
Testing vs. building accounting theory with experimental research: insights from management research

Daniel Reimsbach* and Bastian Hauschild

Business and Economics Faculty,
University of Duesseldorf,
Universitätsstr. 1 (Bld. 24.31, Rm. 02.04),
40225 Düsseldorf, Germany
Fax: +49 (0)211 81-15157
Email: daniel.reimsbach@hhu.de
*Corresponding author

Abstract: With this conceptual paper, we want to raise scholars' awareness of the underused potential of experiments in accounting research. We refer to research taxonomies from the management discipline to explore the role of experimental research in testing and building accounting theory. Although the contribution of laboratory research to testing theory is well acknowledged, we argue that exploratory experimental research efforts also have the potential to build new theory, pointing at the underused potential of experiments in current accounting research. In doing so, we (i) provide a clear understanding of theory building and theory testing in accounting research; (ii) present a guideline for assessing the degree of theory building and theory testing inherent in any (experimental) study. Furthermore, we argue that an inter-study triangulation aligning experimental and archival methodologies will have the potential to produce results that are robust to internal and external validity concerns, thus advancing accounting theory.

Keywords: accounting theory; experimental research; theory building; theory testing; research taxonomies; pro forma earnings.

Reference to this paper should be made as follows: Reimsbach, D. and Hauschild, B. (2015) 'Testing vs. building accounting theory with experimental research: insights from management research', *Int. J. Behavioural Accounting and Finance*, Vol. 5, No. 1, pp.82–98.

Biographical notes: Daniel Reimsbach is an Assistant Professor of Accounting at the University of Duesseldorf, Germany.

Bastian Hauschild is a former Senior Research Fellow at the Finance Department of the University of Duesseldorf, Germany. Currently, he works for Bayer Business Consulting, Germany.

This paper is a revised and expanded version of a paper entitled 'Testing vs. building accounting theory with experimental research' presented at the *Annual Meeting 2014*, Southwestern Accounting Association, Dallas, TX, 12–15 March 2014.

1 Introduction

In recent years, experiments have become very popular with accounting scholars, and editors of top-tier accounting journals still call for further experimental research (most recently, see Moser and Martin, 2012). The use of experiments is often motivated by methodological problems of archival research (e.g., Hirst et al., 1995, p.336; Highhouse, 2009; Koonce et al., 2011a), especially in proving causal relationships among variables of interest. Laboratory experiments have the potential to overcome these methodological problems: such experiments involve the use of simulated treatments in an artificial environment, thus allowing researchers to test assumed causal relationships under controlled conditions (also see Libby et al., 2002; Maines et al., 2006; Falk and Heckman, 2009). Numerous authors thus highlight the contribution of experiments to *testing* accounting theory (e.g., Callahan et al., 2006). However, given that theory is the ‘basic aim of science’ (Kerlinger and Lee, 2000, p.11), advancing the accounting discipline requires not only testing existing theories but also *building* new theories, that is, introducing new relationships and constructs that can serve as the foundation for new theory (Colquitt and Zapata-Phelan, 2007). Against this background, the aim of this paper is twofold. First, we investigate whether experimental research in accounting also contributes to building accounting theory or whether it is (yet) limited to testing theory. Second, given that archival-empirical studies still outnumber experimental approaches by far (see Koonce and Mercer, 2005), we further address if there is an underused potential of experiments in current accounting research. To answer these questions, this paper explores the role of experimental research in testing *and* building accounting theory.

Some authors have alleged that accounting researchers have created “a highly advanced research context which, in effect, is an environment dominated by sophisticated methodology, rather than theory” (Inanga and Schneider, 2005, p.227). However, such allegations need a thorough discourse. Although an active debate on testing and building theory exists in disciplines such as management (e.g., Sutton and Staw, 1995; Colquitt and Zapata-Phelan, 2007; Corley and Gioia, 2011), corresponding work in the accounting discipline remains anecdotal (e.g., Peecher and Solomon, 2001) and focuses on practical advice on how to avoid research traps (Libby et al., 2002). We address this shortcoming

- by providing a clear understanding and definition of theory building and theory testing in accounting research
- by presenting a guideline for assessing the *degree* of theory building and theory testing inherent in any (experimental) study.

In doing so, we refer to Colquitt and Zapata-Phelan’s (2007) taxonomy which originally focuses on research in the management discipline. However, the generic nature of Colquitt and Zapata-Phelan’s (2007) approach allows the degree of a study’s contribution to theory building and theory testing in any other disciplines to be identified. Consequently, we transfer their taxonomy to the accounting discipline. While it can be applied to any empirical accounting research, we focus on experimental approaches. For the sake of brevity, we refer to two topical streams of research in financial accounting, that is, research on pro forma earnings disclosure and environmental, social and governance (ESG) reporting, as they represent a research stream that recently started using experiments (e.g., Frederickson and Miller, 2004; Dilla et al., 2013). We demonstrate that the experimental studies discussed in this paper offer a moderate level

of theory building and theory testing. However, we argue that this is not a consequence of the experimental method in general but rather that it is possible (and reasonable) to enhance the degree of theory building and testing when conducting experimental research. Overall, we conclude that assessing the degree of building and testing theory should be directed at a *stream* of research rather than at a single research study. From this perspective, it seems desirable that a stream of research as a whole builds and tests theory, thus advancing the accounting discipline. Consequently, we advocate the concept of *triangulation of studies* (Hageman, 2008), that is, aligning different *separate* studies on the same topic that use different research methods such as experimental and archival approaches. Within this triangulation however, the role of experimental research is not predetermined and might encompass confirmatory theory-testing approaches as well as more exploratory theory-building approaches. The latter is well acknowledged in other disciplines such as psychology (see Kahneman, 2012) and natural sciences (e.g., Franklin, 2005). We feel that increasing scholars' awareness of these aspects will contribute to pushing the limits of behavioural accounting theory.

The remainder of this paper is organised as follows. In Section 2, we present our understanding of theory building and theory testing. In Section 3, we outline how research on pro forma earnings disclosure and ESG reporting relies on experiments. Furthermore, we use Colquitt and Zapata-Phelan's (2007) taxonomy to reflect on the extent to which experimental studies contribute to testing and/or building accounting theory. In Section 4, we discuss strategies that help realise the full potential of experiments in building and testing accounting theory. Finally, we present a brief conclusion.

