Transparent Practices: Primary and Secondary Data in Business Ethics Dissertations

Shawn W. Nicholson
Terrence B. Bennett

ABSTRACT. We explore the availability and use of data (primary and secondary) in the field of business ethics research. Specifically, we examine an international sample of doctoral dissertations since 1998, categorizing research topics, data collection, and availability of data. Findings suggest that use of only primary data pervades the discipline, despite strong methodological reasons to augment business ethics research with secondary data.

KEY WORDS: business ethics, datasets, doctoral research, open data, primary data, secondary data

Introduction and issues

Popular and scholarly treatment of ethical dealings in business continues to garner considerable attention, with recent books, dissertations, journal articles, and news reports supporting both applied and theoretical advances in this field. In the scholarly arena there has been much attention paid to the methodological choices used to study business ethics (Bryman and Bell, 2003; Crane, 1999; Godfrey and Hatch, 2007; Robertson, 1993). In particular, the past decade has witnessed articles promoting greater use of archival and other secondary data to, among other things, reduce the cost (monetary and political) of collecting primary data, and enhance (or refute) the conclusions derived from primary data, the integrity of which can be compromised by the inevitable interrelationship of the data sources and the ethical practices being observed. In his 1998 article "The Use of Secondary Data in Business Ethics Research," Cowton uses these pages to "stimulate the interest of business ethics researchers in using secondary data" (Cowton, 1998). He worries about the "poor empiricism" in business ethics research and asserts that "as a general rule it seems to be the case that researchers are not as aware as they might be of the potential of secondary data for providing valuable insights into a whole range of questions in a cost effective manner" (Cowton, 1998).

Further advancing Cowton's observations, Harris explains that ethics research is strengthened by analysis of impartial secondary sources. Because the subject matter itself is sensitive, embarrassing, threatening, stigmatizing, or incriminating (Harris, 2001), it can be difficult to collect reliable primary data. His study uses a content analysis of major newspapers to compile secondary data that supports a richer exploration of business ethics. Furthermore, he positions secondary data analysis as a means "to constrain the researcher faced with the temptation to arrive at more extensive conclusions than the [primary] data would support," since classification of the secondary data was clearly specified and able to withstand careful scrutiny (p. 201).

Concomitant with these calls for greater use of secondary data is the evolving discourse surrounding open access and open data. Open data is a philosophy and practice requiring that certain data are freely available to everyone, without restrictions from copyright, patents or other mechanisms of control. The advocates of this model believe that there are substantial benefits that arise from the sharing of research data.[1]

The growing awareness of the advantages of open data is, of course, not exclusive to business research, but reaches across many scholarly disciplines. For example, Wicherts et al. document the slim results from their efforts to collect datasets for reanalysis from articles published in a set of top psychology journals (Wicherts et al., 2006). Responding to

Wicherts et al., a recent editorial in *Nature* laments the poor tradition in psychology regarding data availability and lays out a plan whereby the American Psychological Association shall require the "deposition of data as supplementary electronic material in APA databases" ("A Fair Share," 2006).

In this study we primarily seek to understand the nature of data use in business ethics dissertation research. Secondarily, we consider the overall ease of findability of dissertations relevant to a specific field of inquiry, as well as the ease of ascertaining the presence or use of datasets within these dissertations. Finally, we look for any indication of adherence to trends in open data within the discipline of business ethics. To this end, we analyze an international sample of recent dissertations in the field of business ethics, hoping to identify trends and patterns in the availability and use of secondary data; track the creation and open availability of primary data; and examine the subsequent re-use of primary data.

Our focus on dissertations not only allows us to derive a manageable-sized sample, but also directs our attention to nascent scholarship. Chang and Hsieh assert that "the doctoral dissertation is the major distinguishing feature of education, which traditionally has a dual role: (1) [to] make a positive, original and significant contribution to knowledge and (2) to provide training in research and scholarly techniques" (Chang and Hsieh, 1997). It is reasonable to assume that dissertation authors, as emerging scholars, are exemplars in creating rigorous studies that give effect to the latest trends, theories, and best practices in research and scholarship in their discipline.

