Using content analysis as a research methodology for investigating intellectual capital disclosure

A critique

John Dumay
Department of Accounting & Corporate Governance, Macquarie University, Sydney, Australia, and
Linlin Cai
Department of Industrial & Systems Engineering, Hong Kong Polytechnic University, Hung Hom, Hong Kong

Abstract

Purpose – The purpose of this paper is to build on Dumay and Cai's (2014) prior research to provide a deeper analysis of the problems associated with using content analysis (CA) as a research methodology for investigating intellectual capital disclosure (ICD).

Design/methodology/approach – Totally, 110 articles utilising CA as a research methodology for inquiring into ICD are analysed based on Krippendorff's (2013) conceptual CA research framework and design logic, and tied into issues relating to CA as a research methodology for investigating ICD.

Findings – The authors advocate that ICD CA researchers need to go back to the drawing board and ensure that future studies rigorously apply the basic logic of CA design. In its current state, ICD CA research needs to take a few steps back, before it can move forward. If ICD CA researchers can accomplish this, then there is an opportunity to undertake rigorous research to develop reliable and valid outputs that add to new knowledge about IC.

Research limitations/implications – The main limitations of the research are the chosen sample of CA-based ICD articles and the adoption of the Krippendorff's framework. However, the authors have identified the main corpus of CA-based ICD studies and since Krippendorff is the only recognised comprehensive text on CA as a methodology, the authors use the most appropriate data and framework possible for the analysis.

Originality/value – Prior CA studies have laid the foundation for what is a popular research methodology. However, the authors argue that the popularity of CA as a research method for investigating ICD has become so great that at times the research methodology "drives the research questions" as opposed to the "research questions driving the methodology" Hence, this research examines reasons for CA limited contemporary contribution and recommends how this may be overcome rather than prescribing how to conduct ICD CA research.

Keywords Content analysis, IC research, Critical analysis, Research design, Intellectual capital disclosure

Paper type Literature review

> Content analysis is a research technique for making replicable and valid inferences from texts (or other meaningful matter) to the contexts of their use (Krippendorff, 2013, p. 34).

The authors especially want to thank Professor Eric Tsui and the Department of Industrial & Systems Engineering of Hong Kong Polytechnic University their support provided to the authors throughout the conduct of the research. The authors are also grateful to Jenny Hatton for her editorial feedback on early drafts, and Fiona Crawford of Macquarie University for her sterling efforts to refine the paper's presentation.

Journal of Intellectual Capital
Vol. 16 No. 1, 2015
pp. 121-155

1469-1930
DOI 10.1108/JIC-04-2014-0043

1. Introduction

This paper builds on Dumay and Cai (2014) to provide a deeper analysis of the problems associated with using content analysis (CA) as a research methodology for investigating intellectual capital disclosure (ICD). While this paper has implications from an ICD research perspective, it is also a case study of CA to highlight methodological issues and problems facing accounting and management researchers. In their review and critique of CA as a research methodology for investigating ICD Dumay and Cai (2014) found several problems highlighted by researchers, including the disclosure index used, the subjectivity of the CA research methodology, the unit of analysis and the weighting/quality of disclosures. However, because of length and scope limitations, Dumay and Cai (2014) did not expand on these problems and this paper now takes the opportunity to investigate and critique these and other related problems.

While Dumay and Cai (2014) are not overly optimistic about the future of CA as a research methodology, they argue that researchers need to "transform their understanding and application of CA". Additionally, Dumay and Cai (2014) found that contemporary CA research into ICD adds little new knowledge developing instead conflicting results (see Dumay and Cai, 2014, Table VI). They also claim that they are "not opposed to CA as a research methodology" and argue "CA should be utilised in an innovative way" so that CA contributes to new intellectual capital (IC) knowledge. Therefore, by examining and critiquing how researchers apply CA as a methodology in ICD research this paper aims to outline why contemporary ICD research might have conflicting results and limited contributions to IC knowledge.

This is not the first investigation into how CA is applied as a research methodology and its associated problems; several articles have appeared in different journals over the last decade (see Guthrie *et al.*, 2004b; Abeysekera, 2006; Beattie and Thomson, 2007; Steenkamp and Northcott, 2007; Husin *et al.*, 2012). However, with the exception of Abeysekera (2006), who offers some critique on the definition of IC, coding frameworks, source documents, research methods and theoretical interpretations, none of these articles offer any extensive critique of how the methodology is applied.

These studies have laid the foundation for what is a popular research methodology. However, the popularity of CA as a research method for investigating ICD has become so great that at times the research methodology "drives the research questions" as opposed to the "research questions driving the methodology" (Dumay and Cai, 2014). Hence, the popularity of CA and how it is applied may be the reason for its limited contemporary contribution as evidenced by the declining impact of CA as a research methodology (Dumay, 2014, p. 14; Dumay and Cai, 2014). This paper's insights and critique guide researchers as to how they might apply CA by answering three related research questions:

(1) How have different attributes of CA as a methodology been used for inquiring into ICD?

(2) How might CA be used differently for inquiring into ICD?

(3) What is the future for CA as a research methodology for investigating ICD?

This research extends the data set from Dumay and Cai's (2014) review and critique of 110 articles which use CA as a methodology for inquiring into ICD and is based on Krippendorff's (2013, pp. 35-45) conceptual CA research framework. This framework is

arguably a seminal resource for researchers wanting to utilise CA as a methodology because researchers from a wide range of disciplines cite different versions of Krippendorff's book *Content Analysis: An Introduction to Its Methodology* more than 13,000 times[1]. To our knowledge it is the most cited guide to CA as a research methodology.

A search of Google Scholar points to two other highly cited CA works being Weber (1990) and Neuendorf (2002) which have more than 6,000 and 4,000 citations, respectively. However, while these works add to researchers' knowledge of how to apply CA, both are research guidebooks and they do not offer methodological insights, rather they offer insights into how to apply CA methods. As Guthrie *et al.* (2004a, p. 417) outline "Methods are the means whereby one collects and analyses data. Methodology refers to the philosophical issues which underlie those methods" (see also de Villiers and Dumay, 2013, pp. 893-894). Additionally, these works draw heavily on the works of other researchers, especially Krippendorff, and do not offer an alternate methodological framework. For example, Weber (1990, pp. 17-18) draws exclusively on Krippendorff's first edition (1980) to outline CA reliability rather than develop a separate epistemological view of reliability. Hence, both books contribute to CA methods not methodology which is the focus of our research. Additionally, both Weber (1990) and Neuendorf (2002) are out of print and require updating to keep up with advances in state of the art CA research as Krippendorff's (2013) third edition does.

This paper does not suggest that Krippendorff's methodological framework is flawless and critiques aspects of it. This is in keeping with Krippendorff's (2013, p. 36) espoused use of the conceptual framework, which is to engender "long-term systematic improvements". Similarly, the logic behind CA research design is also critiqued (Krippendorff, 2013, pp. 81-97). Hence, to present this paper's framework, it first outlines Krippendorff's (2013) conceptual CA research framework and design logic, and ties this into issues relating to CA as a research methodology for investigating ICD. Second, it answers research questions one and two by presenting and discussing the research results. Third, it presents a conclusion by answering research question three and outlining the future for CA as a research methodology for investigating ICD.

2. The research framework

This section presents this paper's research framework (see Table I) based on Krippendorff's (2013, pp. 35-45) CA research framework comprising six conceptual components:

(1) a body of text, the data that a content analyst has available to begin the analytical effort;

(2) a research question that the analyst seeks to answer by examining the body of text;

(3) a context of the analyst's choice within which to make sense of the body of text;

(4) an analytical construct that operationalises what the analyst knows about the context;

(5) inferences that are intended to answer the research question, which constitute the basic accomplishment of the CA; and

(6) validating evidence, which is the ultimate justification of the CA.

The framework uses these conceptual components to outline issues relating to the logic and design of CA studies. This is complemented by drawing upon related issues raised in articles written about the use of CA as a research methodology for investigating IC, being Guthrie *et al.* (2004b), Abeysekera (2006),

Research questions	Research question Hypotheses No hypotheses or research questions		
Texts and context	Theory	Stakeholder theory	
		Legitimacy theory	
		Agency theory	
		Signalling theory	
		Resource-based view	
		Political economy of accounting	
		Other theory …	
		No theory	
	ICD index	Source? Is modified?	
	Content unit of analysis	Terms	
		Sentences	
		Words	
		Phrases	
		Theme	
		Events	
		Not disclosed	
	The quality and weighting of ICD	None	
		Yes	Method used
Reliability of CA	None disclosed		
	Agreement coefficient?		
	Other method?		
Inferences	Inductive		
	Deductive		
	Abductive		
Validity of CA	None		
	Corroborating evidence	Interviews	
		Questionnaire	
		Other …	
	Ex post facto research	Study confirmed	

Table I.
The research framework

Beattie and Thomson (2007), Steenkamp and Northcott (2007) and Husin *et al.* (2012), and/or accounting research such as social and environmental disclosures (e.g. Guthrie and Abeysekera, 2006). As Dumay and Cai (2014) outline, these "articles were chosen because they specifically address the use of CA as a research methodology for investigating ICD and have been extensively cited in articles that apply CA as a research methodology". The authors developed the framework iteratively by initially attending to the issues raised in the articles and then by analysing the articles and comparing what they found to Krippendorff's conceptual components. During this process, they discovered other issues, for example, the difference between the reliability and validity of ICD CA research, which they add to the initial analysis.

2.1 A body of text and a research question

According to Krippendorff (2013, p. 36) "most content analyses start with data that are not intended to be analyzed to answer specific research questions" because "they are meant to be *read, interpreted and understood by people other than the analysts*" (author's emphasis)[2]. This makes CA texts different from data normally used in qualitative research, because researchers have no control over producing the data as they would if they were, for example, conducting interviews. Hence, the availability of texts is

usually the starting point for CA research. Unsurprisingly, Guthrie *et al.* (2004b, p. 287) (b argues that annual reports are ideal texts to understand ICD as they are a "major medium for communicating with the public", "highly useful sources of information" and "produced on a regular basis by all companies". Dumay and Cai's (2014) prior research confirms this as they found that 79 per cent of the ICD CA research papers they analysed used annual reports as the primary data source.

However, as Abeysekera (2006, p. 66) argues, just because annual reports are a readily available medium, the information within cannot be guaranteed reliable and "may not reflect the objective reality of the firm". Thus, content analysts must recognise that the texts they analyse, such as annual reports, may be produced and "intended for someone like them" (Krippendorff, 2013, p. 36). As Evans (2013, p. 651) aptly observes "if words chosen to convey a corporation's past, its present situation, and its outlook are subject to manipulation, one should be careful about its interpretation". Considering the popularity of annual reports in many different forms of academic research it would be naïve for content analysts to think that annual reports are not in some way contaminated by their authors. For example, the Danish company Novo Nordisk is famous as a leading light in non-financial reporting and an early proponent of Integrated Reporting (see Dey and Burns, 2010); researchers should be fully aware of this is using Novo Nordisk's annual reports. Similarly, Unerman (2000, p. 667) outlines how analysing annual reports alone "is likely to result in an incomplete picture of reporting practices".

Considering the critique of annual reports as a data source it is advocated, as Guthrie and Abeysekera (2006, p. 122) propose, that other external and internal company disclosures, as well as the data collected through case/field/interview studies, surveys and experiments, offer robust sources of empirical data. According to Dumay and Cai (2014), some novel sources of text used in ICD CA studies are "price sensitive announcements to the Australian Stock Exchange (Dumay and Tull, 2007), competency standards of an accounting firm (Chang and Birkett, 2004) and transcripts of focus group meetings (Chen, 2009)".

While researchers use several sources of texts for CA research, the main related issue is the development of research questions for CA research. However, Abhayawansa (as quoted in Dumay and Cai, 2014) criticizes the development of ICD CA research:

> […] we should be asking more interesting questions in our IC research hoping that CA would be a suitable method to answer some of them. I guess, what I am trying to say is; should the method drive the research questions or the research questions drive the method?

Abhayawansa's observation is crucial because in CA the data comes before the research question (Krippendorff, 2013) whereas in empirical research it is normal to develop a research question by thoroughly analysing the contemporary academic literature to identify potential research gaps prior to embarking on research (de Villiers and Dumay, 2013, p. 893). However, despite the data coming first, Krippendorff (2013, p. 37) still advocates that research questions should come first before the reading and analysis of texts for the sake of "efficiency and empirical grounding". Thus, the research question in CA is important because the context of "research questions affects the perceived relevance of texts and their meanings" (Steenkamp and Northcott, 2007, p. 14). However, Abhayawansa's observation and the fact that many researchers still use annual reports as data sources suggests to us that CA research is somewhat isomorphic and a lack of ingenuity and innovation in developing research questions

may play a role in the continuing decline in the academic impact of articles utilising CA as a methodology for investigating ICD (Dumay and Cai, 2014). Thus, the first part of the analysis will look at the development of research questions and hypotheses in ICD CA research.