2 Experimental research in accounting

2.1 Potential value of accounting experiments

An important aspect in the broad field of accounting research is financial communication between managers and investors, auditors and other information intermediaries. Much of the respective literature focuses on managements' reporting decisions. This focus on judgement and decision-making clearly suggests the use of experimental research methods (also see Libby et al., 2002) especially given the fact that the behaviour of users of financial accounting information often deviates from normative behaviour. However, researchers using archival research methods clearly still outnumber experimental researchers (e.g., Koonce and Mercer, 2005). Dominant archival-empirical research focuses on market prices as a presumed result of managers' reporting decisions. One of the most important challenges associated with this approach is that it does not allow testing causal relationships (Hageman, 2008). Researchers analyse the statistical relationship between certain variables and further add control variables that may have affected the relationship. Still, as possible alternative explanations cannot be ruled out, archival approaches have a relatively low level of internal validity (Aguinis and Bradley, 2014).

In this aspect, experimental research methods are clearly different from the archival approach. The most important advantage of well-designed experiments is their ability to establish causality in relationships between and among phenomena under interest (Libby et al., 2002). Laboratory conditions allow isolation and controlling for other

potentially confounding factors (Aguinis and Bradley, 2014). Thus, researchers can learn about human cognitions and deviations from normative behaviour by systematically manipulating certain factors and observing participants' subsequent behaviour while controlling for alternative explanations, most importantly by randomising participants among experimental conditions (Hageman, 2008). However, high internal validity and control sometimes come at the cost of a lack of generalisability and external validity of the experimental results. Still, laboratory experiments have the primary strengths of control, which allows constructing an environment in which assumed causal phenomena can be tested with a maximum of internal validity (Kerlinger and Lee, 2000). Because of these benefits, a true experiment is considered "a scientific ideal, particularly when testing and building theory" (Hageman, 2008, p.14). However, to further analyse the role of experimental research in testing and building accounting theory, we need an understanding of what constitutes a theory

2.2 Building vs. testing accounting theory with experiments

There is little agreement among accounting academics on what constitutes a theory (Malmi and Granlund, 2009). For some accounting scholars, theory is simply all the existing literature on accounting, while for others theory it is an orienting set of ideas or explanatory concepts (Ahrens and Chapman, 2006). Furthermore, Watts and Zimmerman's (1986) work makes a clear distinction between positive and normative accounting theory. Positive theory aims at explaining empirical phenomena ('the study of what is') while normative theory makes prescriptive outlines ('what ought to be,' Oler et al., 2010). However, elaborating on an experimental study's contribution to building and testing accounting theory requires a clear definition of what actually constitutes a theory. Our view on theory follows Gioia and Pitre's (1990) definition of theory as a statement of concepts and their interrelationships that explain how and why a phenomenon occurs (also see Corley and Gioia, 2011). In this sense, theory emphasises the nature of *causal* relationships (Sutton and Staw, 1995; Koonce et al., 2011b). Our view on theory is thus akin to the view held by most positive accounting researchers.

In the following, to analyse the theoretical contribution of accounting studies, we refer to research taxonomies that represent classification systems categorising research efforts into 'mutually exclusive sets' (Doty and Glick, 1994). However, with a few notable exceptions (e.g., Peecher and Solomon, 2001), research taxonomies that focus on accounting research are rare. Peecher and Solomon (2001) build their taxonomy around two dimensions: the degree of scientific understanding (increasing from exploratory to descriptive, and finally explanatory) and the foci of research (descriptive, normative and prescriptive). We agree that differentiating exploratory and confirmatory research approaches is important when evaluating how research contributes to advancing accounting theory, and we refer to this aspect later on. However, we see three potential pitfalls in Peecher and Solomon's (2001) taxonomy for our reasoning. First, we think that the two dimensions do not represent truly exclusive sets. For example, both dimensions include the category 'descriptive'. This complicates classifying specific research efforts especially because the distinction among the underlying dimensions ('degree of scientific understanding' vs. 'foci of research') remains unclear. Second, the classification of research efforts in Peecher and Solomon's (2001) taxonomy does not consider findings of related prior research. We argue that classifying research as exploratory or confirmatory depends on prior research that has been conducted in that field. Third, and most

importantly, the degree of studies' testing and building of accounting theory is not explicitly incorporated in the taxonomy, which we consider a shortcoming in assessing a study's contribution to the accounting discipline.

Therefore, we will refer to a research taxonomy from the management discipline, where reflections on theory (e.g., Gioia and Pitre, 1990; Sutton and Staw, 1995) and the use of taxonomies (e.g., Corley and Gioia, 2011; Colquitt, 2008) are much more common compared to the accounting discipline. Specifically, we refer to Colquitt and Zapata-Phelan (2007) who develop a taxonomy that allows reflection on the theoretical contribution of empirical studies along the two dimensions of theory testing and theory building. This taxonomy was published in the top-tier *Academy of Management Journal* and has been cited numerous times since publication.¹ Thus, it can be considered very influential in the management discipline and on its current research agenda. The first dimension of Colquitt and Zapata-Phelan's (2007) taxonomy, theory building, refers to "the degree to which an empirical paper clarifies or supplements existing theory or introduces relationships and constructs that serve as the foundation for new theory" (p.1283). Studies on a rather low level of theory building predominantly replicate and thus (cross) validate findings of prior empirical work or examine effects that, while being the subject of prior theorising, have not been at the centre of prior empirical work. Introducing a new mediator or moderator of an existing relationship or process represents a moderate level of theory building. Finally, examining previously unexplored relationships or processes or even introducing a new construct is the highest level of theory building.

The second dimension that characterises a theoretical contribution is theory testing. *Theory testing* is "the degree to which existing theory is applied in an empirical study as a means of grounding a specific set of a priori hypotheses" (Colquitt and Zapata-Phelan, 2007, p.1284; also see Miller and Tsang, 2011). The extent of a study's theory testing remains at a low level if predictions or hypotheses are grounded in logical speculation or past findings that act as 'theory substitutes' (Colquitt and Zapata-Phelan, 2007). Predictions grounded in existing conceptual arguments represent a moderate level of theory testing, whereas a high level is reached if predictions refer to a (complete) model or even a theory.

Furthermore, Colquitt and Zapata-Phelan (2007) identify five distinct types of papers that differ in their overall theoretical contribution. 'Reporters' have relatively low levels of theory building and theory testing. 'Testers' are defined as papers that have a high level of theory testing but only a low level of theory building, whereas, inversely, 'builders' contain high levels of theory building but low levels of theory testing. 'Qualifiers' are papers that have a moderate degree of theory building and theory testing. Finally, 'expanders' offer the highest theoretical contribution because they contain a relatively high level of theory building and theory testing. On a more aggregated level, Colquitt and Zapata-Phelan (2007) further conclude that reporters and qualifiers make an overall low theoretical contribution while builders, testers and expanders make an overall high theoretical contribution.