Methodology

An international sample ($N = 48$) of doctoral dissertations primarily from North America and the United Kingdom from 1998 to 2007 forms the population for this analysis. This sample was derived from a simple keyword search for ("business AND ethics") in two online subscription databases: ProQuest Dissertations & Theses (PQDT, most of the contents of which comes from the United States and Canada) and the Index to Theses (Great Britain and Ireland). When available, we also examined dissertation records from another ProQuest database:

Dissertations & Theses @ CIC Institutions.[2] Records in this database are likely to include more extensive indexing than that found in PQDT. Titles from the results of this initial search were examined in order to exclude results that were clearly unrelated to business ethics. From the remaining sample, non-doctoral-level dissertations were omitted (PQDT and Index to Theses also include masters-level theses), as were results from disciplines unrelated to business or ethics. Ten (21%) of the dissertations came from the UK, while the remaining 79% were from the USA or Canada (except for one English-language dissertation from The Netherlands).

Each dissertation has been content-analyzed and coded for 11 variables (see Appendix A). We reviewed several components of each dissertation for an indication of the use of either primary or secondary data, utilizing a mixture of manifest and latent coding (Monette et al., 2005). Specifically, we paid close attention to mentions in each dissertation's abstract of data collection and use, acknowledging the assertion by Adams and White that authors "typically want to represent their research as accurately as possible in an abstract" (Adams and White, 1994). We also looked at tables of contents, appendices, and the text itself in order to determine the creation (and availability for re-use) of primary data, as well as the presence of primary and secondary data. For purposes of this analysis, we again take our cue from Cowton and define "data" broadly to include not only numeric datasets, but also resources such as survey results, ethnographic studies, or coded text (Cowton, 1998).

Analysis and discussion

We first examined information *about* the dissertations in our sample that could be gleaned from secondary sources, specifically, the citations and abstracts (when available) and other indexing information found in PQDT and Index to Theses. This approach models the typical researcher's level of access to a large universe of dissertations: a browsable collection of full-text dissertations is uncommon; however, an online index or other finding aid is likely to be the most readily available tool for identifying works of potential interest.

The role of title words and subject terms in initial screening

While our primary focus was the presence – and availability for use – of data in business ethics dissertations, we first had to consider the overall ease of findability of dissertations in this broadly multidisciplinary field. How fruitful would a search of dissertation titles (presumably the first and fastest means by which a dissertation's subject reveals itself) be in identifying to a researcher potentially suitable dissertations in the field of business ethics?

The results from our sample were mixed (Figure 1). Not quite half (46%) of the dissertations in our sample had titles that contained the phrase "business ethics," although an additional 10% had titles that contained both words ("business" or a very closely related term, and "ethics") in the title. Another 40% had only the word "ethics" in the title (and were only deemed also to be business–related because of other descriptive features), while 4% had titles that did not mention ethics at all. If a researcher were trying to identify relevant dissertations in business ethics by searching only for results that included both of those words in the title, then nearly half of the items in our sample (44%) would have been excluded.

Titles, of course, do not necessarily serve to fully describe content, and it is reasonable to assume that a careful researcher would also use other indexing cues when screening for relevant content. We expected that when subject or keyword terms were provided in a commercial dissertation database, their presence would increase the findability of relevant results. This proved to be mostly untrue. Seventeen percent of the dissertations in our sample did not include subject or keyword terms as part of their database record. Had these dissertations been given titles without "business" or "ethics" – or did not have appropriately descriptive abstracts – then they would have been incorrectly omitted from our universe of relevant dissertations.
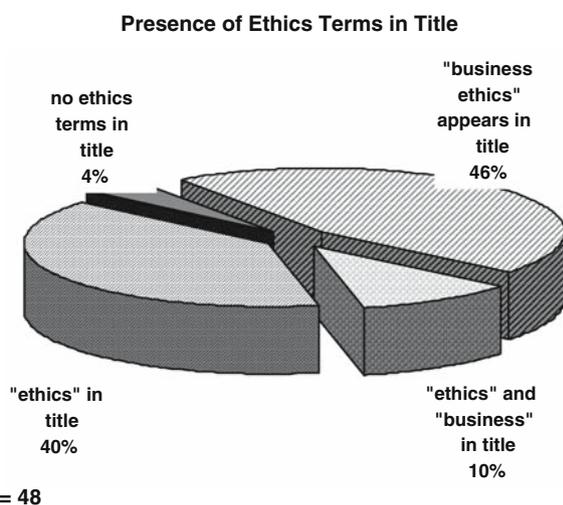
Unfortunately, even when they are available, subject terms supplied by the database vendor can be inadequate for identification of relevant dissertations. In PQDT, dissertations in business ethics fall under broad subject terms such as "business administration – management" or "business education" or even "philosophy." The two dissertations in our sample for which "business ethics" was assigned as a subject term were found in Dissertations & Theses @ CIC Institutions – an indexing tool with more precise subject identifiers, but one that is available to only a small percentage of researchers.