2.2 Context and analytical constructs

According to Krippendorff (2013, p. 38) "texts acquire significance (meanings, contents, symbolic quantities, and interpretations) in contexts of their use". Thus, from the ICD content analysts' perspective the context is about IC, and excludes other meanings the text may have, either intended or unintended by the text's author(s). Similarly, other readers may have different contexts in mind when reading the texts, and thus their approach to analysing the texts will be different. An ICD content analyst applies what he or she knows about IC "to given texts, whether in the form of scientific theories, plausibly argued propositions, empirical evidence, grounded intuitions, or knowledge of reading habits (Krippendorff, 2013, p. 38)". Therefore, what theories researchers apply to analyse texts and how they design the analysis to draw out arguments and empirical evidence is important.

2.2.1 Theory. Guthrie *et al.* (2004b, p. 283) consider applying theory to the methodological processes of CA in ICD research, selecting stakeholder and legitimacy theory as the two "better known" theories. With stakeholder theory "an organisation's management is expected to take on activities expected by their stakeholders and to report on those activities to the stakeholders" Guthrie *et al.* (2004b, p. 283). Legitimacy theory is closely linked to stakeholder theory and posits "a company would voluntarily report on activities if management perceived that the particular activities were expected by the communities in which it operates" based upon a "'social contract' between the company and the society in which it operates" Guthrie *et al.* (2004b, p. 284). Similarly, Beattie and Thomson (2007, p. 130) advocate for the related concept of "positive accounting theory (PAT)" by arguing that PAT explains voluntary ICD when "company managers' interests are aligned with shareholders, IC information will be disclosed if it brings benefits to the company".

However, Abeysekera (2006, p. 69) argues that "'the social contract' may not be sufficient to explain differences in ICD between countries", which is often the differentiating feature of ICD CA studies (Dumay and Cai, 2014). In place of these theories Abeysekera (2006, p. 70) argues "an alternative, critical perspective on ICD is provided by the political economy of accounting (PEA) perspective". PEA contrasts legitimacy theory by positing that firms disclose IC in a way that sets and shapes "the agenda of debate, in order to mediate, suppress, mystify and transform the conflict between the firm and its social, economic, and political arrangements". However, the relevance of PEA is questioned because it has mainly been employed to analyse ICD in annual reports and has not been used to inform ICD in other texts such as analysts' reports (Abhayawansa and Abeysekera, 2009, p. 300).

According to Abeysekera (2006, p. 69), initial ICD CA studies provided "little or no theoretical basis for interpreting their findings" despite Johanson *et al.* (2001, p. 717) arguing "different theories about firms affect the way in which characteristics, and hence definitions of intangibles, are formulated". Additional theories identified by Abeysekera (2006, p. 69) for explaining ICD are the resource-based view of the firm, and signalling and agency theory. Although neither of these theories is discussed in any depth, the latter two theories explain more why firms lack motivation to disclose IC (see Bozzolan *et al.*, 2003). However, Abhayawansa and Abeysekera (2009, p. 298) argue that

signalling theory explains why "management will only make voluntary IC disclosure as long as there is a marginal benefit to be gained from reducing the information asymmetry in the market". Thus, there is little evidence of a consistent use of, or a prominent theory advocated, by ICD CA researchers.

2.2.2 ICD index. The development of an analytical construct is important for CA because it "ensure[s] that an analysis of given texts models the texts' context of use" (Krippendorff, 2013, p. 40). Additionally, it makes visible the "rules of inference that guide the analyst" and make "the context portable to other content analyses or similar contexts". Hence, the idea of a consistent and replicable ICD index could have the advantage of offering ICD CA analysts a consistent framework for analysing ICD in different contexts and allowing for the comparability of studies.

According to Abeysekera (2006), even though there were five major IC frameworks available to ICD CA analysts, only the framework developed by Sveiby (1997) received attention in early ICD CA studies (see Guthrie *et al.*, 1999) and was altered and improved upon by Guthrie and Petty (2000) in what has now become one of the seminal articles on ICD (Dumay and Cai, 2014). This framework was again modified by Guthrie *et al.* (2004b, p. 286) and is presented in Table II. Thus, from the evidence presented by Guthrie *et al.* (2004b, p. 286), in the early stages of ICD CA studies, it appears a consistent framework for ICD was being generally accepted.

In contrast, Beattie and Thomson (2007, p. 132) pointed out "that there is no consensus on a precise definition of IC" and to compound the problem, there is evidence that many synonymous terms have been used for different IC categories. For example, "employees" and "employee competence" are sometimes used in place of "human capital (HC)".

Beattie and Thomson (2007, pp. 132-134) classify these as high-level categories and low-level categories, with their analysis of ICD CA studies revealing 128 different lower-level categories. Thus, their evidence shows that other than an agreement on the three higher-level IC categories there is no consensus on exactly what the lower-level categories should be. Additionally, most studies do not give any detailed explanation of what a specific lower-level category means and thus coders and readers may not understand what is meant by each category and therefore "in the absence of explanation and transparency, interpretations of the findings across studies are potentially meaningless" (Beattie and Thomson, 2007, p. 139).

Steenkamp and Northcott (2007, p. 21) propose "it is useful to adopt previously used categories as far as possible to enhance comparability with other studies", but with reservation. They suggest that modifying existing categories can improve the accuracy of capturing the relevant issues, but this may cause difficulty with the comparability and synergy between studies. Similarly, Husin *et al.* (2012, p. 198) claim "it is more

1. Internal capital	2. External capital	3. Human capital
1. Intellectual property	7. Brands	14. Employee
2. Management philosophy	8. Customers	15. Education
3. Corporate culture	9. Customer satisfaction	16. Training
4. Management processes	10. Company names	17. Work-related knowledge
5. Information/networking systems	11. Distribution channels	18. Entrepreneurial spirit
6. Financial relations	12. Business collaborations	
	13. Licensing agreements	

Table II.
ICD framework

appropriate to read the whole annual report so that relevant information which does not meet the original set of IC items and indicators can be added to the index" to avoid doing a partial CA (Beattie *et al.*, 2004, p. 208). Alternately, ICD CA researchers could start with no framework at all and perform open, axial and selective coding of texts to look for meanings (see Parker and Roffey, 1997, pp. 228-229).

2.2.3 Unit of analysis. Krippendorff (2013, p. 84) defines unitising as "the systemic distinguishing of segments of text – images, voices, and other observables – that are of interest to an analysis". Additionally, "content analysts must justify their methods of unitizing, and to do so, they must show that the information they need for their analysis is represented in the collection of units, not in the relationship between the units, which unitizing discards" (p. 84). Hence, deciding not only what texts to code, but how to divide the text is an important consideration for ICD CA analysts. As Krippendorff (2013, p. 98) outlines:

> […] the first task in any empirical study is to decide what is to be observed as well as how observations are to be recorded and thereafter be considered data" and "recording units are units of description that collectively bear the information that content analysts process and provide the basis for statistical accounts.

Krippendorff (2013, p. 102) also supports that the content unit should be defined first as "context units that delineate the scope of information that coders need to consult in characterizing the recording units" because the context unit is to ensure the "original text can be reconstructed without loss". Thus, a content unit is what is being counted, and a context unit is where the content is located, for example, the Chairman's letter in the annual report (Oliveras *et al.*, 2008).

The issue of unitising is much debated in ICD CA studies. In the articles about CA as a methodology the pros and cons of keywords, sentences, paragraphs, pages and themes are debated (Guthrie *et al.*, 2004b; Beattie and Thomson, 2007; Cinquini *et al.*, 2012; Husin *et al.*, 2012) and no consensus seems to have been achieved. For example, Guthrie *et al.* (2004b, p. 288) support paragraphs as the most suitable unit because, they argue, people establish meaning based on paragraphs instead of particular words or a sentence. In contrast, Beattie and Thomson (2007, pp. 146-147) support sentences as the most reliable unit despite the potential "difficulty involved in allocating information to only one category". However, Husin *et al.* (2012) and Steenkamp and Northcott (2007) argue that all units, except theme, fail to analyse images containing IC, an issue which seems to have been sidestepped by researchers because of the subjectivity involved in coding pictures Guthrie *et al.* (2004b, p. 288). Steenkamp and Northcott (2007) offer a comprehensive example and discussion about the strength and weakness of different units, and argue that the more important issue is to identify the theme rather than use a particular recording unit. Husin *et al.* (2012, pp. 204-206) support this view and give an illustrative example of how they approach analysing IC themes.

To illustrate, the problem of using different coding units is highlighted by Abeysekera (2006, pp. 66-67), who compares and contrasts the differing results obtained by using a frequency count vs a line count in his study of Sri Lankan ICD using the same source documents: "Based on a frequency count external capital has emerged as the most reported category whilst based on line count human capital is the most reported category". The danger here is that choosing one unit over another may result in different interpretations (Abeysekera, 2006, p. 67), which negates the purpose of CA to uncover hidden meanings.

2.2.4 The quality and weighting of ICD. One issue specifically related to the discussion of the unit of analysis is the quality and weighting of ICDs. As Guthrie *et al.* (2004b, p. 289) suggest "studying the quality of disclosure by examining the relative emphasis on each theme, whether the disclosure is quantified or not, and the location of disclosure […] is the approach most likely to yield meaningful results". Similarly, Beattie and Thomson (2007, p. 141) advocate the use of a measure of IC quality rather than just the volume of ICD with the differentiation between quantitative and qualitative disclosures as the primary method of doing so. They identify several ICD CA studies where quantitative disclosures are weighted more heavily than qualitative disclosures (Bozzolan *et al.*, 2003). However, as with the ICD index, there is no consensus found as to whether or not any measure of quality or weighting should occur outside the realm of IC and/or accounting-based CA research.

2.3 Reliability of CA

It is important here to make a distinction between the reliability and validity of ICD CA studies. According to Krippendorff (2008, p. 350) reliability "is the extent to which data can be trusted to represent the phenomena of interest rather than spurious ones". Validity, on the other hand, occurs when "a measuring instrument […] measures what it purports to measure" (Janis, 2008, p. 359). Hence, the general concern for reliability is to ensure, especially where humans are involved, that the data presented is unbiased. This is generally done through the use of strong coding instructions along with some form of statistical agreement by way of "an agreement coefficient, one that is capable of measuring the agreements among the values or categories used to describe the given set of recording units" (Krippendorff, 2008, p. 352).

Of the recommendations from the ICD CA articles examined, Guthrie *et al.* (2004b) propose three ways to increase the reliability of CA by first including "disclosure categories from well-grounded relevant literature", develop "a reliable coding instrument with well-specified decision categories and decision rules", and "training coders and showing that coding decisions made on a pilot sample have reached an acceptable level". However, Guthrie *et al.* (2004b) do not elaborate on how to develop an agreement coefficient. Other articles offering advice on how to measure the reliability of ICD CA studies are almost silent on this issue and at times use the term reliability interchangeably with the term validity (e.g. Abeysekera, 2006, p. 68).

Additionally, while it is beyond the scope of this paper to outline the merits of each coefficient, the following comments outline Krippendorff's advocacy of Krippendorff's α. First, it appears that using Cronbach's α "is unsuitable for evaluating reliability" because it was developed for an "entirely different purpose" being "biometric and educational research" (Krippendorff, 2013, p. 308). Second, Scott's π is suggested as an alternative to Krippendorff's α but has since been discredited as too "liberal" an alternative (p. 308). Third, r_{wg} (J) is not mentioned by Krippendorff as an alternative reliability measure; as it was originally conceived as a measure of agreement and not a measure of reliability (LeBreton *et al.*, 2005) and thus it is argued that he would not support it.

2.4 Inferences

In CA research, inferences must be made about the meaning of the text so that it answers the research question(s) initially posed by the researcher (Steenkamp and Northcott, 2007, pp. 13-14). The application of inference is what makes CA different to other forms of empirical research because inferences are intended to draw what "may be hidden in the human process of coding" and can consist of three types being deductive, inductive and abductive inferences (Krippendorff, 2013, pp. 41-42).

Of these three types "deductive and inductive inferences are not central to content analysis" (pp. 41-42). The issue of inferences is little discussed in the articles examined, with Steenkamp and Northcott's (2007) article being the most comprehensive. However, they do not address the issue of the type of inference an ICD CA analyst makes.