Colquitt and Zapata-Phelan's (2007) taxonomy has a dynamic character insofar, as classifying research efforts explicitly depends on the direction (and findings) of prior research. Thus, this taxonomy differs from Peecher and Solomon's (2001) and provides a dynamic perspective on the theoretical contribution of accounting research studies. Next, we elaborate on how experimental research in accounting can make a theoretical contribution along the two dimensions of theory building and theory testing. To provide a

clear understanding of our reasoning, we refer to two topical streams of research in financial accounting: the disclosure of pro forma earnings and ESG reporting.

3 Theoretical contributions of experimental accounting research

3.1 The example of pro forma earnings disclosure

Analysing investor reaction to the disclosure of pro forma earnings is one example of how accounting researchers use experiments to answer questions where archival research fails. Pro forma earnings represent performance metrics not calculated in accordance with mandatory accounting standards such as US GAAP or IFRS (Marques, 2010). Modifying the bottom-line number or subtotals of the mandatory income statement (GAAP earnings), firms often voluntarily disclose pro forma earnings such as ‘operating income’, ‘recurring earnings’ or ‘adjusted EBIT(DA)’ in their earnings press releases or quarterly/annual reports (Hitz, 2010). Regarding management’s expertise and private information, on the one hand, those performance metrics might supplement GAAP earnings information by providing a more ‘sustainable’ income number that eliminates non-recurring items (Hitz, 2010). On the other hand managers might strategically use pro forma earnings disclosure to portray a better firm performance.

Findings of archival empirical research support the (additional) information (e.g., Bhattacharya et al., 2003; Lougee and Marquardt, 2004) and the strategic reporting hypotheses (Bhattacharya et al., 2004; Entwistle et al., 2006). As management’s true motivation cannot be observed though, both motives are hard to disentangle empirically (Entwistle et al., 2006). Furthermore, to stir further regulatory actions – which are currently much stricter in the USA than in Europe (Hitz, 2010) – proof is needed that investors are indeed influenced and thus potentially misled by pro forma earnings information. Whereas archival empirical research on investor reaction (e.g., Bhattacharya et al., 2007) faces the problem of disentangling the effect of pro forma earnings disclosure on investors’ decisions from other investment-related effects (Allee et al., 2007), controlled experiments might test this assumed causal relationship under pure conditions.

Four experimental studies that explicitly focus on the disclosure of pro forma earnings have been published in international accounting journals (Frederickson and Miller, 2004; Elliott, 2006; Andersson and Hellman, 2007; Dilla et al., 2013). In a seminal work, Frederickson and Miller (2004) find that non-professional investors, upon receiving an earnings press release containing non-GAAP and GAAP information, give significantly higher assessments of a fair value for common stock than other non-professional investors who received GAAP information only. Elliott’s (2006) findings confirm and extend these results, indicating that it is not the mere presence but the emphasis that management places on pro forma earnings that influence non-professional investors’ judgements of a firm’s earnings performance and the investors’ investment decisions. Furthermore, in Elliott’s (2006) experiment, this influence was mitigated by a quantitative reconciliation of pro forma to GAAP earnings. However, Dilla et al.’s (2010) results indicate that the effects of earnings reconciling information on non-professional investor judgements demonstrated by Elliott (2006) are contingent on the financial reporting knowledge level and information viewing behaviour.² Andersson and Hellman’s (2007) study is the only one indicating that

professional analysts are also influenced by the disclosure of pro forma earnings if the difference between (positive) pro forma and (negative) GAAP earnings is substantial. Most recently, Dilla et al. (2013) find that non-professional and professional investors are also influenced by graphical displays of pro forma earnings information when evaluating a company's future earnings potential.

3.2 The example of ESG reporting

A second example of how experimentation has been deployed in topical accounting research is the research stream on ESG reporting. In today's business world, companies increasingly are challenged to consider environmental and social perspectives along with financial goals. Companies react to this trend by reporting on their sustainability efforts and by increasingly disclosing information on their ESG performance (KPMG, 2013). Numerous archival-empirical studies analyse the determinants of such ESG or sustainability reporting (for an overview see Hahn and Kühnen, 2013).³ However, research on the *effects* of ESG reporting still is scarce. Investors are key addressee of ESG reporting and a few studies indicate that they consider non-financial information value relevant (e.g., Berthelot et al., 2012; Dhaliwal et al., 2012). Thus, the role of ESG information in the decision-making process of professional and nonprofessional investors is an important question for experimental research.

Seven experimental studies that explicitly focus on the effects of ESG reporting have been published in international accounting and business journals (Chan and Milne, 1999; Holm and Rikhardsson, 2008, Rikhardsson and Holm, 2008; Reimsbach and Hahn, 2015; Elliott et al., 2014; Cheng et al., 2015; Brown-Liburd and Zamora, 2015). Chan and Milne (1999) find that investors in an allocation-of-funds context react strongly and negatively to poor ESG performance, but that there is no significant reaction to disclosure of good ESG performance. Furthermore, Chan and Milne (1999) observe that long-term decision horizons increase the effect of ESG disclosures, whereas Rikhardsson and Holm's (2008) findings indicate a stronger reliance on ESG information for the short-term horizon. In addition, Rikhardsson and Holm (2008) find that adding quantitative ESG information mitigates rather than extends the directional effect of qualitative ESG information. In a second experiment, Holm and Rikhardsson (2008) again find support for the notion that ESG reporting influences investment-allocation decisions. However, they extend prior research by analysing moderating factors such as the experience level of investors which affect the observed investment allocation decisions. Reimsbach and Hahn's (2015) findings indicate that self-reporting of poor ESG performance does not negatively affect decision-makers' stock-price estimates and investment decisions. However, compared to Chan and Milne (1999), the authors explicitly use negative incidents (e.g., underage labour) to operationalise bad ESG performance in corporate reports. Elliott et al. (2014) find that investors derive higher fundamental-value estimates in response to positive ESG performance, and lower fundamental-value estimates in response to negative ESG performance. However, this effect is driven mainly by unintentional affective reactions to ESG performance. Cheng et al. (2015) and Brown-Liburd and Zamora (2015) extend prior research by analysing the effects of ESG assurance. Cheng et al. (2015) find that assurance increases investors' willingness to invest to a greater extent when ESG indicators have high relevance to the company strategy. Brown-Liburd and Zamora (2015) find that investors are sceptical of reported ESG information if managerial pay is explicitly tied to ESG performance. In this

