UNIVERSITY OF SAO PAULO
COLLEGE OF ECONOMICS, BUSINESS, AND ACCOUNTING
DEPARTMENT OF ACCOUNTANCY AND ACTUARIAL SCIENCE
GRADUATE PROGRAM IN CONTROLLERSHIP AND ACCOUNTING

Quality in the Process of Scientific Production in Accounting in Brazil

Qualidade no Processo de Produção Científica em Contabilidade no Brasil

José Renato Sena Oliveira

Advisor: Prof. Dr. Gilberto de Andrade Martins

SAO PAULO
2016

Dr. Marco Antonio Zago
Rector of the University of Sao Paulo

Dr. Adalberto Américo Fischmann
Dean of the College of Economics, Business, and Accounting

Dr. Gerlando Augusto Sampaio Franco de Lima
Chairman of the Department of Accountancy and Actuarial Science

Dr. Luiz Paulo Lopes Fávero
Coordinator of the Graduate Program in Controllership and Accounting

JOSÉ RENATO SENA OLIVEIRA

Quality in the Process of Scientific Production in Accounting in Brazil

Qualidade no Processo de Produção Científica em Contabilidade no Brasil

Dissertation submitted to the Graduate Program in Controllership and Accounting of the Department of Accountancy and Actuarial Science at the College of Economics, Business, and Accounting at the University of Sao Paulo, as a partial requirement to obtain the Doctoral degree in Science.

Advisor: Prof. Dr. Gilberto de Andrade Martins

Corrected version
(original version available at College of Economics, Business, and Accounting)

SAO PAULO
2016

CATALOGUING DATA
Prepared by Seção de Processamento Técnico do SBD/FEA/USP

Oliveira, José Renato Sena
Quality in the process of scientific production in accounting in Brazil / José Renato Sena Oliveira. – São Paulo, 2016.
167 p.

Tese (Doutorado) – Universidade de São Paulo, 2016.
Orientador: Gilberto de Andrade Martins.

1. Contabilidade – Pesquisa 2. Produção científica 3. Pesquisa científica 4. Pós-graduação I. Universidade de São Paulo. Faculdade de Economia, Administração e Contabilidade. II. Título.

CDD – 657.07

Name: Oliveira, José Renato Sena
Title: Quality in the Process of Scientific Production in Accounting in Brazil

Dissertation submitted to the Graduate Program in Controllership and Accounting of the Department of Accountancy and Actuarial Science at the College of Economics, Business, and Accounting at the University of Sao Paulo, as a partial requirement to obtain the Doctoral degree in Science.

Approved on: 04/28/2016.

Dissertation Committee:

Prof. Dr. **Gilberto de Andrade Martins (advisor and president)**
Institution: University of Sao Paulo Signature: _____

Prof. Dr. **Ana Maria Afonso Ferreira Bianchi**
Institution: University of Sao Paulo Signature: _____

Prof. Dr. **Jacqueline Veneroso Alves da Cunha**
Institution: Federal University of Minas Gerais Signature: _____

Prof. Dr. **Marcia Martins Mendes De Luca**
Institution: Federal University of Ceara Signature: _____

Prof. Dr. **Romualdo Douglas Colauto**
Institution: Federal University of Parana Signature: _____

Sao Paulo (Brazil), April 28, 2016.

This dissertation is dedicated to Brazilian accounting researchers with the hope that they feel challenged both to reflect on the quality of the knowledge produced in the field and to seek continuous improvement in their research practices.

ACKNOWLEDGMENTS

First, I would like to thank God for my life, opportunities, and the strength to overcome challenges.

I am grateful to my family, my safe harbor at all times: my parents, Antonio Aderaldo de Oliveira and Zenilda de Almeida Sena; my siblings, Adauto, Antonio Romeu, and Elenilda; Levi (my nephew and godson) and little Camila (my niece). I also thank Monica and Mosana, my sisters-in-law. I apologize for my absence during these long years. Thank you my uncle Luiz Júlio Carneiro (in memoriam) and my aunt Noélia Maria de Oliveira Carneiro: I am indebted to them for much of what I am today. Thank you for everything. Additionally, I am grateful to my uncle and godfather Otoniel Oliveira: it was hard being away and not being able to say a final goodbye. Rest in peace.

Special acknowledgment goes to my advisor, Professor Dr. Gilberto de Andrade Martins. Since the first moment that we began to discuss ideas about research quality, he has believed and embraced this proposal. Many of the ideas in this study have been inspired by our conversations and our concerns about how to produce knowledge in accounting with rigor and usefulness with the objective of bringing best practices to the science and our best contributions to society. He is tirelessly enthusiastic in discussing issues that might lead accounting knowledge to a higher scientific level, an idea that inspired this proposal. I also thank David Carter and Artur Roberto do Nascimento, who helped me to organize ideas in order to build the research design to apply for the graduate program.