A deductive inference is one that goes from the general to the specific and consists of a major premise, a minor premise and a conclusion. Deductive inferences have little place in the results of a CA study because they do not reveal what is contained in texts or their meaning. Conversely, researchers form inductive inferences on the basis of the probability of a scientific hypothesis being true, and therefore the conclusion drawn from these inferences could still be false. Thus, from an ICD CA analysis perspective a researcher could conclude that because HC is the most reported category of IC, that HC is the most important aspect of IC in a particular context. However, for CA studies, inductive inferences do not look for hidden meaning in the text, rather they base their findings on statistical correlations to determine if a relationship exists and/or some form of cause and effect exists.

In order to fulfil the requirements of looking for hidden meaning using CA as a research methodology, ICD CA analysts need to recognise that they must make abductive inferences that answer their research questions and hypotheses. As Krippendorff (2013, p. 42) outlines:

> […] if one has practical experience with infants' language acquisition, one might be able to infer children's ages from the sounds they make or from the vocabulary they use. Of course, one can make such inferences only with a certain probability, but the probability may be strengthened if one is able to take other variables (contributing conditions) into account.

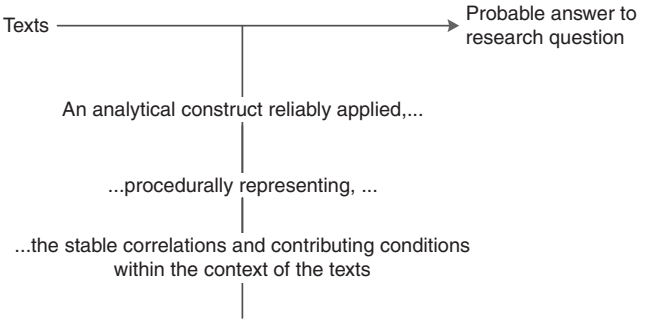
For example, Olsson (2001, p. 51) abductively infers the following from her study of HC disclosures in Swedish companies:

> None of the 18 largest companies disclosed more than 7.0% of HR information as a share of the total information in the annual reports in 1998. This finding reveals a lack of good examples in this group of companies. Very little meaningful changes had taken place during the past 10 years. It must be concluded that a good deal of what was said about the importance of disclosing HR information or having greater transparency regarding a company's human capital is largely lip service and not in accordance with reality.

In the above example, Olsson does not try to generalise the conclusion to all companies in Sweden, but rather relates her conclusion to the particular context she analyses. This may seem like a subtle difference compared to making an inductive inference. However, it is an important one because there is no scientific hypothesis used to determine correlations, rather just reasoning based on her ability to apply "a mixture of statistical knowledge, theory, experience, and intuition to answer [her] research questions from available texts" (Krippendorff, 2013, p. 43). Here Olsson uses her analysis to draw an inference based on the available evidence much in the way the great detective Sherlock Holmes would solve the case. In this case, if the companies in question lived up to their rhetoric about the importance of disclosing HC then Olsson would probably expect to find evidence of these disclosures, but she does not. Hence, the conclusion is the inference she makes to solve her case as outlined in Figure 1.

2.5 Validity of CA

Krippendorff (2013, p. 44) argues "*any content analysis should be validatable in principle*", but in practice this may be impossible, because CA "is intended to be acted upon in the



Source: Adapted from Krippendorff (p. 43)

Figure 1. Developing a probable answer to a CA research question

absence of direct observational evidence". For example, determining if a politician is lying can be hypothesised and inferred from a CA of the politician's speeches but may be impossible to prove. Krippendorff (p. 44) further argues for validation to "prevent analysts from pursuing research questions that allow no empirical validation or that yield results with no backing except by the authority of the researcher". For example, in ICD CA research there is no point concluding from a CA of annual reports that ICD disclosures in annual reports is rising, unless the CA can show that ICD is also on the increase in other forms, such as public IC reports or in practice as corroborating evidence.

The problem with CA is that it can be conducted with little concern for validity and merely to satisfy the curiosity of the content analyst and "such research contributes little to the literature on content analysis" or the phenomenon under investigation (Krippendorff, 2013, p. 44). This is also one of the current criticisms of CA from an ICD perspective as Dumay and Cai (2014, p. 264) observe, "the research into ICD using annual reports and other data sources has added little more than prove that companies are unwilling to disclose IC publicly to their stakeholders".

Another way of validating CA research is through *ex post facto* research (2013, p. 44). For example, Guthrie and Petty (2000, p. 241) found "the key components of intellectual capital are poorly understood, inadequately identified, inefficiently managed, and not reported within a consistent framework when reported at all". One way of following up this research would be to analyse specific ICD in another form of text, such as actual IC statements to determine whether or not public ICD was being embraced by companies and whether those that did disclose IC used a consistent framework. Unfortunately, as Dumay and Cai (2014) discovered, they could only identify two articles from the IC literature that analysed IC statements (Johanson *et al.*, 2006; Pedrini, 2007) and neither of these attempted to validate prior ICD work. However, both articles use additional data to validate their findings.

In this paper's review of the CA methodology articles finds little attention paid to the validity of ICD CA research. For example, Guthrie *et al.* (2004b) discuss reliability and validity synonymously while actually addressing the reliability issue and does not address validity as described by Krippendorff (2013). Abeysekera (2006) suggests "combining more than one complementary research method can improve the relevance and reliability of results, and hence the future credibility of ICD studies", which, in reality, addresses reliability. Additionally, Steenkamp and Northcott (2007), Beattie and Thomson (2007) and Husin *et al.* (2012) make no mention of the validity issue.

3. Results and a critique of CA as a research methodology

The purpose of this section is to answer research questions one and two: "how have different attributes of CA as a methodology been used for inquiring into ICD?" and "how might CA be used differently for inquiring into ICD?" by presenting insights from the analysis of CA studies according to our framework. In conjunction, this paper offers critique as to how these attributes might contribute to the conflicting results and limited contributions to IC knowledge and how to use CA attributes differently for inquiring into ICD. Table III presents a summary of results and the next subsections outline the results.

Issue	Findings
Research questions and hypotheses	Hypotheses in CA should emanate from the research question Only 19 of the 110 articles contained an explicit research question, and only 42 articles propose one or more explicit hypothesis There is little awareness by ICD CA researchers of the need to develop both a research question and related hypotheses
Theory	There is no generally accepted theory to help explain CA research The majority of articles that espouse the use of agency, legitimacy and stakeholder theories are aligned to positivist research and use explicit hypotheses to develop their research An alternative theoretical lens might be needed to understand why firms choose not to disclose IC
ICD index	ICD indexes used in ICD CA research are varied and derived from many different sources, with the most common being those based on the work of Guthrie *et al.* (1999) IC reporting models are generally ignored for the purpose of conducting ICD research and, therefore, questioning their usefulness Many CA ICD researchers do not pay much attention to the potential need to replicate studies in different contexts Using a predefined or modified ICD index only allows researchers to see inside the texts for specific issues rather than uncovering what they do not expect to find
Unit of analysis	In general there is an inconsistent application of counting the number of specific words, sentences or paragraphs, which is often justified by referring to previous research The debate about which context unit to use for ICD CA research appears to have fallen into the category of using IC terms based on a predetermined ICD disclosure index, generally adapted from another study to ascertain ICD incidences
The quality and weighting of ICD	A plethora of qualitative and quantitative schemes are used to categorise and apply weightings to ICD Using different weighting scales may also give different results. Thus, the conclusions of these articles could be empirically questioned
Reliability of CA	Many ICD CA researchers are oblivious to the need to test for coder reliability Most researchers using a coder reliability measure seem unaware of Krippendorff's α and the reasons why Krippendorff advocates it as the best test of coder reliability for CA research
Inferences	Abductive reasoning is not a core issue for ICD CA researchers and what we classify as abductive reasoning is accidental rather than intentional
Validity of CA	Few ICD CA pay attention to the validity of CA, albeit accidently rather than purposefully

Table III.
Summary of results

3.1 Research questions and hypotheses

As highlighted earlier, CA research is different from other forms of qualitative research because the data, being texts, are created before research commences. However, as with all research, research questions play an important role in developing CA research. As Krippendorff (2013, p. 38) outlines, in CA, research questions should have the following characteristics:

(1) they are believed to be answerable (abductively inferable) by examinations of a body of texts;

(2) they delineate a set of possible (hypothetical) answers among which analysts select;

(3) they concern currently inaccessible phenomena; and

(4) they allow for (in)validation, at least in principle, by acknowledging another way to observe or substantiate the occurrence of the inferred phenomena.

Additionally, Krippendorff (2013, pp. 37-38) outlines that in CA the research question and hypotheses are different aspects of his framework:

> Research questions are the targets of the analyst's inferences from available texts. Generally, such questions delineate several possible and initially uncertain answers. In this respect, a research question is analogous to a set of hypotheses.

Thus, hypotheses in CA should emanate from the research question.

Another important distinction to make is that hypotheses in CA are fundamentally different to scientific hypotheses, which rely on different forms of reasoning and although both can rely on a large number of observations, scientific "hypothesis testing is to choose between two conflicting hypotheses about the possible value of a population parameter" (Kenkel, 1989, p. 389). Alternately, CA hypotheses "pertain to phenomena that are not observed during a content analysis, phenomena that are outside the texts and thus retain their hypothetical character until confirmed by validating incidences" (Krippendorff, 2013, p. 37). Therefore, the "counts" of incidences are important, but what they uncover and infer is more important.

For example, if a CA analyst asks the research question "What is the extent of IC in a group of company annual reports?" a hypothesis may be "a substantial amount of human capital (HC) will be disclosed because the company is a renowned advocate of developing its human resources". Hence two other hypotheses could be "structural capital (SC) will be disclosed less than HC because of the expected emphasis on HC" and "relational capital (RC) will be disclosed less than HC because of the expected emphasis on HC".

Therefore the hypothesis is that the CA studies will all have at least one explicit research question and one or more hypothesis. However, as Table IV shows, only 19 of the 110 articles contained an explicit research question, and only 42 articles propose one or more explicit hypothesis. In fact, only two articles, Dumay and Tull (2007) and Sihotang and Winata (2008, pp. 69-71), had both an

Table IV. Research questions and hypotheses

Explicit research question (RQ)	19	17.3%
Explicit hypothesis (H)	42	38.2%
Explicit RQ&H	2	01.8%

explicit research question and a set of hypotheses. However, the hypotheses used by Dumay and Tull (2007) were based on choosing between two conflicting hypotheses rather than trying to infer something from what is hidden in the texts. Most hypothesis testing in ICD is similar, taking a positivist perspective that attempts to predict the trend of ICD and test the effect of factors that influence the extent of ICD (Bozzolan *et al.*, 2003).

However, while many articles did not have an explicit research question, the remaining articles disclosed non-explicit research questions and/or one or more hypothesis. For example, Abeysekera and Guthrie (2004, p. 257) disclose in their article "The aim of this study is to understand the degree of emphasis these firms place on employees and IC, via HCR". While this states the aim of their paper, the authors do not reveal a specific research question. However, the authors could rewrite the aim as "What degree of emphasis do firms in Sri Lanka place on employees and IC, via HCR?" So while the authors can paraphrase the aims and purpose of these articles, the reader cannot assume these to be the research questions because the authors do not explicitly state them.

Among the 19 articles with specific research questions, ten contain research questions generally related to uncovering the extent of ICD (e.g. Brennan, 2001, p. 427), in keeping with Krippendorff's (2013, pp. 12-13) characteristics of CA research questions. However, several articles have research questions that cannot be in keeping with these characteristics. For example, Sihotang and Winata's (2008, p. 66) research question "To what extent do company characteristics such as the industry category, age and size have a relationship with the level of IC disclosures?" is based on research that develops generalisations from their observations. Therefore, these types of research questions create studies that can potentially use abductive reasoning rather than inductive reasoning to answer research questions. A full set of research questions is presented in Appendix 1.

Similarly, some studies contained no hypotheses, yet statistical tests are performed. For example, Rashid *et al.* (2012, p. 32) report, based on their analysis, "In line with Rimmel *et al.* (2009), the present study finds that company age has a negative and significant influence on the IC disclosure score in Model 1" despite not disclosing a specific hypothesis. Again, these authors are clear about what they are testing and they could possibly write the null and alternate hypothesis as:

H0. Age does not have a significant influence on the IC disclosure score in Model 1.

Vs

HA. Age does have a significant influence on the IC disclosure score in Model 1.