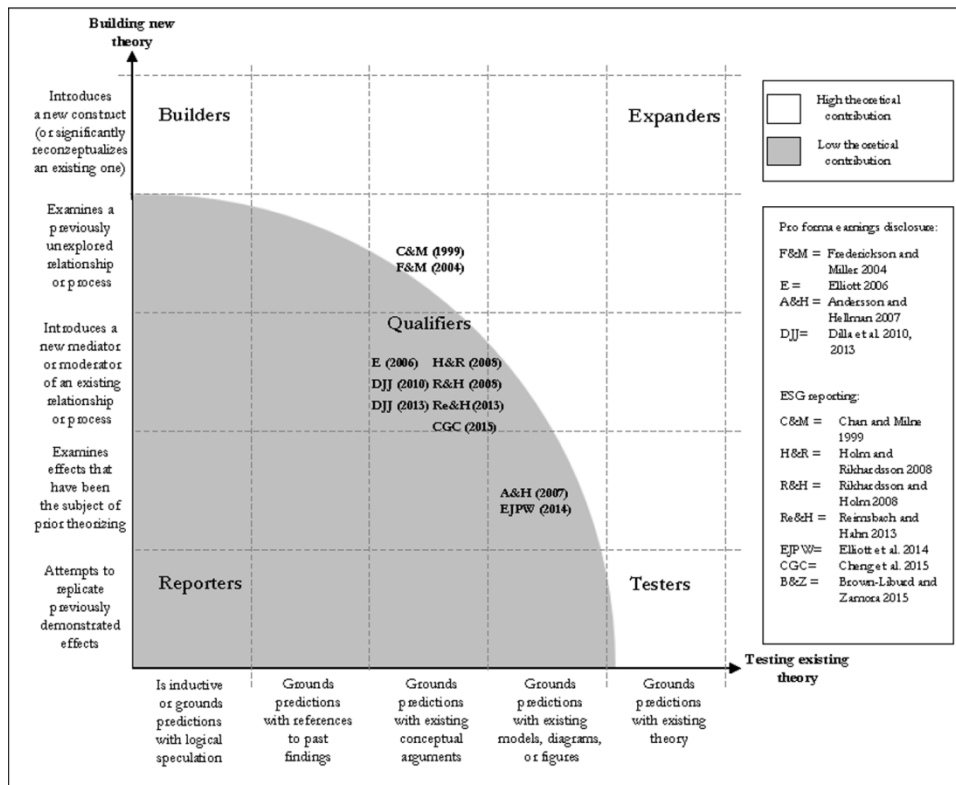
case, good ESG performance positively affects investors' stock-price assessments only when ESG assurance also is present.

Given that the studies on pro forma earnings disclosure and ESG reporting discussed in this paper have been published in top-tier accounting and business journals, it is not surprising that all make a theoretical contribution. However, to further examine how these studies advance the stream of research by testing and building theory, we apply Colquitt and Zapata-Phelan's (2007) taxonomy.

3.3 Theoretical contributions of experimental research on pro forma earnings disclosure and ESG reporting

In Figure 1, Colquitt and Zapata-Phelan's (2007) taxonomy is used to classify the theoretical contribution of the experimental studies discussed in this paper.

Figure 1 Taxonomy of theoretical contributions of experimental research on pro forma earnings disclosure



Source: (cf. Colquitt and Zapata-Phelan 2007, p.283)

First, all studies discussed here follow the *hypothetico-deductive model*, that is, they use theory or theory substitutes to formulate hypotheses that are then tested with observations (Popper, 1979). However, this does not imply that these studies contribute only to theory testing; they might also build theory, although in a different fashion than *inductive* approaches that generate theory through inductive reasoning, starting from empirical

observations (Colquitt and Zapata-Phelan, 2007). In the case of pro forma earnings disclosure, Frederickson and Miller's (2004) study contributes to building accounting theory by examining the relationship between pro forma earnings disclosure and investor reactions. This relationship can be classified as 'previously unexplored' as archival evidence suggested a causal link (e.g., Bhattacharya et al., 2003; Doyle et al., 2003) but, as noted above, faced the problem of disentangling the effect of pro forma earnings disclosure on investors' decisions from other investment-related effects. Furthermore, findings from archival research provided part of the motivation for the experiment and acted as groundings for the experimenters' hypotheses (e.g., Frederickson and Miller, 2004). As these archival findings are linked to existing conceptual arguments from psychological research (e.g., information processing, specifically framing and anchoring), Frederickson and Miller's (2004) work can also be classified as testing theory substitutes of accounting and psychology as the fundamental discipline. However, research questions and hypotheses are not derived from a holistic theoretical framework. Thus, Frederickson and Miller's (2004) study is *exploratory* which inevitably limits the extent of theory testing to a moderate level. However, sometimes no such framework exists that links the rather unexplored relationships of interest (also see Colquitt and Zapata-Phelan, 2007). In the case of ESG reporting, Chan and Milne (1999) are the first to analyse the causal relationship between ESG reporting and investors' allocation decisions. The authors explicitly state that no archival or experimental evidence on investors' reactions to negative ESG disclosures existed at the time. Thus again, they examine a previously unexplored relationship. However, as in the case of Frederickson and Miller (2004), they primarily test theory substitutes (i.e., from psychology and finance) and they do not derive their hypotheses from a holistic theoretical framework.

Subsequent studies inevitably shift focus to exploring additional mediators that explain the core relationships such as the participants' overall level of confidence in their investment decision (Andersson and Hellman, 2007), the investment horizon (Holm and Rikhardsson, 2008) and the specific nature of negative ESG information (Reimsbach and Hahn, 2015). Another starting point is the integration of further antecedents. In the studies on pro forma earnings disclosure, for example, these antecedents encompass the existence of quantitative pro forma to GAAP reconciliations (Elliott, 2006), the graphical display of pro forma information (Dilla et al., 2013) and the level of non-professional investors' financial knowledge (Dilla et al., 2010). In the studies on ESG reporting, the antecedents are, amongst others, assurance of ESG information (Cheng et al., 2015; Brown-Liburd and Zamora, 2015) and (professional) investors' experience level (Holm and Rikhardsson, 2008). One might argue that all these studies do not build 'true theory' (Sutton and Staw, 1995) as they do not examine completely unexplored relationships or introduce new constructs. These studies might contribute, however, to moving (behavioural) accounting theory towards a more solid theoretical foundation, by providing *inputs* for more comprehensive theorising (Colquitt and Zapata-Phelan, 2007). These studies thus 'build' theory in a different sense. In addition, studies subsequent to seminal work might also contribute to theory testing by grounding their hypotheses more thoroughly with existing conceptual arguments or models. Andersson and Hellman (2007) for example provide a more detailed analysis of the framing and (multi-) anchoring heuristics that link the antecedent (pro forma disclosure) to the consequential (investor reaction), thus giving more strength to their experimental results. Elliott et al. (2014) explain the previously examined link between ESG reporting and (unintentional)

investor reaction more thoroughly with the ‘affect-as-information’ theory from psychology.