I would like to thank the members of the qualifying board, Dr. Marcia Martins Mendes De Luca and Dr. Romualdo Douglas Colauto, for their significant contributions to improving this study. Additionally, I am extremely grateful to the faculty members in Brazilian graduate programs in accounting that kindly and voluntarily agreed to participate in this study; they were essential to helping me better understand the issues surrounding the quality and scientific research process of the accounting field in Brazil. I also thank the coordinators of Brazilian graduate programs in accounting who kindly forwarded requests to faculty members to participate in this study.

I thank the University of Sao Paulo (USP) for this great opportunity. I am grateful to the faculty of the Graduate Program in Accounting at FEA/USP for their lessons throughout the course. Special thanks to the professors who went to Feira de Santana City in 2012 to lead us through the early stages of the Interinstitutional Doctorate (DINTER): Dr. Gerlando Lima and Dr. Gilberto Martins (first half) and Dr. Bruno Salotti and Dr. Nelson Carvalho (second half). Thank you to Professor Dr. Edgard Cornacchione Jr, former Chairman of the Department of Accountancy and Actuarial Science at the FEA/USP, who was responsible for facilitating my time as a visiting scholar at University of Illinois at Urbana-Champaign (UIUC). Thank you to the DINTER coordinators at USP: Dr. Silvia Casa Nova and Dr. Luis Afonso (first stage) and Dr. Lucas Barros and Dr. Andson Aguiar (current stage). Finally, thank you to Professor Dr. Andson Aguiar and Professor Dr. Romualdo Colauto, my supervisors in Teaching Monitoring I and II, who gave me the opportunity for new experiences in undergraduate education.

I would like to thank the UIUC for accepting me to the Visiting Scholar Program of the Department of Educational Psychology (Epsy) at the College of Education and for providing me with support for carrying out my activities as visiting scholar during the 2014-2015 academic year. Additionally, I am grateful to Professor Jennifer C. Greene (Epsy/UIUC) both for having supervised me in the program and for her invaluable contributions to improving my research proposal for this dissertation. You are an example for the values of simplicity, competence, and professionalism. Thank you for everything. Thank you to

Professor Thomas Schwandt, who allowed me to audit his Introduction to Evaluation Theory class, and to Mrs. Amy Hager, visa coordinator of the College of Education. Finally, thank you to Jeff Arrigo, lecturer of the Advanced Academic Writing I and Advanced Presentation Skills I classes in the ESL Program of the Department of Linguistics/UIUC, for his valuable contributions to improving my English skills and for inspiring me to accept the challenge of writing this dissertation in English.

I thank the State University of Feira de Santana (UEFS), for all its support. Thank you to my friends from the Finance and Accounting Office, especially Carlos Roberto Suzarte, Chawelly Filgueiras, Fatima Parente, and Flor Morais, for encouraging me to accept the challenge of the doctoral program. Thank you to my colleagues and friends in the Department of Applied Social Sciences, and to Vilânia Santana and the team at the Research and Graduate Dean's Office (PPPG). Special thanks to my friends Professors Rosembergue Valverde and Rossine Cerqueira da Cruz, local coordinators of the DINTER at UEFS, who agreed to conduct the program with all its challenges. Thank you to my friend Tania Cristina Azevedo, who could not be with us on this journey, and Professor Marluce Maria Araújo Assis, former Pro-Rector for Research and Graduate at UEFS. Both of them were directly responsible for the realization of this DINTER, thus contributing immensely to the building project submitted to USP and CAPES.

To my DINTER colleagues – Ana Lucia Carvalho, Edmilson Sousa, Emilio Maltez, Iracema Raimunda Aragão, Marcia D'Souza, Raimundo Lima Filho, Robson Braga, and Sandra Maria Silva – thank you for the friendship, companionship, and persistence. I know how hard it was for each of you. Special thanks to Ana Lucia, Iracema and Sandra Maria, my Feira de Santana partners in the program. Up, up and away! Also to Sandro Soares and Bruna Camargos, friends who we met during the doctoral program.

