reliability notions discussed here. We are therefore confident that the present review gives an accurate and comprehensive reflection of what exactly passes as a rigorous case study. At the same time, we hope that this article may trigger a discussion about alternative ways to conceptualize, ensure, and assess the rigor of case study research.

ACKNOWLEDGEMENTS

REFERENCES

Amabile TM, Patterson C, Mueller J, Wojcik T, Odomirok PW, Marsh M, Kramer SJ. 2001. Academic-practitioner collaboration in management

research: a case of cross-profession collaboration. *Academy of Management Journal* **44**(2): 418–431.

Behling O. 1980. The case for the natural science model for research in organizational behavior and organization theory. *Academy of Management Review* **5**(4): 483–490.

Bergh DG, Perry J, Hanke R. 2006. Some predictors of *SMJ* article impact. *Strategic Management Journal* **27**(1): 81–100.

Burgelman RA. 1983. A model of the interaction of strategic behavior, corporate context and the concept of strategy. *Academy of Management Review* **8**(1): 61–70.

Calder BJ, Phillips LW, Tybout AM. 1982. The concept of external validity. *Journal of Consumer Research* **9**(3): 240–244.

Campbell DT. 1963. From description to experimentation: interpreting trends as quasiexperiments. In *Problems in Measuring Change,* Harris CW (ed). University of Wisconsin Press: Madison, WI; 212–244.

Campbell DT. 1975. Degrees of freedom and the case study. *Comparative Political Studies* **8**: 178–193.

Campbell DT, Stanley J. 1963. *Experimental and Quasi Experimental Designs for Research.* Houghton Mifflin: Boston, MA.

Cook TD, Campbell DT. 1976. Design and conduct of quasi-experiments and true experiments in field settings. In *Handbook of Industrial and Organizational Psychology*, Dunnette MD (ed). Rand McNally: Chicago, IL; 223–326.

Cook TD, Campbell DT. 1979. *Quasi-Experimental Design: Design and Analysis Issues for Field Settings*. Rand McNally: Skokie, IL.

Daft RL, Lewin AY. 1990. Can organization studies begin to break out of the normal science straitjacket? An editorial essay. *Organization Science* **1**(1): 1–9.

Denzin NK, Lincoln YS. 1994. *Handbook of Qualitative Research*. Sage: Thousand Oaks, CA.

Eisenhardt KM. 1989. Building theories from case study research. *Academy of Management Review* **14**(4): 532–550.

Eisenhardt KM, Graebner ME. 2007. Theory building from cases: opportunities and challenges. *Academy of Management Journal* **50**(1): 25–32.

Gibbert M. 2006. Muenchausen, black swans, and the RBV. *Journal of Management Inquiry* **15**(2): 145–151.

Glaser B, Strauss A. 1967. *The Discovery of Grounded Theory*. Aldine Press: Chicago, IL.

Jick TD. Mixing qualitative and quantitative methods: triangulation in action, *Administrative Science Quarterly* **24**(4): 602–611.

Larsson R. 1993. Case survey methodology: quantitative analysis of patterns across case studies. *Academy of Management Journal* **36**(6): 1515–1546.

Leonard-Barton D. 1990. A dual methodology for case studies: synergistic use of a longitudinal single site with replicated multiple sites. *Organization Science* **1**(3): 248–266.

Lincoln YS, Guba EG. 1985. *Naturalistic Inquiry*. Sage: Newbury Park, CA.

March JG, Sproull LS, Tamuz M. 1991. Learning from samples of one or fewer. *Organization Science* **2**(1): 1–13.

McGrath JE. 1981. Dilemmatics: the study of research choices and dilemmas. *American Behavioral Scientist* **25**: 179–210.

McGrath JE, Brinberg D. 1983. External validity and the research process: a comment on the Calder/Lynch dialogue. *Journal of Consumer Research* **10**(1): 115–124.

Miles MB. 1979. Qualitative data as an attractive nuisance: the problem of analysis. *Administrative Science Quarterly* **24**(4): 590–601.

Mir R, Watson A. 2000. Strategic management and the philosophy of science: the case for a constructivist methodology. *Strategic Management Journal* **21**(9): 941–953.

Mook DC. 1983. In defense of external validity. *American Psychologist* **38**: 379–387.

Numagami T. 1998. The infeasibility of invariant laws in management studies: a reflective dialogue in defense of case studies. *Organization Science* **9**(1): 2–15.

Penrose ET. 1960. The growth of the firm. A case study: the Hercules powder company. *Business History Review* **34**(1): 1–23.

Pettigrew A. 1973. *The Politics of Organizational Decision Making*. Tavistock: London, UK.

Rouse MJ, Daellenbach US. 2002. More thinking on research methods for the resource-based perspective. *Strategic Management Journal* **23**(10): 963–967.

Scandura TA, Williams EA. 2000. Research methodology in management: current practices, trends, and implications for future research. *Academy of Management Journal* **43**(6): 1248–1264.

Siggelkow N. 2007. Persuasion with case studies. *Academy of Management Journal* **50**(1): 20–24.

Tahai A, Meyer MJ. 1999. A revealed preference study of management journals' direct influences. *Strategic Management Journal* **20**(3): 279–296.

Weick K. 2007. The generative properties of richness. *Academy of Management Journal* **50**(1): 14–19.

Yin RK. 1981. The case study crisis: some answers. *Administrative Science Quarterly* **26**(1): 58–65.

Yin RK. 1994. *Case Study Research: Design and Methods*. Sage: London, UK.

Yin RK. 1999. Enhancing the quality of case studies in health services research. *Health Services Research* **34**(5): 1209–1224.