As a result, Colquitt and Zapata-Phelan’s (2007) taxonomy allows us to present a graphical display of the degree of theory building and theory testing inherent in the experimental studies discussed in this paper. Overall, we observe a very similar pattern for the experimental research on the effects of pro forma earnings disclosure and ESG reporting. Referring to the five distinct paper types defined by Colquitt and Zapata-Phelan (2007), the experimental studies discussed here would be classified as qualifiers, offering a moderate level of theory building and theory testing. Thus, the taxonomy also identifies a lack of approaches that introduce new constructs or grounds predictions with existing holistic theory. Next, we reflect on whether this is a consequence of the experimental method in general or whether it is possible (and reasonable) to enhance the degree of theory building and testing when conducting experimental research.

4 Realising the full potential of experiments in building and testing accounting theory

We have concluded that the experimental studies on pro forma earnings disclosure and ESG reporting remain somewhat limited in their theoretical contribution. These experiments neither ground their hypotheses in a single holistic theoretical framework nor do their findings actually introduce a truly new construct. This is partly owing to the nature and focus of experiments that often aim at validating archival findings (Libby et al., 2002) but also shows the potential to further advance the contribution of experimental research. One possible direction for enhancing the theoretical contribution of experimental research could thus be to first *develop* a general theory that is then tested using (experimental) observations. This approach requires thorough and explicit theorising first and is thus quite ambitious (Libby et al., 2002). Maines and McDaniel (2000) and Hirst et al. (2007) prove the feasibility of this procedure, however. Maines and McDaniel (2000) analyse the effects of comprehensive-income characteristics on investors’ judgements using a self-developed framework. Hirst et al. (2007) first develop a conceptual model of the effects of forecast disaggregation on management earnings forecast credibility before experimentally testing their hypotheses. Another angle of enhancing the theoretical contribution of experimental research is using a *sequence* of experiments, as proposed in Koonce et al. (2008). The authors first use an experiment to discriminate among different possible theories that explain how investors evaluate managers who choose (not) to use derivatives once the outcomes of those decisions become known. Once the most descriptive theory is identified, Koonce et al. (2008) test its validity and generalisability in several subsequent experiments.

Furthermore, validity of research is key to robust findings that have the potential to move the whole discipline towards theory building and testing (e.g., Grant and Berry, 2011). Although experimental research has the potential to provide a high degree of internal validity, findings derived from other methodologies such as field studies and archival research can offer a higher degree of external validity (Allee et al., 2007). The combined use of multiple complementing methods (Colquitt, 2008), also known as ‘triangulation’ (Hageman, 2008), thus has the potential to increase the theoretical contribution of empirical studies (also see Falk and Heckman, 2009). However, applying

multiple methods *within* a single study (intra-study triangulation), for example experimental and archival methods, is a very ambitious and complex approach (Scandura and Williams, 2000), and often exceeds the scope of a manuscript that editors consider acceptable for publication (also see Molina-Azorin, 2010; Bettis, 2012). Consequently, examples of multiple method studies in accounting are rare, with a few exceptions in management accounting research (e.g., Modell and Lee, 2001).

The degree of building and testing theory provided by a single study critically depends on prior designs and findings within the related stream of research. According to Colquitt and Zapata-Phelan (2007), introducing *new* relationships and constructs is essential for building theory, with replicating studies thus remaining at a low level of theory building. What is classified as ‘new’, however, depends on what previous research has already examined. Theory testing by definition is grounded in existing theory and is inevitably based on past findings. Consequently, assessing a theoretical contribution might reasonably be directed at a *stream* of research, rather than at a single research study. From this perspective, it seems desirable that a stream of research as a whole builds and tests theory, thus advancing the accounting discipline. Consequently, triangulation might be understood as *triangulation of studies* (inter-study triangulation), that is, aligning different *separate* studies on the same topic that use different research methods (also see Hageman, 2008). Different experimental and archival studies could thus be implicitly or explicitly aligned to analyse the same underlying constructs, thus helping to develop a comprehensive view of the research phenomenon (Hageman, 2008). However, triangulation of studies is still rare in accounting (Lillis, 2005; Merchant and van der Stede, 2006; Hageman, 2008). There are notable exceptions though, also in the research on pro forma earnings disclosure that we refer to. Allee et al. (2007) are the first to explicitly validate experimental findings using archival research. Their study explicitly aims at complementing the experimental evidence of Frederickson and Miller (2004) and Elliott (2006) using trade-size-based proxies to distinguish the trading activities of non-professional investors from professional investors. The results reveal the abnormal trading behaviour of less-sophisticated investors after press releases that contain a pro forma number are disclosed. This result is consistent with Frederickson and Miller’s (2004) experimental evidence. Consistent with Elliott’s (2006) results, the authors also find that less-sophisticated investors rely more on pro forma earnings when they are emphasised by placing them before the GAAP earnings number in the release.

Obviously, the role of experimental research in triangulation is neither predetermined nor limited to theory testing or validating previously derived archival-empirical results. Rather, as Modell (2005) stresses, theory-testing approaches are primarily geared towards external and construct validation of converging findings, whereas theory-building approaches are typically based on internal validation efforts stemming from further probing of unexpected or inconclusive findings. Although experimental research has the potential to provide a high degree of internal validity, findings derived from other methodologies such as field studies and archival research can offer a higher degree of external validity (Allee et al., 2007). We therefore see an underused potential of exploratory experimental research directed at building accounting theory and call for further exploration of the full potential of experiments in advancing accounting theory. Again referring to the examples of pro forma earnings disclosure and ESG reporting, this becomes even more evident, considering the number of published experimental studies (four and seven) vs. the numerous archival-empirical studies (for an overview of the extensive literature, see e.g., Entwistle et al., 2010 and Hahn and Kühnen, 2013).