I could not forget my friends who believed me and who kept me stimulated throughout. Thank you to Cláudia Cruz, my friend and my example of humility and competence as a researcher. Thank you also to Josue Braga and Renata Braga, for their friendship, encouragement, and support: I will forever remember the good moments in Urbana, IL, particularly those accompanied by Papa John's pizzas. Thank you to so many friends from Bahia who always cheered for my success, particularly Adriano Cosme Lima, Alcione Carvalho, Carlos Adriano Gomes, Cleberton Santos, Erivaldo Carneiro Júnior, Giminiano Santos, Graciela Barbosa, Jorge Rodrigues, Madalena Sousa, Maria Helena Martins, Naiana Vasconcelos, Norma Bernardo, Robério Souza, Ronaldo Nascimento, Thais Gabriele Ribeiro, and Wilson Jesus. Finally, thank you to my friends Gerlando Lima, Rodolfo Ferrari and Veronica Santana, partners in learning and fun moments at the office of the USP Conference in Accounting: the doctorate was less difficult because of those pranks and jokes: Good morning, Congress! Good afternoon, Congress! Thank you to the staff of PPGCC and EAC/FEA/USP for your support and friendship: Marcia Bento, Janaína Nogueira, Janaína Silva, Juliana Meneghim, Belinda Ludovici, Evandro Figueiredo, and Maria José Medeiros.

I thank my friends from Urbana-Champaign, who made my days in the US more pleasant: Débora Gonçalves and her family, Hélio Emmendoerfer Jr, Karulliny Vianna, Maria Carolina Hidalgo, and Rejane Canha and her family. Additionally, I want to thank my ESL colleagues at Parkland College (Champaign, IL), where I learned to overcome the challenges of a new language and learned to 'build bridges' with the help of Cheryl Lehman, George Uricoechea, Tracey Brown, and Wendy Patriquin, our ESL instructors.

Finally, I thank the Bahia State Support Foundation for Research (FAPESB), whose funding supported the Interinstitutional Doctorate in Accounting in the partnership between FEA/USP, UEFS and associates.

The value of scientific research is partially dependent on the ability of individual researchers to demonstrate the credibility of their findings.
LeCompte and Goetz (1982).

The most important purpose of evaluation is not to prove, but to improve.
Daniel Stufflebeam (2000).

ABSTRACT

Oliveira, José Renato Sena. **Quality in the process of scientific production in Accounting in Brazil.** (Doctoral dissertation). College of Economics, Business, and Accounting at University of Sao Paulo. Sao Paulo, 2016.

This study aims to analyze the building process of scientific production in accounting in light of the quality attributes of good research. In so doing, it attempts to identify the stages of the research process in which these characteristics are revealed, to compare the literature on research attributes with the attributes identified in respondents' practices, and to present an approach based on the attributes of good research to judge the quality of scientific production in the accounting field. This research uses the Modified Delphi Technique, which indicates that the first inputs came from the literature. The expert panel consisted of 37 faculty members from 19 Brazilian graduate programs in accounting who were recommended/recognized by the CAPES Foundation. Based on the literature, especially Brinberg and McGrath (1985), Spencer et. al. (2003), and Mays and Pope (2006), an orientation matrix was developed with 53 attributes/relationships related to general quality criteria and nine key features. Experts gave each proposition a grade from 1 to 10 based on their level of agreement regarding adherence to their research practices. The Delphi was applied in two rounds using online questionnaires with customized access. The findings reveal that most of the respondents obtained their doctoral degree in accounting at a national institution other than the one at which they currently work and that more than 70% of the respondents have been working as teachers or coordinators in graduate programs for fewer than seven years. With respect to the respondents' research experience, most serve on journals' editorial boards, act as journal referees, and at some point have obtained research funding from development institutions. Approximately one-third have received research productivity grants and almost one-half either currently serve or has served as a journal editor. Approximately 3/4 of the propositions achieved a strong level of agreement, and the following 10 propositions achieved a Delphi relative score of more than 90%: voluntary participation of subjects, goal/problem shown precisely, confidentiality of participants' data, conclusions versus aim, results comparison with other studies, checklist of findings versus purpose, literature review versus main concepts, theory to support propositions, useful strategy for purpose, and previous findings versus hypotheses. Those attributes that demonstrated low or moderate levels of agreement involved items that may compromise the quality and integrity of research, including those related to ethical principles, demonstrations of how researchers addressed errors and biases, and disclosure of the impact of the research team's participation on the results. The approach chosen meant discussing the relationships between criteria and fundamental features versus levels of agreement, followed by presentation of a logic model to evaluate the research process. The research concludes that certain current practices employed by Brazilian researchers in accounting – combined with the local institutional environment – contribute to reduce the quality of accounting research. This position is supported by the high dispersion of answers on various items and the low acceptance of attributes related to ethics, both of which are mandatory requirements under Brazilian law. Additionally, the low level of agreement on issues regarding the criteria of rigor and internal validity/credibility or defensibility – in addition to items related to rigor, integrity, and feasibility – reached only the moderate level.

Keywords: accounting – research, scientific production, scientific research, graduate studies, research quality, quality criteria.