So while the hypotheses can be written for these studies in scientific form, they do not use the type of hypothesis advocated by Krippendorff (2013, p. 37) and will thus use inductive rather than abductive reasoning to present findings and conclusions.

In the studies specifying specific hypotheses, there was a clustering of similar attributes attempting to understand what drives firms to disclose IC voluntarily with the two main issues being industry type and company size (see Table V). Alternately, there were also many different hypotheses put forward to try to predict what attributes influenced the extent of ICD.

Selected citation	Hypothesis	Number of hypotheses
Petty and Cuganesan (2005, p. 43)	Industry type	18
García-Meca *et al.* (2005, p. 71)	Company size	18
Gan *et al.* (2013, p. 57)	The impact of the firm's auditor or audit committee	11
Li *et al.* (2008, p. 140)	Company ownership	10
Ferreira *et al.* (2012, pp. 282-284)	Profitability	9
Cerbioni and Parbonetti (2007, p. 797)	The make-up of board committee	7
Cordazzo (2007, p. 295)	Listing status	7
Jindal and Kumar (2012, p. 229)	Leverage	7
Ferreira *et al.* (2012, p. 284)	The level of IC	7
Branco *et al.* (2010, p. 263)	The trend of ICD	6
Cordazzo (2007, p. 294)	Firm age	6
Li *et al.* (2008, p. 140)	Role duality	6
Singh and Kansal (2011, p. 311)	The value of IC	5
Joshi (2012, p. 229)	The organisation structure	4
Yau *et al.* (2009, p. 22)	The growth of the organisation	3
Oliveira *et al.* (2006, p. 17)	Internationalisation	3
Abeysekera (2011, pp. 324-325)	External economic status	3
Orens *et al.* (2009, pp. 1538-1539)	Internal economic status	4
Dammak *et al.* (2008, p. 422)	Firm performance	2
García-Meca and Martínez (2007, p. 63)	Firm risk	2
Bozzolan *et al.* (2006)	Other hypotheses[a]	12

Note: [a]Other hypotheses that were mentioned once

Table V.
Hypotheses

As a result, the analysis infers that there is little awareness by ICD CA researchers of the need to develop both a research question and related hypotheses, and thus the explicit intent of ICD CA research is not clear. However, it is clear that ICD CA researchers have developed research questions and hypotheses based on inductive reasoning rather than abductive reasoning as advocated by Krippendorff.

3.2 Context and analytical constructs

3.2.1 Theory. As highlighted, when developing the research framework, there is no generally accepted theory to help explain CA research. This is due to the abductive nature of CA research – if a theory is used then it is expected that the researchers will test the theory. However, that has not prevented ICD CA researchers from applying theory to their research. As shown in Table VI, there is a wide variety of theories utilised. Among the 48 articles that weave theory into their literature reviews or analysis the top five theories are agency theory, legitimacy theory, resource-based view of the firm, stakeholder theory and signalling theory. These theories are also widely used in the accounting and management literature to explain other forms of disclosure, such as social and environmental information (e.g. Archel *et al.*, 2009) as well as IC disclosure in non-CA-based IC research (e.g. An *et al.*, 2011).

Unsurprisingly, the majority of articles that espouse the use of agency, legitimacy and stakeholder theories are aligned to positivist research and use explicit hypotheses to develop their research (e.g. Cordazzo and Vergauwen, 2012). Again, this goes against the use of abductive reasoning. Why have so many researchers undertaken CA

Agency theory	14
Legitimacy theory	11
Resource-based view	7
Stakeholder theory	7
Signalling theory	6
Information asymmetry	3
Institutional theory	3
Political economy of accounting theory	1
Political cost	1
Human capital theory	1
Capital market theory	1
Economic theory	1
Cost of disclosure theory	1
Cost-benefit theory	1
Structuration theory	1
Modern finance theory	1
Agenda-setting theory	1
Liquidity-based theory	1
Decision-usefulness perspective	1
Broad system-wide perspective	1
Internalisation theory	1
Capital market theory	1
Critical theory	1
Fuzzy set theory	1

Table VI.
Type and number of theories used in ICD articles

research based on evaluating inductive hypotheses when Krippendorff (2013, p. 42) states clearly "Deductive and inductive inferences are not central to content analysis"? However, these theories can also be the basis of inductive research questions (Guthrie and Abeysekera, 2006, p. 257) from a legitimacy theory perspective to measure the extent of ICD:

> Legitimacy theory is closely tied to the reporting of IC and to the use of content analysis methods as a measure of such reporting. Companies are more likely to report on their IC if they have a specific need to do this. This may happen when companies find themselves unable to legitimise their status on the basis of the hard assets that are traditionally recognised as the symbols of corporate success. The extent of IC reporting is, at this juncture, best measured using content analysis. Thus, legitimacy theory, IC reporting, and content analysis are linked.

However, some CA ICD studies find that the term "intellectual capital" is not mentioned in the analysed texts (Bontis, 2003; Abeysekera and Guthrie, 2005) and the extent of ICD is relatively low (Guthrie and Petty, 2000; Cordazzo and Vergauwen, 2012). Thus, an alternative theoretical lens might be needed to understand why firms choose not to disclose IC because not many firms do so explicitly, nor is it extensively hidden in their texts.

3.2.2 ICD index. As Table VII indicates, the ICD indexes used in ICD CA research are varied and derived from many different sources, with the most common being those based on the work of Guthrie *et al.* (1999) (25), Sveiby (1997) (12) and Guthrie and Petty (2000) (9). However, as Petty and Guthrie (2000, p. 245) outline, they "followed the contemporary classification scheme for intangibles derived from Sveiby's (1997) intellectual capital framework: internal structures (organisational capital); external structures (customer/relational capital); and employee competence (human capital)".

Miscellaneous or not explicitly mentioned	45
Guthrie *et al.* (1999)	25
Sveiby (1997)	12
Guthrie and Petty (2000)	9
Bontis (2003)	7
Guthrie *et al.* (2004b)	5
Bukh *et al.* (2005)	4
MERITUM (2002) model	2
Japanese IC Guideline	1
Indicators from the Global Reporting Initiative (GRI)	1
Schmalenbach Gesellschaft Work Group on Financial Accounting in Kristandl and Bontis (2007)	1

Table VII.
ICD indexes used

Thus, most studies use Sveiby's (1997) framework in some form or another, and there have been no substantially new ICD indexes (see Table I); rather Sveiby's framework appears to be modified not replaced.

Table VII also highlights the use of specific reporting models such as the MERITUM (2002) and Japanese IC frameworks. However, it is noticeable how little influence these models have had on ICD research in general. This infers that IC reporting models are generally ignored for the purpose of conducting ICD research and, therefore, questioning their usefulness.

Of the studies analysed, 39 reported changing an existing framework for their study. A common reason for changing the index was to make it more in keeping with the context of the research project. For example, Lee *et al.* (2007, p. 62) "added a further 15 extra items (not contained in previous indices) after further consideration for the Australian socio-political and economic environment, and healthcare system". The advantage of using an unmodified ICD index is that it can facilitate comparison between different research projects. However, the accuracy and comparability of CA results decrease as the number of terms used increases because they are utilised differently and can thus have conflicting meanings (see Lock Lee and Guthrie, 2010). Additionally, there are 45 articles using ICD indexes based on different sources or that do not disclose the source, which infers that many CA ICD researchers do not pay much attention to the potential need to replicate studies in different contexts. It appears researchers are more often comparing "apples with oranges" than "apples with apples".

Some researchers develop ICD indexes from different sources, showing a commendable degree of innovation in their research. For example, Oliveira *et al.* (2006) deploy an index based on analysing the Management Report and Chairman's Letter in annual reports. This kind of ICD index is constructed based on the practical use of specific texts rather than just analysing entire annual reports, which has been questioned by several researchers (Abeysekera, 2006; Dumay, 2014). Another innovative example is Jindal and Kumar's (2012) ICD index, developed from the IC literature, with which they filter data based on the Indian company Infosys' annual report. This method of construction has the same difficulty of comparability with other research. While innovative, these two ICD index examples may not be suitable for analysing texts in other contexts.

Another innovative ICD index is outlined by Dumay and Lu (2010), who use an open coding method based on grounded theory. The analysis of company reports and news media revealed three major themes, job security, employee commitment and legitimacy. The three themes form the basis of their analysis of HC management practices of an

Australian bank. Thus, they have developed an index that other researchers can use even though it cannot be guaranteed to be suitable to other organisations. Additionally, Dumay and Lu (2010, p. 76) outline the process used to identify the three major themes so the process can be replicated by other researchers. However, they use this method to uncover what is hidden in the texts rather than looking for predefined IC elements – this allows for the discovery of specific elements but prevents hidden elements from coming forward.

This last point is important because CA is concerned with finding what is hidden in the texts rather than what the researcher wants to find. Using a predefined or modified ICD index only allows researchers to see inside the texts for specific issues rather than uncovering what they do not expect to find.

3.2.3 Unit of analysis. Only in one case is the content unit and context unit the same, being price sensitive announcements to the Australian Stock Exchange (ASX) (Dumay and Tull, 2007). The results in Table VIII show 68 articles reveal that terms, items, attributes or themes are used to count ICD (Boedker *et al.*, 2004). Typically, these articles use an ICD index to count ICD incidences rather than count the number of specific words, sentences or paragraphs, which is often justified by referring to previous research (Bozzolan *et al.*, 2006). Therefore, one or more ICD incidence can be found in a sentence, paragraph, line or context unit. However, some articles do not exactly disclose the content unit used. For example, Uyar and Kiliç (2012, p. 366) do not reveal the content unit used: "The information gathering process was conducted by content analysis in annual reports of the sample firms which were downloaded from corporate web sites".

Another issue discovered when classifying the content unit was the way researchers apply it. The analysis of words is a prime example because, even though 18 articles use words, the researchers apply their analysis of words in different ways. For example, Li *et al.* (2008, p. 142) develop an intellectual capital word count (ICWC) measure to capture the number of words dedicated to IC in an annual report because "Words are the smallest unit of measurement for analysis and can be expected to provide the maximum robustness to the study in assessing the quantity of disclosure". However, some studies use an index of IC words and phrases, and specifically search for exact matches in the texts aided by a computer (Brüggen *et al.*, 2009). The problem with this kind of analysis is that it potentially misses what is hidden and only shows up that which is intentionally disclosed. This is not consistent with the basic principles of CA.

In summary, the debate about which context unit to use for ICD CA research appears to have fallen into the category of using IC terms based on a predetermined ICD disclosure index, generally adapted from another study to ascertain ICD incidences. The main argument for this approach is that IC can appear anywhere in the text and is not necessarily confined to one content element. Thus, it seems rational to count incidences of IC as terms, items, attributes or themes. However, doing so can only

Recording unit	Terms/items/attributes/themes/not specific	63
	Words	18
	Sentences	15
	Line count	4
	Paragraphs	2
	Context unit	1

Table VIII.
Units of content analysis

offer a narrow predetermined view of IC rather than attempting to discover something hidden. For example, the analysis could not find any article that performed a raw word frequency count to find out the most commonly used words and/or phrases in the texts. This might expose additional words, terms, items, attributes or themes hidden in the texts, potentially providing additional insights beyond a predetermined ICD index.

3.2.4 The quality and weighting of ICD. What makes one form of disclosure different from another or more valuable than another and how does one justify attaching a higher value to the disclosure? The issue of disclosure qualities and weighting is an interesting aspect of ICD studies because it is highly subjective (Cerbioni and Parbonetti, 2007, p. 804). Our evidence shows a plethora of qualitative and quantitative schemes to categorise and apply weightings to ICD.

As Table IX shows, eight articles made a qualitative assessment of ICD without adding a numerical weighting score. However, apart from Dumay and Tull (2007, p. 240) and Lock Lee and Guthrie (2010, p. 11), who propose that news-tenor (positive, neutral and negative) has an impact on share prices, the remaining articles mainly use the categories as part of a wider statistical analysis. For example, Cerbioni and Parbonetti (2007, p. 797) test the hypothesis "*Ceteris paribus*, there is a negative association between board size and quantity and quality of intellectual capital voluntary disclosure" and Kristandl and Bontis (2007, p. 579) examine whether "There is a negative association between COEC and the level of voluntary disclosure". Thus, different studies use subjective qualitative classifications of ICD depending on the different research questions and hypotheses examined.