The examples of pro forma earnings disclosure and ESG reporting also show that archival research findings in which causality is diffuse always have been part of the motivation to conduct experiments. In this context, recent findings of archival research could be tested under controlled laboratory conditions. In the case of pro forma earnings disclosure, such findings are, for example, the mediating effect of corporate governance on investor reaction to pro forma earnings disclosure (Jennings and Marques, 2011). Corporate governance elements, such as non-executive directors, foreign nationals (e.g., Khan, 2010) and women representation in the board (Fernandez-Feijoo et al., 2014) also are topical issues in current archival ESG reporting research. Thus, incorporating corporate governance aspects could be a promising avenue, in general, for future experimental research on reporting practices. Inversely, following Allee et al.'s (2007) idea, recent experimental findings could be further validated in terms of their generalisability to real-world settings using archival methods explicitly designed to analyse the same underlying constructs. For example Dilla et al.'s (2010) experimental findings indicate that investor reactions to the disclosure of pro forma earnings do not vary based on the pure existence but rather on the presentation form and direction (GAAP to pro forma vs. pro forma to GAAP) of a quantitative reconciliation. Where Allee et al. (2007) use a proxy simply for the existence of a reconciliation, future research should use a more refined indicator variable that differentiates between the possible categories of reconciliations (also see Marques, 2010; Zhang and Zheng, 2011), thus testing the robustness of Dilla et al.'s (2010) results in the archival setting. In the ESG reporting context, the important role of assurance as indicated by the experimental results of Cheng et al. (2015) and Brown-Liburd and Zamora (2015) could be further supported by additional archival evidence. Combining laboratory and field studies might thus strengthen the discipline's overall confidence in the validity of the findings if they persist in the archival and experimental setting (Allee et al., 2007; Colquitt, 2008). However, the (chronological) order of experimental and archival-empirical research is not predetermined, but often a scholar's deliberate choice. Within triangulation, a method's resource efficiency (see Libby et al., 2002) can thus be an important criterion when choosing between alternative research approaches. Another criterion that influences the order of experimental and archival-empirical research is illustrated by Kachelmeier and King (2002). Often, when evaluating accounting policy issues, an ex-ante perspective of research is essential for analysing 'likely reactions' to different alternative policies that could be implemented (Kachelmeier and King, 2002, also see Sprinkle, 2003). In these situations where the policies (and thus any real-world data) do not yet exist, experiments still have the potential to build policy-relevant theoretical insights, or in other words, theories (Kachelmeier and King, 2002). In those cases, experimental research is characterised by a high 'practical utility' (Corley and Gioia, 2011) because it explicitly builds a link from accounting research to accounting policy (Kachelmeier and King, 2002). Again picking up the approach of triangulation, results of such ex ante policy-directed experiments can be further investigated using archival approaches, once the policies are in place.

5 Conclusion

In this paper, we reflected on what experimental studies on topics related to financial accounting can or cannot do: testing or building theory? This is an important question,

given that theory is the ‘basic aim of science’ (Kerlinger and Lee, 2000, p.11). We argue that Colquitt and Zapata-Phelan’s (2007) taxonomy is a means of guiding critical discourse on the theoretical contribution inherent in any (experimental) study. This taxonomy has been very influential in guiding the research agenda in the management discipline and, given its generic nature, it also has high potential for mapping past and guiding future (experimental) accounting research. Here, it is well accepted that well-designed experiments allow researchers to test causal relationships under controlled conditions, or in other words, to test theory. However, we argue that experimental research has an underused potential to build accounting theory (Callahan et al., 2006). Kahneman (2012), reflecting on his contribution to developing behavioural decision theory, emphasises the necessity of exploratory experimental research in the early stages of a stream of research, using the internal validity advantages of experimental efforts to investigate phenomena where causality is diffuse. Following this line of reasoning, we shift our focus from the level of a single study to the more aggregated level of a stream of research. Here, a triangulation of studies aligning experimental and archival methodologies has the potential to produce results that are robust to internal and external validity concerns (Allee et al., 2007), thus advancing accounting theory. Therefore, triangulating studies, which is also suggested and used in several other disciplines such as organisational sciences (e.g., Grant and Berry, 2011) and supply chain management (Siemsen, 2011), seem a promising venue for future research in accounting. However, within triangulating studies, the role of experimental research is not predetermined and might encompass confirmatory theory-testing approaches and more exploratory theory-building approaches. The latter is also well acknowledged in other disciplines such as psychology (see again Kahneman, 2012) and natural sciences (e.g., Franklin, 2005). With this paper, we want to raise scholars’ awareness of these often-neglected issues and broaden the appeal of experimental research. Although we referred to two specific examples of experimental research in financial accounting, we believe that our arguments generalise to various other aspects in accounting.

Acknowledgements

We thank Editor Ian Crawford as well as two reviewers for their very helpful and constructive comments.

References

- Aguinis, H. and Bradley, K.J. (2014) ‘Best practice recommendations for designing and implementing experimental vignette methodology studies’, *Organizational Research Methods*, Vol. 17, No. 4, pp.351–371.
- Ahrens, T. and Chapman, C.S. (2006) ‘Doing qualitative field research in management accounting: Positioning data to contribute to theory’, *Accounting, Organizations and Society*, Vol. 31, No. 8, pp.819–841.
- Allee, K.D., Bhattacharya, N., Black, E.L. and Christensen, T.E. (2007) ‘Pro forma disclosure and investor sophistication: external validation of experimental evidence using archival data’, *Accounting, Organizations and Society*, Vol. 32, No. 3, pp.201–222.
- Andersson, P. and Hellman, N. (2007) ‘Does pro forma reporting bias analyst forecasts?’, *European Accounting Review*, Vol. 16, No. 2, pp.277–298.