RESUMO

Oliveira, José Renato Sena. **Qualidade no processo de produção científica em Contabilidade no Brasil.** (Tese de doutorado). Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo. São Paulo, 2016.

O propósito deste estudo é analisar o processo de construção da produção científica em Contabilidade à luz dos atributos de qualidade de uma boa pesquisa. Buscou-se observar em quais estágios do processo de pesquisa estes atributos são revelados, comparar os atributos da literatura com aqueles identificados nas práticas dos respondentes, e apresentar uma abordagem baseada em atributos de uma boa pesquisa para o julgamento da qualidade da produção científica na área. A metodologia utilizada foi a Técnica Delphi Modificada, em que as primeiras entradas têm origem na literatura. O painel de especialistas foi composto por 37 docentes de 19 Programas Brasileiros de Pós-Graduação (PPG) Stricto Sensu em Contabilidade recomendados/reconhecidos pela CAPES. Com base na literatura, especialmente Brinberg e McGrath (1985), Spencer et. al. (2003) e Mays e Pope (2006), foi elaborada uma matriz de orientação com 53 atributos/relações vinculados a critérios gerais de qualidade e a nove características-chave. Os especialistas atribuíram uma nota de 1 a 10 para cada atributo de acordo com o nível de concordância quanto à aderência às suas práticas de pesquisa. A Delphi foi aplicada em duas rodadas com o uso de questionários online e com acesso personalizado. Os achados revelaram que a maioria dos (das) respondentes tem Doutorado em Contabilidade obtido em uma Instituição nacional diferente da que atua, e mais de 70% têm até 7 anos de docência ou coordenação de PPG. Sobre a experiência de pesquisa, a maioria faz parte de conselhos editoriais de periódicos da sua área, atua como avaliador de periódicos e tem, ou já obteve, financiamento de instituições de fomento para pesquisa. Cerca de um terço tem bolsa de produtividade em pesquisa e quase metade atua ou já atuou como editor de periódico. Aproximadamente 3/4 das proposições receberam nível forte de concordância, e 10 delas alcançaram percentual superior a 90% no escore relativo da Delphi: participação voluntária dos sujeitos, objetivo/problema mostrado precisamente, confidencialidade dos dados dos participantes, conclusões versus objetivos, comparação de resultados com outros estudos, checar links entre achados versus propósito, revisão de literatura versus principais conceitos, teoria para suportar as proposições, estratégia útil ao propósito, e achados anteriores versus hipótese. Por outro lado, atributos que atingiram níveis baixo ou moderado de aceitação mostraram itens que podem comprometer a qualidade e integridade da investigação, como aqueles relacionados aos princípios éticos, à demonstração de como os pesquisadores lidaram com erros e vieses, e à divulgação dos impactos da participação da equipe de pesquisa sobre os resultados. A abordagem proposta consistiu na discussão das relações entre critérios e características-chave versus níveis de concordância, seguida da apresentação de um modelo lógico para avaliar o processo de pesquisa. Concluiu-se que algumas práticas correntes utilizadas por pesquisadores da área contábil no Brasil, combinadas com o ambiente institucional local, contribuem para a diminuição da qualidade da pesquisa no campo. Esta posição é apoiada pela elevada dispersão das respostas em vários itens e pela baixa aceitação de atributos relacionados à ética e que são requisito obrigatório pela legislação brasileira. Adicionalmente, pelo baixo nível de concordância sobre questões relacionadas com o rigor e a validade interna/credibilidade ou defensibilidade, bem como para itens relativos ao rigor, integridade e viabilidade, que atingiram o nível moderado.

Palavras-chave: contabilidade – pesquisa, produção científica, pesquisa científica, pós-graduação, qualidade da pesquisa, critérios de qualidade.

LIST OF ILLUSTRATIONS

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

AEA	American Evaluation Association
ANPAD	Associação Nacional de Pós-Graduação e Pesquisa em Administração (National Association of Graduate and Research in Administration)
CAPES	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Coordination for the Improvement of Higher Level Personnel)
CAR	Contemporary Accounting Research
CEP	Comitê de Ética em Pesquisa com Seres Humanos (Ethics Committee for Research with Human Beings)
CNS	Conselho Nacional de Saúde (Brazilian National Health Council)
CONEP	Comissão Nacional de Ética em Pesquisa (National Research Ethics Commission)
DAV	Diretoria de Avaliação (Evaluation Board)
DINTER	Doutorado Interinstitucional (Interinstitutional Doctorate)
ENANPAD	Encontro da Associação Nacional dos Programas de Pós-Graduação em Administração (Meeting of the National Association of Graduate Programs in Administration)
FAPESB	Fundação de Amparo à Pesquisa do Estado da Bahia (Bahia State Support Foundation for Research)
FAPESP	Fundação de Amparo à Pesquisa do Estado de São Paulo (Sao Paulo Research Foundation)
FEA	Faculdade de Economia, Administração e Contabilidade (College of Economics, Business, and Accounting)
HEI	Higher Education Institutions
IF	Impact Factor
ISI	Institute for Scientific Information
JAR	The Journal of Accounting Research
JCR	Journal Citation Report
KEC	Key Evaluation Checklist
OADS	Office of the Associate Director for Science
PPG	Programa de Pós-Graduação (Graduate Program)
PPGCC	Programa de Pós-Graduação em Controladoria e Contabilidade (Graduate Program in Controllership and Accounting)
PPPG	Pró-Reitoria de Pesquisa e Pós-Graduação (Research and Graduate Dean's Office)