Table X shows the extent to which articles added a weighted value to a particular ICD beyond a dichotomous rating (e.g. 0 or 1, or "no" or "yes") and assigned to any specific disclosure. Appendix 2 outlines the full detail of each scale. Typically, each weighting scale begins with a zero for no disclosure, although several scales do not record the zeros. This generally does not matter because the purpose of most of these articles is to prefer one type of ICD over another and develop a measure based on the sum of the weights assigned to each ICD found in the text. These measures help answer the hypotheses using statistical inferences. In support, Appendix 2 identifies the articles that specifically outline the formulae used to sum up their respective disclosure weightings.

Additionally, Appendix 2 reveals that the classifications of what merits a score of 0, 1, 2, 3, 4 or 5 is subjective and inconsistent. The simpler scales, with three weights, commonly attribute values to qualitative and quantitative disclosures. More expansive scales of four weights or more differ by typically distinguishing between numerical (non-financial) and financial disclosures. However, Kristandl and Bontis (2007) use a scale from 1 (qualitative, unspecified) to 4 (quantitative, specified) to develop a "VRSCORE which measures and ranks the level of corporate disclosure on an ordinal scale". As a result, it infers that once the weighting scales go beyond the simple distinction between quantitative and qualitative disclosures (as a 1 and a 2) the inconsistency between the weights used makes it difficult to compare research using different scales.

Similarly, as with Abeysekera's (2006, p. 67) contention that using different content units for analysis may give different results, this suggests that using different weighting scales will also give different results, empirically questioning these articles' conclusions. However, one could also argue that these articles may reveal new insights that may otherwise have gone unnoticed if they use the same scales in the same way.

Table IX.
Qualitative categorising ICD

Article	Positive	Neutral	Negative	Historical	Forward-looking	Narrative	Quantitative	Visual	Monetary	Non-monetary	Non-time specific	Fact	Merely managerial perception	Financial
Cerbioni and Parbonetti (2007, p. 62)	X		X	X	X									
Dumay and Tull (2007, p. 240)	X	X	X											
Lock Lee and Guthrie (2010, p. 13)	X	X	X											
Abhayawansa and Guthrie (2012, p. 404)	X	X	X	X	X		X	X	X		X			
Kristandl and Bontis (2007, p. 581)				X	X									
Abeysekera (2011, p. 321)						X	X	X						
Striukova *et al.* (2008, p. 305)						X			X	X				
Campbell and Abdul Rahman (2010, p. 62)						X	X					X	X	X

Note: This is self-explanatory "X" denotes which attribute (column heading) applies to which paper

3.3 Reliability of CA

From the articles analysed 69 do not disclose anything about coding reliability measures even though this is an essential element of the CA methodology. Of the 41 papers that consider reliability, 26 articles use the agreement of different coders, and two articles (Abdolmohammadi, 2005; Bezhani, 2010) use pilot testing with coders rather than a reliability coefficient as endorsed by Krippendorff (2008, p. 352). In these cases the level of agreement is subjective.

Only 11 articles specifically disclose using a reliability coefficient, and only five articles utilise Krippendorff's α, which Krippendorff (2013, p. 278) advocates as the "most general agreement measure with appropriate reliability interpretations in content analysis". In addition to these articles, three articles use Scott's π (Bozzolan *et al.*, 2006; Abeysekera, 2010, 2011), two use Cronbach's α (Kristandl and Bontis, 2007; Orens *et al.*, 2009) and one uses James *et al.*'s (1984) r_{wg} (J) (Gerpott *et al.*, 2008). This demonstrates that the majority of articles measuring reliability do not use the coder reliability coefficient specifically designed for CA research (Krippendorff, 2013, p. 305).

The analysis and evidence infers that many ICD CA researchers are oblivious to the need to test for coder reliability. Additionally, most researchers using a coder reliability measure seem unaware of Krippendorff's α and the reasons why Krippendorff advocates it as the best test of coder reliability for CA research. This problem is not unique to ICD CA research because other studies have found a general lack of reporting of reliability assessments with a similar low usage rate of Krippendorff's α (Krippendorff, 2013, p. 301). This suggests that ICD CA researchers have not emphasised reliability measures because the articles reviewed on how to apply CA to ICD research do not emphasise the issue either.

3.4 Inferences

As outlined previously, Krippendorff advocates abductive inferences are central to CA. However, because of the way researchers have developed their research questions the analysis does not reveal any articles following the CA methodology rigorously using abductive reasoning. As part of the research process, the authors placed PDF copy of all 110 articles into a common folder in an Endnote library. Then, using the full search capability of Adobe Acrobat X Professional a search for the terms "abductive" and "abductively" did not find a single use of the term. This evidence infers that abductive reasoning is not a core issue for ICD CA researchers and any abductive reasoning found in the articles is accidental rather than intentional. When examining the articles for abductive inferences only two examples are found, as detailed below (Dumay and Lu, 2010, pp. 70, 74):

> RQ: Is the rhetoric of HC disclosure achieved in practice?
>
> The paper illustrates how highly exposed HC disclosures are to scrutiny by both internal and external stakeholders and if the rhetoric is not transformed into practice how the disclosures can be used as a weapon by adversarial stakeholders to attack

Article example	Total
Three weights	12
Four weights	12
Five weights	3
Six weights	2

Table X.
Weighting of ICD

> the organisation and/or attempt to change the balance of power between management and employees.

Thus, in the above example, the rhetoric of HC disclosure is the focus of the research question and the answer also exposes the hidden inference that stakeholders can use empty rhetoric for their own ends (April *et al.*, 2003, pp. 165-166):

> RQ: Is IC at all relevant or meaningful to South African mining companies and, if so, what are they doing in this regard?
>
> From these findings it is concluded that mining companies value intellectual capital but lack the appropriate systems and structures to manage intellectual capital meaningfully.

Similarly, April *et al.* (2003) infer that IC is meaningful (valued) by South African (SA) mining companies and uncovers the hidden issue of not having appropriate SC to manage their IC. Again, a critical element of CA research is not consistently found in the articles, which further reinforces the inference that ICD CA researchers are oblivious to the way in which to conduct robust CA studies as advocated by Krippendorff (2013).

3.5 Validity of CA

Of the 110 articles, only nine pay attention to the validity of CA, albeit accidently rather than purposefully, in line with implementing CA methodology rigorously. The analysis found seven studies that used corroborating evidence to support their findings and two that developed *ex post facto* research, specifically comparing IC reports with other texts (Johanson *et al.*, 2006; Pedrini, 2007). However, as the concept of CA reliability is not applied, the lack of attention to validity is unsurprising. Coupled with Krippendorff's (2013, p. 44) earlier comment that much CA research is ad hoc, this apparent lack of concern for validity suggests that rigorous ICD CA studies are mostly absent.

Because there is no overt reference to the validity of CA as part of a rigorous methodology, the examples in Table XI use different data sources, such as interviews and surveys, as evidence for validity to confirm or even question the findings of a text analysis. An example of the latter is Carrington and Tayles' (2012, p. 244) use of interviews alongside an analysis of internal and external documents. Their initial finding was "there is no formal recognition of the construct of IC", but after examining interview texts, they found "the embedded practices within the organisations suggested the presence of such". From this seemingly conflicting discovery, it exemplifies how testing for validity can also uncover hidden data that the analysis of the texts may not reveal. Similarly, April *et al.* (2003) show how external capital is

Table XI. Articles using another source of data for validity

Article	Corroborating evidence	Source
Abeysekera (2008)	Interviews	Human resource executives
April *et al.* (2003)		Senior individuals
Murthy and Abeysekera (2007)		HR heads
Carrington and Tayles (2012)		Top managers
Khan and Ali (2010)	Questionnaires	Stakeholders
Ax and Marton (2008)		Chief financial officers
Bezhani (2010)		Not clear

reported most in the annual reports of SA mining companies. However, their data from interviews with senior individuals shows how HC is rated highest (April *et al.*, 2003, pp. 178-179) thus contradicting their analysis of annual reports. Therefore, testing for validity can either confirm or contradict a text analysis, both of which are also relevant findings.

4. The future of CA as a research methodology

The purpose of this section is to answer the last research question "What is the future for CA as a research methodology for investigating ICD?" to do so, this discussion returns to the original purpose of the paper, which is to build on Dumay and Cai's (2014) prior research to provide a deeper analysis of the problems associated with using CA as a research methodology for ICD. The main finding of the analysis in this paper is that there is much inconsistency and an absence of rigour in ICD CA studies applying CA as a methodology. This is a likely reason why contemporary ICD CA studies add little new knowledge and develop conflicting results (see Dumay and Cai, 2014, Table VI).

Applying rigour to future CA ICD studies will not automatically advance CA-based ICD research. However, it is a necessary first step. Should CA researchers be able to prove that their research is rigorous, replicable and can make valid inferences from texts, then there should be a subsequent decline in conflicting results and more innovative research which examines specific research questions. Only, then can researchers demonstrate that CA is the right methodology to use to answer these questions and allow the research question drive the methodology, rather than the methodology drive the research questions.

The analysis of the 110 CA ICD articles discovers that the main reason for the lack of rigour is that few studies – in fact not a single example – follow exactly the most basic design of CA (see Krippendorff, 2013, p. 83), although a few good examples come close (e.g. Bozzolan *et al.*, 2003; Li *et al.*, 2008). These articles mix the CA aspects of their research with other variables, instead of taking a more structured approach, which sees the result of the CA as input into a wider research undertaking (see Figure 2).

Many researchers use CA intermixed with other data to answer a larger research question, which requires data from sources beyond analysing texts. Once the CA component of the study is completed, and the CA results prove reliable and are validated to answer the initial research question, then the researcher can begin to design a further study "exceeding content analysis" to "investigate the worlds of others" and "other phenomena" (Krippendorff, 2013, pp. 94-97). Therefore, the results of the initial CA become the data for answering a new research question and testing an associated set of hypotheses (Figure 3).

Unfortunately, the studies examined using CA as the basis for developing inferences from incorporating variables and data beyond the texts do not follow this research design. The fact that not one of the articles analysed even alludes to developing abductive inferences is ample evidence of this claim. Instead, ICD researchers mix CA data with data and variables from other sources in the name of a CA study when, in fact, two studies are required: the first to provide validated CA data for inclusion into the second wider study based on other methodologies. In contrast, CA studies measure the extent of CA rather than the significance of ICD, and this was enough for the researcher to justify including CA data with data from other worlds and phenomena.

It is thus recommend that an assessment of relevance be included when examining the validity of any ICD CA to ensure the data passing from an ICD CA to the wider

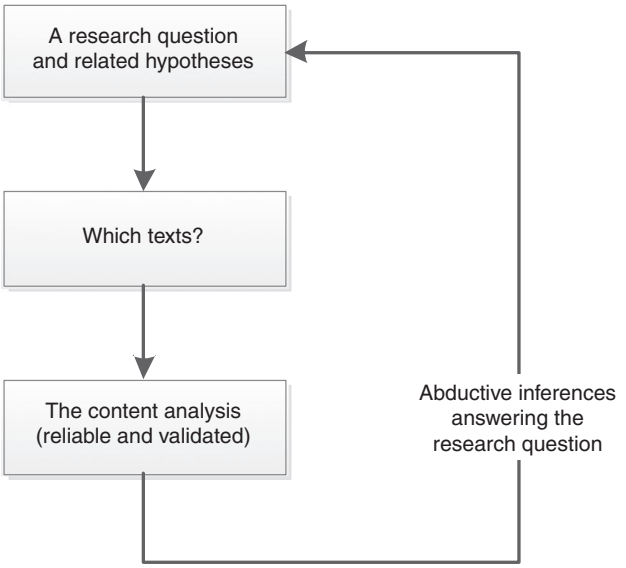


Figure 2.
Basic CA rigour

Source: Adapted from Krippendorff (2013, p. 83)

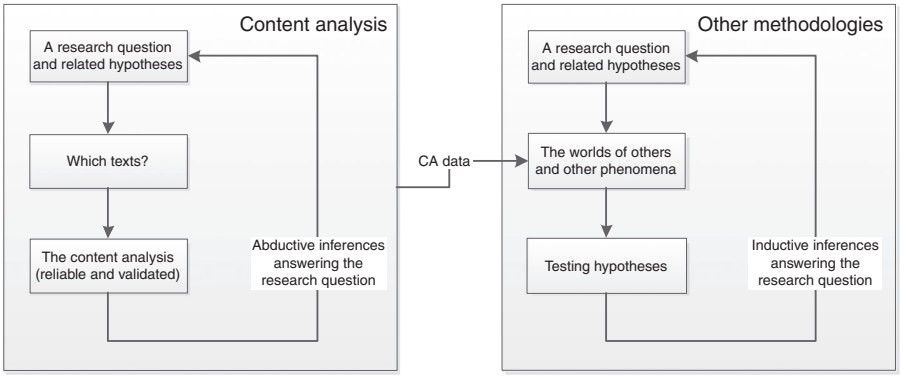


Figure 3.
Designs exceeding CA

Source: Adapted from Krippendorff (2013, p. 97)

analysis is sufficiently robust. Even better would be for researchers to choose the most relevant text to analyse for ICD, rather than blindly following previous studies, most of which use annual reports (Dumay and Cai, 2014). For example, Dumay and Lu (2010) performed their CA research "utilising corporate social responsibility (CSR) reports, newsletters, annual reports, and other publicly available information, with an emphasis on media reports" as these texts were deemed the most relevant to answering their research question.