- Berthelot, S., Coulmont, M. and Serret, V. (2012) 'Do investors value sustainability reports? A Canadian study', *Corporate Social Responsibility and Environmental Management*, Vol. 9, No. 6, pp.355–363.
- Bettis, R. A. (2012) 'The search for asterisks: compromised statistical tests and flawed theories', *Strategic Management Journal*, Vol. 33, No. 1, pp.108–113.
- Bhattacharya, N., Black, E.L., Christensen, T.E. and Larson, C. (2003) 'Assessing the relative informativeness and permanence of pro forma earnings and GAAP operating earnings', *Journal of Accounting and Economics*, Vol. 36, Nos. 1–3, pp.285–319.
- Bhattacharya, N., Black, E.L., Christensen, T.E. and Mergenthaler, R.D. (2004) 'Empirical evidence on recent trends in pro forma reporting', *Accounting Horizons*, Vol. 18, No. 1, pp.27–43.
- Bhattacharya, N., Black, E.L., Christensen, T.E. and Mergenthaler, R.D. (2007) 'Who trades on pro forma earnings information?', *The Accounting Review*, Vol. 82, No. 3, pp.581–619.
- Brown-Liburd, H. and Zamora, V.L. (2015) 'The role of corporate social responsibility (CSR) assurance in investors' judgments when managerial pay is explicitly tied to csr performance', *Auditing: A Journal of Practice & Theory*, Vol. 34, No. 1, pp.75–96.
- Callahan, C.M., Gabriel, E.A. and Sainty, B.J. (2006) 'A review and classification of experimental economics research in accounting', *Journal of Accounting Literature*, Vol. 25, pp.59–126.
- Chan, C.C.C. and Milne, M.J. (1999) 'Investor reactions to corporate environmental saints and sinners: an experimental analysis', *Accounting and Business Research*, Vol. 29, No. 4, pp.265–279.
- Cheng, M.M., Green, W.J. and Chi Wa Ko, J. (2015) 'The impact of strategic relevance and assurance of sustainability indicators on investors' decisions', *Auditing: A Journal of Practice & Theory*, Vol. 34, No. 1, pp.131–162.
- Colquitt, J.A. (2008) 'Publishing laboratory research in AMJ: a question of when, not if', *Academy of Management Journal*, Vol. 51, No. 4, pp.616–620.
- Colquitt, J.A. and Zapata-Phelan, C.P. (2007) 'Trends in theory building and theory testing: five-decade study of the academy of management journal', *Academy of Management Journal*, Vol. 50, No. 6, pp.1281–1303.
- Corley, K.G. and Gioia, D.A. (2011) 'Building theory about theory building: what constitutes a theoretical contribution?', *Academy of Management Review*, Vol. 36, No. 1, pp.12–32.
- Dhaliwal, D.S., Radhakrishnan, S., Tsang, A. and Yang, Y.G. (2012) 'Nonfinancial disclosure and analyst forecast accuracy: international evidence on corporate social responsibility disclosure', *The Accounting Review*, Vol. 87, No. 3, pp.723–759.
- Dilla, W.N., Janvrin, D.J. and Jeffrey, C. (2013) 'The impact of graphical displays of pro forma earnings information on professional and nonprofessional investors' earnings judgments', *Behavioral Research in Accounting*, Vol. 25, No. 1, pp.37–60.
- Dilla, W.N., Janvrin, D.J. and Jeffrey, C. (2010) *Pro Forma Accounting Reconciliation Disclosures: The Effect of Financial Reporting Knowledge and Information Viewing Behavior on Judgments of Nonprofessional Investors*, 21 July, 2010, Available at SSRN: <http://ssrn.com/abstract=1646618> or <http://dx.doi.org/10.2139/ssrn.1646618>
- Doty, D.H. and Glick, W.H. (1994) 'Typologies as a unique form of theory building: toward improved understanding and modeling', *Academy of Management Review*, Vol. 19, No. 2, pp.230–251.
- Doyle, J.T., Lundholm, R.J. and Soliman, M.T. (2003) 'The predictive value of expenses excluded from pro forma earnings', *Review of Accounting Studies*, Vol. 8, Nos. 2–3, pp.145–174.
- Elliott, W.B. (2006) 'Are investors influenced by pro forma emphasis and reconciliations in earnings announcements?', *The Accounting Review*, Vol. 81, No. 1, pp.113–133.

- Elliott, W.B., Jackson, K.E., Peecher, M.E. and White, B.J. (2014) 'The unintended effect of corporate social responsibility performance on investors' estimates of fundamental value', *The Accounting Review*, Vol. 89, No. 1, pp.275–302.
- Entwistle, G., Feltham, G. and Mbagwu, C. (2006) 'Misleading disclosure of pro forma earnings: an empirical examination', *Journal of Business Ethics*, Vol. 69, No. 4, pp.355–372.
- Entwistle, G.M., Feltham, G.D. and Mbagwu, C. (2010) 'The value relevance of alternative earnings measures: a comparison of pro forma, GAAP, and I/B/E/S earnings', *Journal of Accounting, Auditing and Finance*, Vol. 25, No. 2, pp.261–288.
- Falk, A. and Heckman, J.J. (2009) 'Lab experiments are a major source of knowledge in the social sciences', *Science*, Vol. 326, No. 5952, pp.535–538.
- Fernandez-Feijoo, B., Romero, S. and Ruiz-Blanco, S. (2014) 'Women on boards: Do they affect sustainability reporting?', *Corporate Social Responsibility & Environmental Management*, Vol. 21, No. 6, pp.351–364.
- Franklin, L.R. (2005) 'Exploratory Experiments', *Philosophy of Science*, Vol. 72, No. 5, pp.888–899.
- Frederickson, J.R. and Miller, J.S. (2004) 'The effects of pro forma earnings disclosures on analysts' and nonprofessional investors' equity valuation judgments', *The Accounting Review*, Vol. 79, No. 3, pp.667–686.
- Gioia, D.A. and Pitre, E. (1990) 'Multiparadigm perspectives on theory building', *The Academy of Management Review*, Vol. 15, No. 4, pp.584–602.
- Grant, A.M. and Berry, J.W. (2011) 'The necessity of others is the mother of invention: intrinsic and prosocial motivations, perspective taking, and creativity', *Academy of Management Journal*, Vol. 54, No. 1, pp.73–96.
- Hageman, A.M. (2008) 'A review of the strengths and weaknesses of archival, behavioral, and qualitative research methods: recognizing the potential benefits of triangulation', in Arnold, V., Clinton, B.D., Lillis, A., Roberts, R., Wolfe, C. and Wright, S. (Eds.): *Advances in Accounting Behavioral Research*, Emerald, Bingley, Vol. 11, pp.1–30.
- Hahn, R. and Kühnen, M. (2013) 'Determinants of sustainability reporting: a review of results, trends, theory, and opportunities in an expanding field of research', *Journal of Cleaner Production*, Vol. 59, pp.5–21.
- Highhouse, S. (2009) 'Designing experiments that generalize', *Organizational Research Methods*, Vol. 12, No. 3, pp.554–566.
- Hirst, D.E., Koonce, L. and Simko, P.J. (1995) 'Investor reactions to financial analysts' research reports', *Journal of Accounting Research*, Vol. 33, No. 2, pp.335–351.
- Hirst, D.E., Koonce, L. and Venkataraman, S. (2007) 'How disaggregation enhances the credibility of management earnings forecasts', *Journal of Accounting Research*, Vol. 45, No. 4, pp.811–837.
- Hitz, J-M. (2010) 'Press release disclosure of 'pro forma' earnings metrics by large german corporations – empirical evidence and regulatory recommendations', *Accounting in Europe*, Vol. 7, No. 1, pp.63–86.
- Holm, C. and Rikhardsson P. (2008) 'Experienced and novice investors: Does environmental information influence investment allocation decisions?' *The European Accounting Review*, Vol. 17, No. 3, pp.537–557.
- Inanga, E.L. and Schneider, B.W. (2005) 'The failure of accounting research to improve accounting practice: a problem of theory and lack of communication', *Critical Perspectives on Accounting*, Vol. 16, No. 3, pp.227–248.
- Jennings, R. and Marques, A. (2011) 'The joint effects of corporate governance and regulation on the disclosure of manager-adjusted non-GAAP earnings in the US', *Journal of Business Finance and Accounting*, Vol. 38, Nos. 3–4, pp.364–394.