Presence of data revealed in abstracts and tables of contents

Our next point of inquiry was to determine the adequacy of author-supplied abstracts in revealing the presence of data within a dissertation. First, it is worth noting that it became evident to us that in some instances the abstract included in PQDT or the Index to Theses was a foreshortened version of the abstract found in the dissertation itself. When portions of an abstract have been omitted in a database, this has obvious negative implications for the findability of relevant dissertations by researchers. Based on our examination of abstracts, it was disappointing to discover that only 27% of the abstracts in our sample explicitly indicate that data is included in the research (see Figure 2), while a full 58% make no mention of data whatsoever (the remaining 15% of the abstracts offer an inconclusive indication that data *may* be present).

A subsequent examination of the tables of contents of our sample set of dissertations reveals that a much larger portion of our sample does contain data. The tables of contents explicitly cite data in 58% of

Presence of Ethics Terms in Title

no ethics terms in title 4%

"business ethics" appears in title 46%

"ethics" in title 40%

"ethics" and "business" in title 10%

N = 48

Figure 1. Appearance of "business ethics," "ethics" or "business," and "ethics" in titles of a sample of dissertations dealing with business ethics.

Abstract Indicates Presence of Data
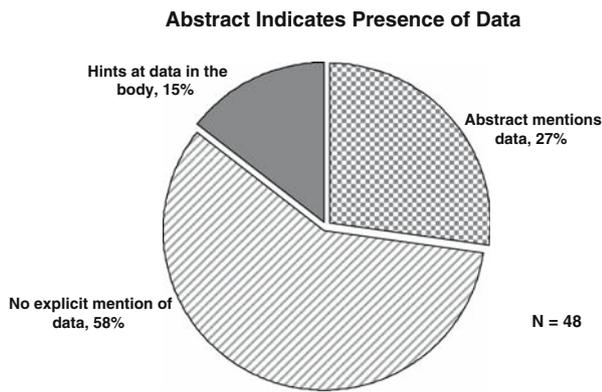
Figure 2. Abstract in dissertation database cites the presence of data in the dissertation.

our sample, and another 19% hint at the presence of data. Only 23% of our sample included tables of contents that offered no explicit mention of data. Considered another way, we found that among all of the dissertations with *abstracts* that did not mention the presence of data (73% of our sample, $N = 35$), an examination of their *tables of contents* revealed that 43% of these ($N = 15$) actually do offer data-supported research.

These results refute the claim by Adams and White (1994) that abstracts are an accurate reflection of an author's work. The implication of this finding is serious: while abstracts are generally included in dissertation-finding tools (i.e. online indexes to dissertations, such as PQDT), tables of contents are not included. Therefore, the use of only an online index to identify dissertations in business ethics that offer data-supported research – and that may include datasets – would fail to serve up a significant proportion of relevant dissertations. These could only be discovered by a researcher who is willing to complete the laborious process of obtaining the full text of a dissertation in order to examine the table of contents. And, as we discovered, even a detailed table of contents may not explicitly reveal the presence of data within a dissertation, as discussed below.