While the issue of validity is paramount, what takes place between choosing a suitable set of texts to analyse and validity remains problematic. Most studies lack an explicit research question and hypotheses suited to abductive inferences, even though

abductive reasoning is essential to CA research. Additionally, the common view that ICD CA research needs to utilise an ICD index also limits its ability to uncover hidden meanings in text because most researchers are inevitably looking for the same thing. Meanwhile, the subjectivity and inconsistency between the unit of analysis which, when applied to the same data, can bring about different results coupled with the subjectivity applied to the quality and weighting of ICD, highlights that researchers have yet to decide on how to measure ICD consistently and reliably.

5. Conclusion

In summary, this research has discovered inconsistency in regards to how ICD researchers apply CA as a methodology. These findings are consistent with the previous literature as outlined in the review based on the method articles by Guthrie *et al.* (2004b), Abeysekera (2006), Beattie and Thomson (2007), Steenkamp and Northcott (2007) and Husin *et al.* (2012). First, the analysis finds no generally accepted approach for developing research questions and/or hypotheses. Similarly, while positivist theories and Guthrie *et al.*'s (1999) ICD framework dominate, there is substantial variation in the application of units of analysis and the quality of ICD to make it almost impossible to compare findings and form a basis for replicable studies which might add to the validity of the ICD project. Similarly, the tools needed to ensure reliability and validity seem not to have been of concern to many ICD researchers, and this requires a return to the methodological drawing board as far as CA-based ICD studies are concerned.

5.1 Implications for research

While the findings paint a bleak picture of ICD CA research, there is light at the end of the tunnel. That is it is advocated that ICD CA researchers need to go back to the drawing board and ensure that future studies rigorously apply the designs outlined in Figures 2 and 3 as a start (see also Krippendorff, 2013, pp. 82-97 for an in depth logic of CA design). While it is possible that too much "rigour perhaps lead[s] to rigor mortis" (Otley, 2003, p. 324), the patient is barely breathing. A return to a more rigorous application of Krippendorff's methodology might revive the patient to a state of health whereby variations of applying CA to understand ICD will add to IC knowledge. Krippendorff's design logic may not be perfect but it is a starting point. In its current state, ICD CA research needs to take a few steps back, before it can move forward. If ICD CA researchers can accomplish this, then there is an opportunity to undertake rigorous research to develop reliable and valid outputs that add to new knowledge about IC. Once this is accomplished, researchers can also develop ideas on how use CA in the future within the IC field to develop interesting and publishable papers (see Dumay, 2014, pp. 82-84).

5.2 Possible areas for future research

Once researchers have come to grips with developing reliable and valid research outputs, this then opens up a multitude of possibilities for IC research. As Dumay and Cai (2014, p. 284) outline, research needs to be innovative and deal with IC research not previously investigated. For example, researchers appear to be stuck in analysing the same theories, variables and frameworks using similar data sources. This type of research has little opportunity of making a significant contribution to knowledge because it is at best incremental. Therefore, no matter how good Krippendorff's α is, no matter what confirmatory evidence establishes validity, if the research does not answer a pertinent research question, then it will have little if any impact.

For example, the last two decades have seen organisations developing additional communication channels to engage with and disclose information to stakeholders, especially via the internet and corporate web sites. Therefore, just because another researcher uses a data source such as annual reports, their future use in ICD research projects cannot be justified unless the researcher(s) can argue convincingly why it is the best data source for revealing new IC knowledge. Similarly, researchers need to broaden the scope of theories used to investigate IC and may be even go back to the basics of communication theory which underpins all forms of CA (Krippendorff, 2013).

Additionally, researchers should also be concerned with IC practices inside organisations (see Guthrie *et al.*, 2012; Dumay and Garanina, 2013). Therefore, research analysing the various texts which organisations produce as they manage and create value on a day to day basis is one way of achieving this. Internal documents, interviews, field notes, tweets, blogs and news media are examples of data sources produced on a day-to-day basis which may prove valuable in understanding how IC impacts organisations today rather than in the past. There appears to be little value in analysing formal reports or disclosures, which by the time a person receives the text, the text has lost its currency, and thus has lost its value.

5.3 Limitations

The main limitations of the research are the chosen sample of CA-based ICD articles and the adoption of the Krippendorff's framework. However, the authors feel that they have identified the main corpus of CA-based ICD studies and since Krippendorff is the only recognised comprehensive text on CA as a methodology, the most appropriate data and framework possible for the analysis is used.

Notes

1. According to Google Scholar as at 15 February 2014 (http://scholar.google.com.au/).
2. In this paper we cite Krippendorff several times and he often uses italics to emphasise his points. Where italics are included in the citations these are all contained in the original text.

References

Abdolmohammadi, M.J. (2005), "Intellectual capital disclosure and market capitalization", *Journal of Intellectual Capital*, Vol. 6 No. 3, pp. 397-416.

Abeysekera, I. (2006), "The project of intellectual capital disclosure: researching the research", *Journal of Intellectual Capital*, Vol. 7 No. 1, pp. 61-77.

Abeysekera, I. (2008), "Intellectual capital disclosure trends: Singapore and Sri Lanka", *Journal of Intellectual Capital*, Vol. 9 No. 4, pp. 723-737.

Abeysekera, I. (2010), "The influence of board size on intellectual capital disclosure by Kenyan listed firms", *Journal of Intellectual Capital*, Vol. 11 No. 4, pp. 504-518.

Abeysekera, I. (2011), "The relation of intellectual capital disclosure strategies and market value in two political settings", *Journal of Intellectual Capital*, Vol. 12 No. 2, pp. 319-338.

Abeysekera, I. and Guthrie, J. (2004), "Human capital reporting in a developing nation", *The British Accounting Review*, Vol. 36 No. 3, pp. 251-268.

Abeysekera, I. and Guthrie, J. (2005), "An empirical investigation of annual reporting trends of intellectual capital in Sri Lanka", *Critical Perspectives on Accounting*, Vol. 16 No. 3, pp. 151-163.

Abhayawansa, S. and Abeysekera, I. (2009), "Intellectual capital disclosure from sell-side analyst perspective", *Journal of Intellectual Capital*, Vol. 10 No. 2, pp. 294-306.

Abhayawansa, S. and Guthrie, J. (2012), "Intellectual capital information and stock recommendations: impression management?", *Journal of Intellectual Capital*, Vol. 13 No. 3, pp. 398-415.

An, Y., Davey, H. and Eggleton, I.R.C. (2011), "Towards a comprehensive theoretical framework for voluntary IC disclosure", *Journal of Intellectual Capital*, Vol. 12 No. 4, pp. 571-585.

April, K.A., Bosma, P. and Deglon, D.A. (2003), "IC measurement and reporting: establishing a practice in SA mining", *Journal of Intellectual Capital*, Vol. 4 No. 2, pp. 165-180.

Archel, P., Husillos, J., Larrinaga, C. and Spence, C. (2009), "Social disclosure, legitimacy theory and the role of the state", *Accounting, Auditing & Accountability Journal*, Vol. 22 No. 8, pp. 1284-1307.

Ax, C. and Marton, J. (2008), "Human capital disclosures and management practices", *Journal of Intellectual Capital*, Vol. 9 No. 3, pp. 433-455.

Beattie, V. and Thomson, S.J. (2007), "Lifting the lid on the use of content analysis to investigate intellectual capital disclosures", *Accounting Forum*, Vol. 31 No. 2, pp. 129-163.

Beattie, V., McInnes, B. and Fearnley, S. (2004), "A methodology for analysing and evaluating narratives in annual reports: a comprehensive descriptive profile and metrics for disclosure quality attributes", *Accounting Forum*, Vol. 28 No. 3, pp. 205-236.

Bezhani, I. (2010), "Intellectual capital reporting at UK universities", *Journal of Intellectual Capital*, Vol. 11 No. 2, pp. 179-207.

Boedker, C., Guthrie, J. and Cuganesan, S. (2004), "The strategic significance of human capital information in annual reporting", *Journal of Human Resource Costing and Accounting*, Vol. 8 No. 2, pp. 7-22.

Boekestein, B. (2006), "The relation between intellectual capital and intangible assets of pharmaceutical companies", *Journal of Intellectual Capital*, Vol. 7 No. 2, pp. 241-253.

Bontis, N. (2003), "Intellectual capital disclosures in Canadian corporations", *Journal of Human Resource Costing and Accounting*, Vol. 7 No. 1, pp. 9-20.

Bozzolan, S., Favotto, F. and Ricceri, F. (2003), "Italian annual intellectual capital disclosure: an empirical analysis", *Journal of Intellectual Capital*, Vol. 4 No. 4, pp. 543-558.

Bozzolan, S., O'Regan, P. and Ricceri, F. (2006), "Intellectual capital disclosure (ICD): a comparison of Italy and the UK", *Journal of Human Resource Costing & Accounting*, Vol. 10 No. 2, pp. 92-113.

Branco, M.C., Delgado, C., Sá, M. and Sousa, C. (2010), "An analysis of intellectual capital disclosure by portuguese companies", *EuroMed Journal of Business*, Vol. 5 No. 3, pp. 258-278.

Branswijck, D. and Everaert, P. (2012), "Intellectual capital disclosure commitment: myth or reality?", *Journal of Intellectual Capital*, Vol. 13 No. 1, pp. 39-56.

Brennan, N. (2001), "Reporting intellectual capital in annual reports: evidence from Ireland", *Accounting, Auditing & Accountability Journal*, Vol. 14 No. 4, pp. 423-436.

Brüggen, A., Vergauwen, P. and Dao, M. (2009), "Determinants of intellectual capital disclosure: evidence from Australia", *Management Decision*, Vol. 47 No. 2, pp. 233-245.

Bukh, P.N., Nielsen, C., Gormsen, P. and Mouritsen, J. (2005), "Disclosure of information on intellectual capital in Danish IPO prospectuses", *Accounting, Auditing & Accountability Journal*, Vol. 18 No. 6, pp. 713-732.

Campbell, D. and Abdul Rahman, M.R. (2010), "A longitudinal examination of intellectual capital reporting in Marks & Spencer annual reports, 1978-2008", *The British Accounting Review*, Vol. 42 No. 1, pp. 56-70.

Carrington, D. and Tayles, M. (2012), "Intellectual capital in the Caribbean hospitality industry: two case studies", *Electronic Journal of Knowledge Management*, Vol. 10 No. 3, pp. 244-257.

Cerbioni, F. and Parbonetti, A. (2007), "Exploring the effects of corporate governance on intellectual capital disclosure: an analysis of European biotechnology companies", *European Accounting Review*, Vol. 16 No. 4, pp. 791-826.

Chang, L. and Birkett, B. (2004), "Managing intellectual capital in a professional service firm: exploring the creativity-productivity paradox", *Management Accounting Research*, Vol. 15 No. 1, pp. 7-31.

Chen, S.-Y. (2009), "Identifying and prioritizing critical intellectual capital for e-learning companies", *European Business Review*, Vol. 21 No. 5, pp. 438-452.

Cinquini, L., Passetti, E., Tenucci, A. and Frey, M. (2012), "Analyzing intellectual capital information in sustainability reports: some empirical evidence", *Journal of Intellectual Capital*, Vol. 13 No. 4, pp. 531-561.

Cordazzo, M. (2007), "Intangibles and Italian IPO prospectuses: a disclosure analysis", *Journal of Intellectual Capital*, Vol. 8 No. 2, pp. 288-305.

Cordazzo, M. and Vergauwen, P.G. (2012), "Intellectual capital disclosure in the UK biotechnology IPO prospectuses", *Journal of Human Resource Costing & Accounting*, Vol. 16 No. 1, pp. 4-19.

Dammak, S., Triki, M. and Boujelbene, Y. (2008), "A study on intellectual capital disclosure determinants in the European context", *International Journal of Learning and Intellectual Capital*, Vol. 5 No. 3, pp. 417-430.

de Villiers, C. and Dumay, J. (2013), "Construction of research articles in the leading interdisciplinary accounting journals", *Accounting, Auditing & Accountability Journal*, Vol. 26 No. 6, pp. 876-910.