PTC	Points to Consider
RAE	Research Assessment Exercises
SciELO	Scientific Electronic Library Online
TCLE	Termo de Consentimento Livre e Esclarecido (Informed Consent)
UEFS	Universidade Estadual de Feira de Santana (State University of Feira de Santana)
UIUC	University of Illinois at Urbana-Champaign
UK	United Kingdom
USP	Universidade de São Paulo (University of Sao Paulo)
VNS	Validity Network Schema

TABLE OF CONTENTS

page

CHAPTER 1 –
INTRODUCTION

The purpose of this chapter is to show a comprehensive and detailed overview of this research. The first topic provides the study's context and presents the research problem and goals. The second topic posits the study's hypotheses and the statement of the thesis to be defended. The third topic discusses the study's justifications, the research gap, and the study's expected contributions. Finally, the fourth topic acknowledges the study's limitations.

1.1 Context, Research Problem, and Goals

Research is an important activity at educational institutions. Research may develop and improve techniques, work strategies, and action courses in multiple fields of activity. Additionally, it is essential to both advancing knowledge and finding or framing solutions to social problems. The Belmont Report (National Institutes of Health, 1979), which establishes ethical principles for the practice of research involving human subjects, describes research as follows:

> The term 'research' designates an activity designed to test an hypothesis, permit conclusions to be drawn, and thereby to develop or contribute to generalizable knowledge (expressed, for example, in theories, principles, and statements of relationships). Research is usually described in a formal protocol that sets forth an objective and a set of procedures designed to reach that objective.

This definition highlights certain features of research that every person or institution involved in research activity must consider, particularly when that research activity is conducted for the purpose of acquiring knowledge. According to the scientific view, which is presented in the Belmont Report, the production of scientific knowledge implies organized and concatenated actions that can both generate generalizable knowledge and expand the boundaries of science. Moreover, this concept could be associated with the "step-by-step" logic of knowledge building. Another definition of research is "a systematic investigation including research development, testing, and evaluation, designed to develop or contribute to generalizable knowledge" (OHRP & DHHS, 2009). In this context, generalization relates to the researcher's intention (i.e., whether he/she intends to contribute to expanding knowledge). Both definitions are consistent with the idea of domains in scientific investigation (Brinberg & McGrath, 1985; Brinberg, 1982), in which research is essentially a study of relations

between units (or unit elements). These units are embedded and involve some content of interest (the substantive domain), ideas that yield meaning (the conceptual domain), and techniques and procedures by which those ideas can be studied (the methodological domain).

The institutional environment of research in Brazil sets basic productivity standards for graduate programs (Masters and Ph.D.). These patterns are consistent with the requirements of governmental agencies that provide funding and grants to researchers and institutional infrastructure for research; these requirements are established by the Brazilian Federal Agency for the Support and Evaluation of Graduate Education. On the one hand, these standards encourage scientific production in several fields; on the other, the need to reach productivity goals may reduce the quality of the knowledge produced. In other words, as a result of the eagerness to publish the results of research, the attention to scientific rigor and relevance maybe be pushed into the background. Thus, the institutional environment justifies studies that analyze the quality of scientific production and "how it is done," also known as the research process. Under this approach, the research process "is the identification, selection, combination, and use of elements and relations from the conceptual, methodological, and substantive domains" (Brinberg & McGrath, 1985, p. 16).

Crotty (1998) argues that research planning can be based on two main questions: "First, what methodologies and methods will we be employing in the research we propose to do? Second, how do we justify this choice and use of methodologies and methods?" To do so, the research process in the context of the social sciences involves four elements: (a) methods; (b) methodology; (c) theoretical perspective; and (d) epistemology."

The order of these elements is not fixed, perhaps because philosophical (theoretical and epistemological) assumptions will typically be revealed prior to the research. Assumptions must be considered by a researcher during the planning, i.e., before the implementation and accomplishment of the research. A carefully planned research project gives the researcher more confidence about each step that he/she should follow, how to mitigate risks and limitations, and how to identify matters that can compromise feasibility. By following a plan, it becomes possible to obtain consistent outcomes that are both valid and reliable.