Collection of primary data and secondary data

While 58% of the abstracts in our sample made no mention of data, and 23% of the tables of contents did not reveal the presence of data, an examination of the full text of the dissertations in our sample revealed that only 16% truly did not include any primary data, and a mere 5% were without either primary or secondary data (See Table I). A further

TABLE I

Collection of primary and secondary data

Description	Number	Percent (%) of total
Dissertations with primary data collected		
Total with no primary data	7	16
Total with only one type of primary data	26	61
Total with two types of primary data	10	23
	43	100
Dissertations with secondary data collected		
Total with no secondary data	24	56
Total with only one type of secondary data	11	25
Total with two types of secondary data	2	5
Total with three or more types of secondary data	6	14
	43	100
Presence of primary and secondary data		
No primary or secondary data	2	5
Primary data only	22	51
Secondary data only	5	12
Primary and secondary data	14	32
	43	100

examination of the type of primary and/or secondary data reveals some interesting trends.

Among the 84% of dissertations in our sample that included primary data, an overwhelming majority (61% of the total sample, see Table I) reflected collection of just one type of data – generally results from a survey or questionnaire (see Table II). Most of the remaining dissertations with primary data reported findings from ethnographic or other observations, or from in-depth interviews. Only two dissertations (4% of the total sample) reported results from actual tests or experiments. From this, it follows logically that half of dissertations that offered two types of primary data (23% of the total sample, see Table I) included at least one dataset arising from a structured or in-depth interview, or from the results of a survey or questionnaire (see Table II).

The field of business ethics is informed by multiple disparate disciplines and inevitably involves examinations of human behavior and psychological motivation; so, it is not surprising that the dominant forms of primary data discovered in these dissertations are not numeric datasets but text-based data such as ethnographic observations, survey results, and transcribed interviews. But if researchers in business ethics are heeding the advice of Cowton (1998) and Harris (2001), then we would expect to see an abundant use of secondary data to reinforce the conclusions derived from analyses of primary data. However, the dissertations in our sample do not match up to this expectation.

Among the dissertations in our sample, fully 56% are without any secondary data (see Table I). Of the 44% that do make use of some form of secondary data, we found only one dissertation (3% of all instances of secondary data within the sample, see Table II) that uses numeric datasets, and nine (24% of all instances of secondary data) that are supported by information from corporate financial reports or other SEC filings (in some cases, each of these was present in a single dissertation). Only five dissertations (13%) made use of publicly available corporate reports such as mission statements or ethics statements, while four dissertations (11%) cited government documents other than SEC filings. Surprisingly, only two dissertations (5%) reported the use of internal corporate reports or correspondence.

Because our inquiry was partly inspired by Cowton's well-reasoned argument for the use of secondary data in business ethics research, we also examined the bibliographies of all of the dissertations in our sample to determine if his article ''The Use of Secondary Data in Business Ethics Research'' (Cowton, 1998) was included. We were disheartened to discover that only one of our authors cited this article by Cowton.

Availability of collected data

In support of our interest in the growing conversation surrounding open data, we also examined the extent to which the authors in our sample made their data available for further use. We focused not only on the authors' primary data, but also on secondary data to which the author may have added value through time-consuming collection and re-presentation.

By invoking a generous interpretation of the concept, we were able to ascertain that at least some data were made available within the text or appendices of 26% of the dissertations in our sample, although it is arguable whether data such as selected excerpts from structured interviews or coding from the content analysis of interview transcripts would be particularly useful to future researchers. In just five instances (11%), dissertations include a clear citation to secondary data, generally government reports from which the author has collected and assembled data.

Not surprisingly, for a large majority (63%) of the dissertations in our sample, direct access to the data was unavailable to readers, and in most of these cases (47%) it remained unclear whether the author would (or could) provide data to those who might request it. Given the sensitive nature of much of the data that must be collected for research in this field, it is understandable that authors need to be cautious about making it available, and it was not unexpected to find qualifying statements such as ''confidentiality was guaranteed to interviewees'' or ''transcripts of confidential interviews will be destroyed within three years.''