- Kachelmeier, S.J. and King, R.R. (2002) 'Using laboratory experiments to evaluate accounting policy issues', *Accounting Horizons*, Vol. 16, No. 3, pp.219–232.
- Kahneman, D. (2012) *Thinking, Fast and Slow*, Penguin, London.
- Kerlinger, F.N. and Lee, H.B. (2000) *Foundations of Behavioral Research*, Thomson Learning, New York.
- Khan, Md. (2010) 'The effect of corporate governance elements on corporate social responsibility (CSR) reporting Empirical evidence from private commercial banks of Bangladesh', *International Journal of Law and Management*, Vol. 52, No. 2, pp.82–109.
- Koonce, L. and Mercer, M. (2005) 'Using psychology theories in archival financial accounting research', *Journal of Accounting Literature*, Vol. 24, pp.175–214.
- Koonce, L., Lipe, M.G. and McNally, M.L. (2008) 'Investor reactions to derivative use and outcomes', *Review of Accounting Studies*, Vol. 13, No. 4, pp.571–597.
- Koonce, L., Nelson, K.K. and Shakespeare, C.M. (2011a) 'Judging the relevance of fair value for financial instruments', *The Accounting Review*, Vol. 86, No. 6, pp.2075–2098.
- Koonce, L., Seybert, N. and Smith, J. (2011b) 'Causal reasoning in financial reporting and voluntary disclosure', *Accounting, Organizations and Society*, Vol. 36, Nos. 4–5, pp.209–225.
- KPMG (2013) *KPMG International Survey of Corporate Social Responsibility Reporting 2013*, retrieved from <http://www.kpmg.com/Global/en/IssuesAndInsights/ArticlesPublications/corporate-responsibility/Documents/corporate-responsibility-reporting-survey-2013.pdf>
- Libby, R., Bloomfield, R. and Nelson, M.W. (2002) 'Experimental research in financial accounting', *Accounting, Organizations and Society*, Vol. 27, No. 8, pp.775–810.
- Lillis, A.M. (2005) 'Cross-sectional field studies in management accounting research-closing the gaps between surveys and case studies', *Journal of Management Accounting Research*, Vol. 17, No. 1, pp.119–141.
- Lougee, B.A. and Marquardt, C.A. (2004) 'Earnings informativeness and strategic disclosure: an empirical examination of "pro forma" earnings', *The Accounting Review*, Vol. 79, No. 3, pp.769–795.
- Maines, L.A. and McDaniel, L.S. (2000) 'Effects of comprehensive-income characteristics on nonprofessional investors' judgments: the role of financial-statement presentation format', *The Accounting Review*, Vol. 75, No. 2, pp.179–207.
- Maines, L.A., Salamon, G.L. and Sprinkle, G.B. (2006) 'An information economic perspective on experimental research in accounting', *Behavioral Research in Accounting*, Vol. 18, pp.85–102.
- Malmi, T. and Granlund, M. (2009) 'In search of management accounting theory', *European Accounting Review*, Vol. 18, No. 3, pp.597–620.
- Marques, A. (2010) 'Disclosure strategies among S&P.500 firms: evidence on the disclosure of non-GAAP financial measures and financial statements in earnings press releases', *The British Accounting Review*, Vol. 42, No. 2, pp.119–131.
- Merchant, K.A. and van der Stede, W.A. (2006) 'Field-based research in accounting: accomplishments and prospects', *Behavioral Research in Accounting*, Vol. 18, No. 1, pp.117–134.
- Miller, K.D. and Tsang, E.W.K. (2011) 'Testing management theories: critical realist philosophy and research methods', *Strategic Management Journal*, Vol. 32, No. 2, pp.139–158.
- Modell, S. (2005) 'Triangulation between case study and survey methods in management accounting research: an assessment of validity implications', *Management Accounting Research*, Vol. 16, No. 2, pp.231–254.
- Modell, S. and Lee, A. (2001) 'Decentralization and reliance on the controllability principle in the public sector', *Financial Accountability and Management*, Vol. 17, No. 3, pp.191–218.

- Molina-Azorin, J.F. (2010) 'The use and added value of mixed methods in management research', *Journal of Mixed Methods Research*, Vol. 5, No. 1, pp.7–24.
- Moser, D.V. and Martin, P.R. (2012) 'A broader perspective on corporate social responsibility research in accounting', *The Accounting Review*, Vol. 87, No. 3, pp.797–806.
- Oler, D.K., Oler, M.J. and Skousen, C.J. (2010) 'Characterizing accounting research', *Accounting Horizons*, Vol. 24, No. 4, pp.635–670.
- Peecher, M.E. and Solomon, I. (2001) 'Theory and experimentation in studies of audit judgments and decisions: avoiding common research traps', *International Journal of Auditing*, Vol. 5, No. 3, pp.193–203.
- Popper, K.R. (1979) *Truth, Rationality and the Growth of Scientific Knowledge*, Klostermann, Frankfurt am Main.
- Reimsbach, D. and Hahn, R. (2015) 'The effects of negative incidents in sustainability reporting on investors' judgments – an experimental study of third-party versus self-disclosure in the realm of sustainable development', *Business Strategy and the Environment*, Vol. 24, No. 4, pp.217–235.
- Rikhardsson P, Holm C. (2008) 'The effect of environmental information on investment allocation decisions – an experimental study', *Business Strategy and the Environment*, Vol. 17, No. 6, pp.382–397.
- Scandura, T.A. and Williams, E.A. (2000) 'Research methodology in management: current practices, trends, and implications for future research', *Academy of Management Journal*, Vol. 43, No. 6, pp.1248–1264.
- Sprinkle, G.B. (2003) 'Perspectives on experimental research in managerial accounting', *Accounting, Organizations and Society*, Vol. 28, Nos. 2–3, pp.287–318.
- Sutton, R.I. and Staw, B.M. (1995) 'What theory is not', *Administrative Science Quarterly*, Vol. 40, No. 3, pp.371–384.
- Watts, R.L. and Zimmerman, J.L. (1986) *Positive Accounting Theory*, Prentice-Hall Inc., Englewood-Cliffs.
- Zhang, H. and Zheng, L. (2011) 'The valuation impact of reconciling pro forma earnings to GAAP earnings', *Journal of Accounting and Economics*, Vol. 51, Nos. 1–2, pp.186–202.

Notes

¹We count 129 citations in the web of Science database and 333 in Google Scholar.

²Dilla et al.'s (2010) study has not been published yet but is available as a working paper on SSRN.

³We use the terms 'sustainability report(ing)/disclosure' and 'ESG report(ing)/disclosure' interchangeably, because both describe corporate reporting that aims to communicate ESG-related information.