By adopting an international approach, some accounting studies that involve different views of accounting research might be highlighted, such as those studies involving innovation and relevance in the field (Reiter & Williams, 2002), experimental research in financial accounting with a predictive validity framework (Libby, Bloomfield, & Nelson, 2002), and different approaches in managerial accounting (Baxter & Chua, 2003). Another example

involves a survey of UK accounting and finance academics that analyzes why those academics do not publish in top American journals (Brinn, Jones, & Pendlebury, 2001). That study analyzes respondents' perceptions to present a set of reasons for this failure to publish in such journals, including methodological and research-quality problems.

Studies on the evaluation of accounting research at international level address multiple aspects, such as productivity and quality involving university business departments (Chan, Chang, Tong, & Zhang, 2012; M.J. Jones, Brinn, & Pendlebury, 1996; Lowe & Locke, 2005). Other studies address research quality criteria, such as impact (Brown & Gardner, 1985; Carmona, 2006); relevance (Reiter & Williams, 2002); rigor (Evans, Feng, Hoffman, Moser, & Van der Stede, 2015; Williams, 2014); and validity (Libby et al., 2002). Finally, a set of studies explores citation analysis (Brown & Gardner, 1985; Dunbar & Weber, 2014); and perceptions of accounting journals quality, also called peer reviews (Ballas & Theoharakis, 2003; Brinn et al., 2001; Brown & Huefner, 1994; Lowe & Locke, 2005; Lowensohn & Samelson, 2006; Taylor, 2011). Such studies, though not specifically focused on the quality attributes of good research, explore quality criteria and their relationship with productivity and quality evaluation of what is published in scientific journals in the area.

In the Brazilian environment, epistemological studies of accounting research have adopted various approaches. For example, some studies present a bibliometric perspective and address aspects of methodological paths used in theses and other publications in major journals and conferences (Mendonça Neto, Riccio, & Sakata, 2009; Miranda, Azevedo, & Martins, 2011; Theóphilo & Iudícibus, 2005). Other studies address ethical issues in accounting research (Andrade, 2011; Antunes, Mendonça Neto, Oyadomari, & Okimura, 2011), citation analysis in accounting journals (Aragão, Oliveira, & Lima, 2014). Conversely, another study analyzes academic productivity (Martins & Lucena, 2014), and specifically faculty productivity in Brazilian graduate accounting programs. That study investigates the profile and main scientific production practices of those programs' faculty. It highlights the issue that arises when various parts of the results from a single study are published in various papers (a practice known as "salami science"). Additionally, following rejection by journals with higher impact, researchers submit papers to less-prestigious journals until they are accepted for publication.

These studies and their issues stress that bibliometric characterization based on reports and publications is inadequate to reveal research shortcomings in the accounting field research process. Accounting research must be reflected throughout the construction and development process – from the topic choice of the investigation until submission of the

report for scientific dissemination. Therefore, the practices that can provide those studies with both suitability and scientific character must be understood; thus, a reflection on the elements that contribute to increasing scientific quality is important in this regard. Notably, other knowledge areas, such as health (Davies & Dodd, 2002; Kalichman & Plemmons, 2007) and education (Hostetler, 2005) have explored the concept of good research in their investigation practices.

Good research practices are defined as rules that researchers can follow to "assure the quality, objectivity, and integrity of research data" (Shamoo & Resnik, 2003). A good research idea is associated with choices of research strategies characterized as "fit for purpose," which means that some strategies are more appropriate than others for addressing specific research questions (Denscombe, 2010). In particular, the latter author advocates that the most important aspect of good research is that the research report shows that the choices made are reasonable and explicit.

Aiming to discuss and define research quality, Valentine (2009) argues that the answer to the question of "'what are the characteristics of a high quality study' depends in part on why the judgment is being done". Each study can have a wide range of stakeholders among whom the meaning of quality varies. The author describes quality as a fit (i.e., an adequate link) "between the study's goals and the study's design and implementation characteristics". This description might be associated both with the idea of domains (Brinberg & McGrath, 1985; Brinberg, 1982) and with the description of the research process (Crotty, 1998). In turn, the perception of quality in accounting research has been presented from different angles. For instance, an essay published in Accounting and Finance (Clarkson, 2012) defends the notion of research quality associated with three main factors: contribution, scientific credibility, and communication. Another study (Lowe & Locke, 2005) analyzes the association (in the perception of accounting researchers) between research paradigms and the quality of journals in the UK. Journal quality, as highlighted by these authors, has essentially been studied from two approaches: citation-based studies and perception studies based on surveys.