Dey, C. and Burns, J. (2010), "Integrated reporting at Novo Nordisk", in Hopwood, A., Unerman, J. and Fries, J. (Eds), *Accounting for Sustainability: Practical Insights*, Earthscan, London, pp. 215-232.

Dumay, J. (2014), "15 years of the journal of intellectual capital and counting: a manifesto for transformational IC research", *Journal of Intellectual Capital*, Vol. 15 No. 1, pp. 2-37.

Dumay, J. and Cai, L. (2014), "A review and critique of content analysis as a methodology for inquiring into IC disclosure", *Journal of Intellectual Capital*, Vol. 15 No. 2, pp. 264-290.

Dumay, J. and Garanina, T. (2013), "Intellectual capital research: a critical examination of the third stage", *Journal of Intellectual Capital*, Vol. 14 No. 1, pp. 10-25.

Dumay, J.C. and Lu, J. (2010), "Disclosing improvements in human capital: comparing results to the rhetoric", *Journal of Human Resource Costing & Accounting*, Vol. 14 No. 1, pp. 70-97.

Dumay, J. and Tull, J. (2007), "Intellectual capital disclosure and price sensitive Australian stock exchange announcements", *Journal of Intellectual Capital*, Vol. 8 No. 2, pp. 236-255.

Ensslin, S.R. and De Carvalho, F.N. (2007), "Voluntary disclosure of intellectual capital in the Brazilian context: an investigation informed by the international context", *International Journal of Accounting, Auditing and Performance Evaluation*, Vol. 4 No. 4, pp. 478-500.

Evans, S. (2013), "A thousand words?", *Accounting, Auditing & Accountability Journal*, Vol. 26 No. 4, pp. 650-651.

Ferreira, A.L., Branco, M.C. and Moreira, J.A. (2012), "Factors influencing intellectual capital disclosure by Portuguese companies", *International Journal of Accounting and Financial Reporting*, Vol. 2 No. 2, pp. 278-298.

Gan, K., Saleh, Z., Abessi, M. and Huang, C.C. (2013), "Intellectual capital disclosure in the context of corporate governance", *International Journal of Learning and Intellectual Capital*, Vol. 10 No. 1, pp. 52-70.

García-Meca, E. (2005), "Bridging the gap between disclosure and use of intellectual capital information", *Journal of Intellectual Capital*, Vol. 6 No. 3, pp. 427-440.

García-Meca, E. and Martínez, I. (2007), "The use of intellectual capital information in investment decisions: an empirical study using analyst reports", *The International Journal of Accounting*, Vol. 42 No. 1, pp. 57-81.

García-Meca, E., Parra, I., Larrán, M. and Martínez, I. (2005), "The explanatory factors of intellectual capital disclosure to financial analysts", *European Accounting Review*, Vol. 14 No. 1, pp. 63-94.

Gerpott, T., Thomas, S. and Hoffman, A. (2008), "Intangible asset disclosure in the telecommunications industry", *Journal of Intellectual Capital*, Vol. 9 No. 1, pp. 37-61.

Guthrie, J. and Abeysekera, I. (2006), "Content analysis of social, environmental reporting: what is new?", *Journal of Human Resource Costing & Accounting*, Vol. 10 No. 2, pp. 114-126.

Guthrie, J. and Petty, R. (2000), "Intellectual capital: Australian annual reporting practices", *Journal of Intellectual Capital*, Vol. 1 No. 3, pp. 241-251.

Guthrie, J., Parker, L. and Gray, R. (2004a), "Requirements and understandings for publishing academic research: an insider view", in Humphrey, C. and Lee, W. (Eds), *The Real Life Guide to Accounting Research: A Behind-the-Scenes View of Using Qualitative Research Methods*, Elsevier, Amsterdam, pp. 411-432.

Guthrie, J., Petty, R., Ferrier, F. and Wells, R. (1999), "There is no accounting for intellectual capital in Australia: a review of annual reporting practices and the internal measurement of intangibles within Australian organisations", paper presented at the International Symposium on Measuring and Reporting of Intellectual Capital: Experiences, Issues and Prospects, Amsterdam, June.

Guthrie, J., Petty, R. and Ricceri, F. (2006), "The voluntary reporting of intellectual capital: comparing evidence from Hong Kong and Australia", *Journal of Intellectual Capital*, Vol. 7 No. 2, pp. 254-271.

Guthrie, J., Ricceri, F. and Dumay, J. (2012), "Reflections and projections: a decade of intellectual capital accounting research", *British Accounting Review*, Vol. 44 No. 2, pp. 68-92.

Guthrie, J., Petty, R., Yongvanich, K. and Ricceri, F. (2004b), "Using content analysis as a research method to inquire into intellectual capital reporting", *Journal of Intellectual Capital*, Vol. 5 No. 2, pp. 282-293.

Haji, A.A. and Ghazali, N.A.M. (2012), "Intellectual capital disclosure trends: some Malaysian evidence", *Journal of Intellectual Capital*, Vol. 13 No. 3, pp. 377-397.

Husin, N.M., Ahmad, N. and Sapingi, R. (2013), "Intellectual capital: a focus on human capital reporting practices of top Malaysian listed companies", *The South East Asian Journal of Management*, Vol. 5 No. 1, pp. 51-72.

Husin, N.M., Hooper, K. and Olesen, K. (2012), "Analysis of intellectual capital disclosure: an illustrative example", *Journal of Intellectual Capital*, Vol. 13 No. 2, pp. 196-220.

James, L.R., Demaree, R.G. and Wolf, G. (1984), "Estimating within-group interrater reliability with and without response bias", *Journal of Applied Psychology*, Vol. 69 No. 1, pp. 85-98.

Janis, I. (2008), "The problem of validating content analysis", in Krippendorff, K. and Bock, M.A. (Eds), *The Content Analysis Reader*, Sage, Los Angeles, CA, pp. 358-366.

Jindal, S. and Kumar, M. (2012), "The determinants of HC disclosures of Indian firms", *Journal of Intellectual Capital*, Vol. 13 No. 2, pp. 221-247.

Johanson, U., Martensson, M. and Skoog, M. (2001), "Mobilizing change through the management control of intangibles", *Accounting, Organizations and Society*, Vol. 26 Nos 7/8, pp. 715-733.

Johanson, U., Koga, C., Skoog, M. and Henningsson, J. (2006), "The Japanese government's intellectual capital reporting guideline: what are the challenges for firms and capital market agents?", *Journal of Intellectual Capital*, Vol. 7 No. 4, pp. 474-491.

Joshi, M. (2012), "Intellectual capital disclosures by Indian and Australian information technology companies: a comparative analysis", *Journal of Intellectual Capital*, Vol. 13 No. 4, pp. 582-598.

Kenkel, J.L. (1989), *Introductory Statistics for Management and Economics*, PWS-KENT Publishing Company, Boston, MA.

Khan, H.-U.-Z. and Ali, M. (2010), "An empirical investigation and users' perceptions on intellectual capital reporting in banks: evidence from Bangladesh", *Journal of Human Resource Costing & Accounting*, Vol. 14 No. 1, pp. 48-69.

Krippendorff, K. (2008), "Testing the reliability of content analysis data", in Krippendorff, K. and Bock, M.A. (Eds), *The Content Analysis Reader*, Sage, Los Angeles, CA, pp. 350-357.

Krippendorff, K. (2013), *Content Analysis: An Introduction to its Methodology*, Sage, Los Angeles, CA.

Kristandl, G. and Bontis, N. (2007), "The impact of voluntary disclosure on cost of equity capital estimates in a temporal setting", *Journal of Intellectual Capital*, Vol. 8 No. 4, pp. 577-594.

LeBreton, J.M., James, L.R. and Lindell, M.K. (2005), "Recent issues regarding rWG, rWG, rWG(J), and rWG(J)", *Organizational Research Methods*, Vol. 8 No. 1, pp. 128-138.

Lee, A., Neilson, J., Tower, G. and Van der Zahn, J.W.M. (2007), "Is communicating intellectual capital information via the internet viable?: case of Australian private and public hospitals", *Journal of Human Resource Costing & Accounting*, Vol. 11 No. 1, pp. 53-78.

Li, J., Pike, R. and Haniffa, R. (2008), "Intellectual capital disclosure and corporate governance structure in UK firms", *Accounting and Business Research*, Vol. 38 No. 2, pp. 137-159.

Liao, L., Low, M. and Davey, H. (2013), "Chinese and English language versions: intellectual capital disclosure", *Journal of Intellectual Capital*, Vol. 14 No. 4, pp. 661-686.

Lock Lee, L. and Guthrie, J. (2010), "Visualising and measuring intellectual capital in capital markets: a research method", *Journal of Intellectual Capital*, Vol. 11 No. 1, pp. 4-22.

Mention, A.-L. (2011), "Exploring voluntary reporting of intellectual capital in the banking sector", *Journal of Management Control*, Vol. 22 No. 3, pp. 279-309.

MERITUM (2002), *Guidelines for Managing and Reporting on Intangibles (Intellectual Capital Statements)*, Vodafone Foundation, Madrid.

Murthy, V. and Abeysekera, I. (2007), "Human capital value creation practices of software and service exporter firms in India", *Journal of Human Resource Costing & Accounting*, Vol. 11 No. 2, pp. 84-103.

Neuendorf, K.A. (2002), *The Content Analysis Guidebook*, Sage, Thousand Oaks, CA.

Nurunnabi, M., Hossain, M. and Hossain, M. (2011), "Intellectual capital reporting in a South Asian country: evidence from Bangladesh", *Journal of Human Resource Costing & Accounting*, Vol. 15 No. 3, pp. 196-233.

Oliveira, L., Rodrigues, L.L. and Craig, R. (2006), "Firm-specific determinants of intangibles reporting: evidence from the Portuguese stock market", *Journal of Human Resource Costing & Accounting*, Vol. 10 No. 1, pp. 11-33.

Oliveras, E., Gowthorpe, C., Kasperskaya, Y. and Perramon, J. (2008), "Reporting intellectual capital in Spain", *Corporate Communications: An International Journal*, Vol. 13 No. 2, pp. 168-181.

Olsson, B. (2001), "Annual reporting practices: information about human resources in corporate annual reports in major Swedish companies", *Journal of Human Resource Costing and Accounting*, Vol. 6 No. 1, pp. 39-52.

Olsson, B. (2004), "Intellectual capital disclosure through annual reports: a study of the Swedish retail industry", *Journal of Human Resource Costing & Accounting*, Vol. 8 No. 2, pp. 57-72.

Orens, R., Aerts, W. and Lybaert, N. (2009), "Intellectual capital disclosure, cost of finance and firm value", *Management Decision*, Vol. 47 No. 10, pp. 1536-1554.

Otley, D. (2003), "Management control and performance management: whence and whither?", *The British Accounting Review*, Vol. 35 No. 4, pp. 309-326.

Parker, L.D. and Roffey, B.H. (1997), "Methodological themes: back to the drawing board: revisiting grounded theory and the everyday accountant's and manager's reality", *Accounting, Auditing & Accountability Journal*, Vol. 10 No. 2, pp. 212-247.

Pedrini, M. (2007), "Human capital convergences in intellectual capital and sustainability reports", *Journal of Intellectual Capital*, Vol. 8 No. 2, pp. 346-366.

Petty, R. and Cuganesan, S. (2005), "Voluntary disclosure of intellectual capital by Hong Kong companies: examining size, industry and growth effects over time", *Australian Accounting Review*, Vol. 15 No. 36, pp. 40-50.

Petty, R. and Guthrie, J. (2000), "Intellectual capital literature review: measurement, reporting and management", *Journal of Intellectual Capital*, Vol. 1 No. 2, pp. 155-176.

Rashid, A.A., Ibrahim, M.K., Othman, R. and See, K.F. (2012), "IC disclosures in IPO prospectuses: evidence from Malaysia", *Journal of Intellectual Capital*, Vol. 13 No. 1, pp. 57-80.

Rimmel, G., Nielsen, C. and Yosano, T. (2009), "Intellectual capital disclosures in Japanese IPO prospectuses", *Journal of Human Resource Costing & Accounting*, Vol. 13 No. 4, pp. 316-337.

Schneider, A. and Samkin, G. (2008), "Intellectual capital reporting by the New Zealand local government sector", *Journal of Intellectual Capital*, Vol. 9 No. 3, pp. 456-486.

Shareef, F. and Davey, H. (2005), "Accounting for intellectual capital: evidence from listed English football clubs", *Journal of Applied Accounting Research*, Vol. 7 No. 3, pp. 78-116.