Conclusions

Analysis of primary and secondary data is an integral component of the research and scholarship taking

TABLE II

Types of primary and secondary data

Code	Description	Number[a]	Percent (%) of total
Types of primary data collected			
8.0	No primary data collected	7	13
8.1	Observation: ethnographic study	5	9
8.2	Observation: other	2	4
8.3	Survey or questionnaire	24	45
8.5	Results of tests or experiments	2	4
8.6	Other (10 are structured interviews)	13	25
		53	100

Code[b]	Description	All dissertations		Dissertations with secondary data	
		Number[c]	Percent (%) of total	Number	Percent (%) of total
Types of secondary data collected					
9.0	No secondary data collected	24	39	0	0
9.1	Numeric datasets: from ICPSR	0	0	0	0
9.11	Numeric datasets: other	1	2	1	3
9.2	Corporate reports (annual reports; etc.)	6	10	6	16
9.21	Other public company documents	5	8	5	13
9.3	Internal corporate reports	2	3	2	5
9.4	Government: SEC filings	3	5	3	8
9.41	Government: regulatory data	1	2	1	3
9.42	Government: other	3	5	3	8
9.5	Legal/court cases	0	0	0	0
9.6	Academic reports	3	5	3	8
9.7	Synthesis of literature	6	10	6	16
9.8	Other	8	13	8	21
		62	100	38	100

[a]N = 53 (43 dissertations; 10 with two types of primary data).
[b]See Appendix for fuller description of codes.
[c]N = 62 (43 dissertations: 24 with no secondary data; 2 with 2 types; 6 with 3 or more types).

place at research institutions, yet methodological choices continue to be challenged within the business ethics literature. Specifically, recent articles implore greater use of secondary data to augment the collection and analysis of primary data. This study sought to investigate the nature of data collection and use in recent doctoral dissertations in business ethics.

While most of the dissertations examined for this study reported on the analysis of primary data gathered by the authors, we were somewhat surprised to find very few dissertations that made use of secondary data to bolster the exploration of business ethics concerns.

These observations about the use and availability of data within business ethics research are made worse by the additional layer of opacity resulting from finding aids – specifically online indexes to dissertations – that do not enhance the findability of relevant research in this field. Finally, we note that within our sample there is scant evidence of adherence to the tenets of the open data movement.

Limitations and future research

This analysis explores issues related to the use and availability of primary and secondary data in doctoral dissertations. By focusing on the multidisciplinary field of business ethics, we sought to make discoveries that might reflect practices of data use in other disciplines. However, comparative analyses from other disciplines – especially from fields in which research data is more commonly characterized by numeric datasets rather than coded text – is essential for a deeper understanding of the applicability of our findings to research in other disciplines.

Also, while we note the relatively low use of secondary data in business ethics research – despite a clear call for its use by established scholars in this field – we do not attempt to discern what drives the continued exclusive reliance among a majority of new scholars on primary data collection. Is there a culture that dictates that original research can only be demonstrated by original data collection? Further studies must be undertaken to explore this issue.

We have also observed researchers do not expressly promote the availability of their primary data for re-use and further analysis. Again, we do not

attempt to explore the reasons behind this apparent unwillingness to make original research sufficiently transparent to draw other scholars with similar interests into the conversation. Is there a research culture that supports data hoarding? Is there fear among researchers of encroachment by others into their field of original inquiry? Not only must these questions by probed further, but it will also be interesting to observe whether the nascent stirrings among proponents of the open data movement will evolve into a sea change among researchers and scholarly publishers concerning the findability and availability for use of research data.

Notes

[1] "Open data" is an evolving concept whose advocates are affiliated with a variety of disciplines. Their opinions are often most ably captured in blogs (http://wwmm.ch.cam.ac.uk/blogs/murrayrust/?p=913), list servs (http://www.arl.org/sparc/opendata/), and even the online contributed encyclopedia Wikipedia (http://en.wikipedia.org/wiki/Open_data).

[2] This is a database of full-text dissertations from the schools that comprise the Committee on Institutional Cooperation (CIC): http://www.cic.uiuc.edu/.

Appendix A

NB: The raw data for this paper includes dissertations from North America and the United Kingdom. The authors will gladly share the list with interested parties, yet the text itself remains the property of authors and/or universities and/or ProQuest/UMI.