Based on this variety of concepts, it is understood that the attributes of good research (or research with quality) involve the entire research-building process. In other words, good research involves the foundations and adequate practices of scientific investigation to achieve the most appropriate answer to the proposed research problem. Given this context, the research question of the present study is as follows: What quality attributes can draw good research practices in the building process of scientific production in accounting? Moreover, the study is guided by two evaluation questions about research quality: Which practices are

presented in the research production process in accounting? Which attributes can delineate the quality evaluation of these practices? The building process of scientific production in this study involves the key features or different stages (Mays & Pope, 2006; Spencer, Ritchie, Lewis, & Dillon, 2003) in the process of planning, execution, and outcomes evaluation in a scientific investigation (Brinberg & McGrath, 1985).

This study aims to analyze the building process of scientific production in accounting in light of the quality attributes of good research. Additionally, as its specific goals, this study intends to do the following: (a) to identify the stages of the research process in which these attributes are revealed in the practices of Brazilian accounting researchers; (b) to compare features of the literature with those identified in accounting research practices; and (c) to present an approach based on the attributes of good research to assess the quality of scientific production in the field of accounting.

1.2 Hypotheses

The investigation is focused on assessing the adequacy of the following hypotheses as possible answers to the research problem:

Hypothesis 1: Current practices used in the building process of scientific production in accounting in Brazil do not fulfill the quality attributes of good research that are described in the literature;

Hypothesis 2: The Brazilian institutional environment contributes to the lesser quality of scientific production in accounting as a result of institutional failures that may impact research integrity.

It is assumed as the stated thesis that certain current practices of investigation used by Brazilian accounting researchers, combined with the local institutional environment, contribute to the lower quality scientific production in the field of accounting in Brazil.

1.3 Justification, Gap, and Contributions

The first subject is academic relevance. This study is justified by the importance of research production based on new approaches to and alternative views about its construction. This study can contribute to improving knowledge in accounting based on new methods of analysis. The comprehension and analysis of scientific production practices are relevant to evaluating research quality. Thus, this study can help redirect actions, redefine practical and

methodological strategies, and facilitate access to research funding in the field by illuminating how to achieve greater acceptance of academic production by the scientific community.

The Coordination for the Improvement of Higher Level Personnel (CAPES), a Brazilian governmental foundation of the Ministry of Education, is responsible for evaluating graduate programs in Brazil. For this purpose, CAPES has a system called “Qualis” to assess journals, and “books classification” to assess books, as part of its quality evaluation. The “Qualis” system formulates a grade (denominated “stratum”) by classifying each journal into scaled strata according to specific criteria based on formal features and impact. This mechanism helps classify and judge programs’ scientific production during their quadrennial evaluation process.

With respect to the criteria for evaluating journals, the Area Documents for 2012 and 2013 (CAPES, 2012, 2013) have presented a set of standards to assess the scientific production of faculty members in Brazilian graduate programs in administration, accounting, and tourism that are notably based on shape criteria. However, the second document indicates that homogeneity of form has ceased to be a quality differentiating criterion for classifying journals in the Qualis/CAPES strata, as discussed in the 2012 document. This difficulty in stratification has had practical consequences for assessing the journals; In particular, it led to framing some national journals relevant to the area in a stratum higher than their classification if the established criteria were followed. According to the Area Document (CAPES, 2013), those journals’ editors should make extra efforts to consolidate and include the periodical in databases that extend the impact of national production in their areas of interest. The need to find alternative and complementary ways to evaluate the quality of scientific production in graduate programs creates the opportunity to do this by evaluating aspects of the research process to increase the quality of scientific production during the stages of planning, implementing, and disseminating results.

Although CAPES is concerned with evaluating the quality of graduate programs (and to do so, also analyzes the quality of journals and books), it focuses on outcomes, i.e., publications. Thus, a gap emerges because the evaluation that uses indicators to measure and classify the scientific production of graduate programs is limited to informing programs and researchers about their failures in the research process. Therefore, it is necessary to engage in a systematic reflection about possible improvements in the process by which scientific knowledge is produced in graduate accounting programs in Brazil. The purpose of this reflection should be to offer ways to improve such practices by thinking about every step of

the research process in an integrated manner to address planning through publication and by taking quality criteria into account.

The second aspect is that of practical relevance. Studies focused on evaluating research practices may impact the field by helping researchers to identify weaknesses, limitations, and points to consider to improve the steps and methodological choices used to develop their studies. Each stage of the research process has a set of procedures that are designed to meet best scientific practices, to improve the possibility of obtaining research funding, and to have the chance to publish studies in journals with higher impact. Additionally, the aspect of practical relevance may present an opportunity to be both more productive and more notable in the field. The expectation is that the approach presented herein may be useful to increase the acceptability of Brazilian accounting research in the research community.