Sihotang, P. and Winata, A. (2008), "The intellectual capital disclosures of technology-driven companies: evidence from Indonesia", *International Journal of Learning and Intellectual Capital*, Vol. 5 No. 1, pp. 63-82.

Singh, S. and Kansal, M. (2011), "Voluntary disclosures of intellectual capital: an empirical analysis", *Journal of Intellectual Capital*, Vol. 12 No. 2, pp. 301-318.

Steenkamp, N. and Northcott, D. (2007), "Content analysis in accounting research: the practical challenges", *Australian Accounting Review*, Vol. 17 No. 43, pp. 12-25.

Striukova, L., Unerman, J. and Guthrie, J. (2008), "Corporate reporting of intellectual capital: evidence from UK companies", *The British Accounting Review*, Vol. 40 No. 4, pp. 297-313.

Sujan, A. and Abeysekera, I. (2007), "Intellectual capital reporting practices of the top Australian firms", *Australian Accounting Review*, Vol. 17 No. 2, pp. 71-83.

Sveiby, K.E. (1997), *The New Organisational WEALTH: Managing and Measuring Knowledge Based Assets*, Berrett Koehler, SanFrancisco, CA.

Unerman, J. (2000), "Methodological issues – reflections on quantification in corporate social reporting content analysis", *Accounting, Auditing & Accountability Journal*, Vol. 13 No. 5, pp. 667-681.

Uyar, A. and Kiliç, M. (2012), "Value relevance of voluntary disclosure: evidence from Turkish firms", *Journal of Intellectual Capital*, Vol. 13 No. 3, pp. 363-376.

Vafaei, A., Taylor, D. and Ahmed, K. (2011), "The value relevance of intellectual capital disclosures", *Journal of Intellectual Capital*, Vol. 12 No. 3, pp. 407-429.

Vandemaele, S.N., Vergauwen, P.G.M.C. and Smits, A.J. (2005), "Intellectual capital disclosure in the Netherlands, Sweden and the UK: a longitudinal and comparative study", *Journal of Intellectual Capital*, Vol. 6 No. 3, pp. 417-426.

Weber, R.P. (1990), *Basic Content Analysis*, Sage, London.

Whiting, R.H. and Miller, J.C. (2008), "Voluntary disclosure of intellectual capital in New Zealand annual reports and the 'hidden value'", *Journal of Human Resource Costing & Accounting*, Vol. 12 No. 1, pp. 26-50.

Whiting, R.H. and Woodcock, J. (2011), "Firm characteristics and intellectual capital disclosure by Australian companies", *Journal of Human Resource Costing & Accounting*, Vol. 15 No. 2, pp. 102-126.

Yau, F.S., Chun, L.S. and Balaraman, R. (2009), "Intellectual capital reporting and corporate characteristics of public-listed companies in Malaysia", *Journal of Financial Reporting and Accounting*, Vol. 7 No. 1, pp. 17-35.

Yi, A. and Davey, H. (2010), "Intellectual capital disclosure in Chinese (mainland) companies", *Journal of Intellectual Capital*, Vol. 11 No. 3, pp. 326-347.

Further reading

Murthy, V. and Mouritsen, J. (2011), "The performance of intellectual capital: mobilising relationships between intellectual and financial capital in a bank", *Accounting, Auditing and Accountability Journal*, Vol. 24 No. 5, pp. 622-646.

Corresponding author

John Dumay can be contacted at: john.dumay@mq.edu.au

Appendix 1

Article	Research question
Brennan (2001, p. 427)	What is the extent of IC reporting?[a]
Olsson (2004, p. 60)	What is the difference of ICD between different years/different type of organisation/different countries?[a]
Sujan and Abeysekera (2007, p. 74)	Can the media agenda-setting theory explain the state of IC reporting in relation to reporting differences among industry sectors (classified according to the Global Industry Classification Standard), and reporting differences between the knowledge-based and service sector and other sectors? How do the results of the current study compare with those of the previous studies and can the media agenda-setting theory explain the comparison?
Khan and Ali (2010, p. 49)	What are users' opinions with regards to Bangladeshi banks' IC reporting initiatives?
Dumay and Lu (2010, p. 74)	Is the rhetoric of HC disclosure achieved in practice?
Sihotang and Winata (2008, p. 66)	What is the relationship between the number of the annual reports' pages and the level of IC disclosures?
Carrington and Tayles (2012, p. 246)	Do hotel managers use intellectual capital information to assist them in their operational decisions relating to staffing, customer and supplier relationships?
April *et al.* (2003, p. 166)	Is intellectual capital at all relevant or meaningful to South African mining companies and, if so, what are they doing in this regard?
Vafaei *et al.* (2011, p. 409)	Whether the extent of textual information in annual reports on intellectual capital, aggregated across the human, structural, relational and general categories, has value-relevance in share markets? Whether ICD moderates the incremental value-relevance of reported IFRS adjustments to earnings (i.e. net profit or loss) and equity (i.e. book value of net assets)?
Dumay and Tull (2007, pp. 240-241)	Is there any effect on the CAR (cumulative abnormal return) of a stock when an announcement is classified as "price sensitive" by the ASX? Is there any effect on the CAR of a stock when an announcement is classified as "price sensitive" by the ASX and when that announcement may be classified as "good", "bad" or "neutral news"? Is there any effect on the CAR of a stock when an announcement is classified as "price sensitive" by the ASX and when such an announcement is aligned?
Nurunnabi *et al.* (2011, p. 202)	How do we develop a disclosure index that is applicable to developing countries?
Sihotang and Winata (2008, pp. 69-71)	To what extent do company characteristics such as the industry category, age and size have a relationship with the level of IC disclosures?
García-Meca (2005, p. 429)	This essay is to examine the information related to intellectual capital reported to analysts, comparing if this data is finally considered by financial analysts in their decision-making process?
Brennan (2001, p. 427)	The paper tests whether in Ireland book values and market values are materially different
Boekestein (2006, p. 244)	How visible are intangible assets on the balance sheet?

Note: [a]These are generic questions summarised from the originals

Table AI.
Research questions

Appendix 2

Table AII.
ICD quality measures

Three measures	0	1	2	3	4	5	Weighting formulae specified
Oliveira *et al.* (2006, p. 20)	No	Qualitative	Quantitative				Yes
Bozzolan *et al.* (2006, p. 99)	No	Qualitative	Quantitative				No
Whiting and Miller (2008, pp. 35-36)	No	Qualitative	Quantitative				No
Whiting and Woodcock (2011, p. 113)	No	Qualitative	Quantitative				Yes
Bozzolan *et al.* (2003, p. 549)	No	Qualitative	Quantitative				No
Vandemaele *et al.* (2005, p. 420)	No	Qualitative	Quantitative				No
Cerbioni and Parbonetti (2007, p. 810)	No	Yes	Quantitative				No
Yi and Davey (2010, p. 335)	No	Immaterial	Obscure				No No
Sujan and Abeysekera (2007, p. 75)		Narrative	Numerical	Monetary			No
Branswijck and Everaert (2012, p. 43)		General	Qualitative	Quantitative			Yes
Mention (2011, pp. 288-289)		Fact, qualitative, backward oriented	Perception, quantitative, forward-looking				No
Abeysekera (2008, p. 728)			Quantitative	Numerical	Monetary		No
Four measures							
Petty and Cuganesan (2005, p. 44)	No	Discursive	Numerical	Monetary			No
Guthrie *et al.* (2006, p. 259)	No	Discursive	Numerical	Monetary			No
Guthrie *et al.* (1999, p. 28)	No	Discursive	Numerical	Monetary			No
Haji and Ghazali (2012, p. 385)	No	Narrative	Numerical	Monetary			No
Yau *et al.* (2009, pp. 22-23)	No	Narrative	Numerical	Monetary			No
Gan *et al.* (2013, p. 60)	No	Narrative	Numerical	Monetary			Yes
Bezhani (2010, p. 185)	No	Discursive	Numerical	quantified			No
Nurunnabi *et al.* (2011, p. 201)	No	Qualitative	Quantitative	Qualitative and quantitative			Yes

(*continued*)

Table AII.

Three measures	0	1	2	3	4	5	Weighting formulae specified
Gerpott *et al.* (2008, p. 42)	No	Yes	Qualitative or quantitative	Qualitative and quantitative			No
Ensslin and De Carvalho (2007, p. 485)	No	Not narrative	Not numerical	Monetary			No
Husin *et al.* (2012, p. 241)		Obscure	Descriptive	Strongly descriptive			Yes
Kristandl and Bontis (2007, p. 581)		Qualitative, unspecified		Quantitative, specific			Yes
Five weights							
Husin *et al.* (2013, p. 64)	No	Obscure	Descriptive	Quantitative or monetary	Quantitative, monetary, visual image or descriptive		No
Liao *et al.* (2013, p. 664)	No	Narrative	Numerical	Monetary	Both qualitative and quantitative		No
Singh and Kansal (2011, p. 306)	No	Narrative	Quantitative	Monetary	formula-based comparative		Yes
Six weights							
Schneider and Samkin (2008, p. 407)	No	Immaterial	Obscure	Descriptive	Quantitative or monetary	Quantitative, monetary and descriptive	No
Shareef and Davey (2005, p. 91)	No	Immaterial	Obscure	Descriptive	Quantitative or monetary	Quantitative, monetary and descriptive	No

This article has been cited by:

1. Qianyu Wang Department of Accounting, Waikato Management School, University of Waikato, Hamilton, New Zealand Umesh Sharma Department of Accounting, Waikato Management School, University of Waikato, Hamilton, New Zealand Howard Davey Department of Accounting, Waikato Management School, University of Waikato, Hamilton, New Zealand . 2016. Intellectual capital disclosure by Chinese and Indian information technology companies. *Journal of Intellectual Capital* **17**:3, 507-529. [Abstract] [Full Text] [PDF]
2. Maurizio Massaro Department of Economic Sciences and Statistics, University of Udine, Udine, Italy John Dumay Department of Accounting and Corporate Governance, Macquarie University, Sydney, Australia James Guthrie Department of Accounting and Corporate Governance, Macquarie University, Sydney, Australia . 2016. On the shoulders of giants: undertaking a structured literature review in accounting. *Accounting, Auditing & Accountability Journal* **29**:5, 767-801. [Abstract] [Full Text] [PDF]
3. Doni Federica, Gasperini Andrea, Pavone Pasquale. 2016. Early adopters of integrated reporting: The case of the mining industry in South Africa. *African Journal of Business Management* **10**:9, 187-208. [CrossRef]
4. Richard Slack Durham University Business School, Durham University, Stockton on Tees, UK Matthias Munz Newcastle Business School, Northumbria University, Newcastle upon Tyne, UK . 2016. Intellectual capital reporting, leadership and strategic change. *Journal of Applied Accounting Research* **17**:1, 61-83. [Abstract] [Full Text] [PDF]
5. Dr Stefano Zambon Pierluigi Catalfo Economia e Impresa, University of Catania, Catania, Italy Inge Wulf Institute of Management and Economics, Clausthal University of Technology, Clausthal-Zellerfeld, Germany . 2016. Intangibles disclosure in Management Commentary regulation in Germany and Italy. *Journal of Intellectual Capital* **17**:1, 103-119. [Abstract] [Full Text] [PDF]
6. Mary Low Department of Accounting, University of Waikato, Hamilton, New Zealand Grant Samkin Department of Accounting, University of Waikato, Hamilton, New Zealand AND Department of Financial Accounting, University of South Africa, Pretoria, South Africa Yuanyuan Li Department of Accounting, University of Waikato, Hamilton, New Zealand . 2015. Voluntary reporting of intellectual capital. *Journal of Intellectual Capital* **16**:4, 779-808. [Abstract] [Full Text] [PDF]
7. Gaia Melloni Department of Accounting, Bocconi University, Milano , Italy . 2015. Intellectual capital disclosure in integrated reporting: an impression management analysis. *Journal of Intellectual Capital* **16**:3, 661-680. [Abstract] [Full Text] [PDF]
8. Professor James Guthrie and Associate Professor John Dumay John Dumay Department of Accounting and Corporate Governance, Macquarie University, Sydney, Australia James Guthrie Department of Accounting and Corporate Governance, Macquarie University, Sydney, Australia AND Institute of Advanced Studies, University of Bologna, Bologna, Italy Pina Puntillo Department of Management and Law, University of Calabria, Cosenza, Italy . 2015. IC and public sector: a structured literature review. *Journal of Intellectual Capital* **16**:2, 267-284. [Abstract] [Full Text] [PDF]