Coding scheme

Code	Label
1	Year
2	Country
Series 3	Word in title
3.0	No ethics terms in title
3.1	"Business ethics" appears in title

continued

Code	Label
3.2	"Ethics" and "business" in title
3.3	"Ethics" in title
Series 4	Subjects/Keyword terms
4.0	No Subjects/Keyword terms
4.1	Ethics (business)
4.11	Ethics (philosophy)
4.2	Business Administration
4.3	Management or Managers
4.4	Other
Series 5	Abstract browse
5.0	No abstract available
5.1	Abstract mentions data
5.2	No explicit mention of data
5.3	Hints at data in the body
Series 6	Table of contents browse
6.0	No table of contents
6.1	Table of contents cites data explicitly
6.2	No explicit mention of data
6.3	Hints at data, but unclear without further probing
Series 7	Research design
7.1	Exploratory
7.2	Descriptive
7.3	Causal
7.4	Other
Series 8	Primary data
8.0	No primary data collected
8.1	Observation: ethnographic study
8.2	Observation: other
8.3	Survey or questionnaire
8.4	Findings from focus groups/meetings
8.5	Results of tests or experiments
8.6	Other
Series 9	Secondary data
9.0	No secondary data collected
9.1	Numeric datasets: from Inter-University Consortium for Political and Social Research (ICPSR)
9.11	Numeric datasets: other
9.2	Corporate reports (annual reports; mission statements; ethics statements)
9.21	Other public company documents (mission or ethics statements; press releases)
9.3	Internal (private) corporate reports or correspondence
9.4	Government: SEC filings

continued

Code	Label
9.4	Government: regulatory data (EPA, FEC, etc.)
9.4	Government: other
9.5	Legal/court cases
9.6	Academic reports (journal and books)
9.7	Synthesis of literature
9.8	Other
Series 10	Data available
10.0	Data not available
10.1	Yes, in document (appendix, in text, etc.)
10.2	Yes, not in document, but available via ICPSR
10.3	Yes, not in document, but document contains a good citation
10.4	Available on Web
10.5	Unclear: follow-up with author is required
10.6	Data has been destroyed
Series 11	Cites Cowton (1998)
11.1	Yes
11.2	No

References

Adams, G. B. and J. D. White: 1994, 'Dissertation Research in Public-Administration and Cognate Fields – An Assessment of Methods and Quality', *Public Administration Review* **54**(6), 565–576.

Bryman, A. and E. Bell: 2003, *Business Research Methods* (Oxford University Press, Oxford, New York).

Chang, P. L. and P. N. Hsieh: 1997, 'A Qualitative Review of Doctoral Dissertations on Management in Taiwan', *Higher Education* **33**(2), 115–136.

Cowton, C. J.: 1998, 'The Use of Secondary Data in Business Ethics Research', *Journal of Business Ethics* **17**(4), 423–343.

Crane, A.: 1999, 'Are You Ethical? Please Tick Yes □ Or No □ on Researching Ethics in Business Organizations', *Journal of Business Ethics* **20**(3), 237–248.

A Fair Share: 2006, *Nature* **444**(7120), 653–654.

Godfrey, P. C. and N. W. Hatch: 2007, 'Researching Corporate Social Responsibility: An Agenda for the 21st Century', *Journal of Business Ethics* **70**(1), 87–98.

Harris, H.: 2001, 'Content Analysis of Secondary Data: A Study of Courage in Managerial Decision Making', *Journal of Business Ethics* **34**(3/4), 191–207.

Monette, D. R., T. J. Sullivan and C. R. DeJong: 2005, *Applied Social Research : A Tool for the Human Services*, 6th Edition (Thomson, Belmont, CA).

Robertson, D. C.: 1993, 'Empiricism in Business Ethics: Suggested Research Directions', *Journal of Business Ethics* **12**(8), 585–599.

Wicherts, J. M., D. Borsboom, J. Kats and D. Molenaar: 2006, 'The Poor Availability of Psychological Research Data for Reanalysis', *American-Psychologist* **61**(7), 726–728.

Shawn W. Nicholson
East Lansing, MI 48824-1048, U.S.A.
E-mail: nicho147@msu.edu

Terrence B. Bennett
College of New Jersey,
Ewing, NJ 08628, U.S.A.