Obtaining a detailed understanding of the building process of scientific production in accounting contributes to the evaluation of research quality. As a result, this understanding may aid in the self-assessment process of what is produced. This self-assessment helps generate studies with greater adherence to scientific criteria that are generally recognized in the scientific community. As an effect, studies tend to receive greater acceptance in the submission process for publication in the official channels of scientific dissemination (Godoy, 2005). Self-assessment can also help shorten the submission process because it tends to diminish the number of required revisions for acceptance in journals. Furthermore, studies that are of higher quality, that make a clear contribution and that have an impact are taken into account in resource allocation, faculty recruitment, student enrollment decisions and accreditation, among other factors (Chan et al., 2012).

As with expected impact, quality improvement of the scientific production of graduate accounting programs contributes to increasing the likelihood that papers will be published in relevant international journals. Additionally, quality improvement may create new opportunities for research funding, for collaboration in international partnerships, and for increased relevance of the country's scientific output in the international scientific community.

1.4 Limitations of the Study

Scientific research requires the researcher to make numerous decisions and choices in order to make the study both feasible and scientifically acceptable. Brinberg and McGrath (1985) emphasize that the research process involves three conflicting desiderata, i.e.,

generalizability, precision, and realism, and that it is up to the researcher to consider the intensity of these factors in his or her design choices and in conducting the investigation. In effect, these options lead to study limitations because it is impossible to maximize all three characteristics. In the present study, some limitations were identified:

(a) The choice of the general criteria for the researcher because there may be discrepancies in various stakeholders' assessments of them;
(b) The risk of having biases due to the number of panelists;
(c) The limitations of the technique that was chosen, such as the lack of interaction between experts, which might lead to different results;
(d) the four stakeholder groups interested in the research process (researchers, bureaucrats – development institutions and the regulator –, academic journals editors, and readers) are not adequately represented in Delphi.

To mitigate the limitations of the choice of criteria, an association was made with similar standards used by other theoretical lenses that was based on the common element of connection between them. Regarding the risk of limitations because of the number of panelists, a decision was made to invite all faculty members in Brazilian graduate accounting programs to participate in the Delphi experiment. To minimize the possible effects of the lack of interaction among the participants, data collection tools helped induce suggestions for inclusion and change to increase reliability. Finally, to minimize the lack of an adequate representation of the interested groups, it was decided to identify the experts who participated from the panel and who develop activities representing these groups and explore their representation in the sample characterization.

.

CHAPTER 2 –
LITERATURE REVIEW

The purpose of this chapter is to provide an understanding of the main concepts that involve research quality. The chapter is organized into three sections. The first section presents the concepts of quality, attribute, and the research process. Additionally, it introduces the integration of the research process and scientific investigation domains (Brinberg & McGrath, 1985). Finally, it describes some quality elements such as relevance, validity, rigor, and integrity, along with a discussion of good research practices. The second section discusses how to evaluate the quality of the research process, considering not only research traditions but also stakeholders and their interests. Complementarily, the types of studies used to judge the quality of journals, along with a debate about criteria and indicators, are also presented. The third section, which discusses the quality of accounting research, covers prior studies of accounting research quality, the role of the regulator in graduate programs, and the quality criteria used to evaluate the steps in the research process.

2.1 What is Quality in the Research Context?

To properly understand the purpose of this study, it is important to be clear about the meaning of its central elements: quality, attribute and the research process. In a generic sense, quality is described as an attribute, a natural condition, a property by which something or someone is individualized that distinguishes it from others. Additionally, quality is defined as a way of being, essence, or nature. It can be further defined as the degree of perfection, precision, and compliance with a particular standard (Michaelis, 2012). In scientific research, Valentine (2009) notes that the answer to the question, "What are the characteristics of a high quality study?" depends, in part, on why the judgment is being made because different stakeholders attribute different meanings to their judgments. He argues that quality refers to an adequate link between the goals of the study and the design and purpose involved in implementing the study. Regarding accounting research, Clarkson (2012) argues as follows:

> Here, the notion of 'quality' speaks to the study's ability to 'inform the debate' which in turn is tied directly to three fundamental factors: (i) contribution – the importance of the study's focus and the extent of its innovation; (ii) the rigor with which the study has been conducted (scientific credibility); and (iii) the ability of the document itself to convey the study in a transparent and accessible fashion (communication).

These elements may guide an evaluation of research quality in various fields because they represent the general attributes of good research. Similar elements are presented by Webster and Watson (2002), whose study mentions the following questions associated with each of these elements: (a) contribution ('what's new?'); (b) impact ('so what?'); (c) logic ('why so?'); and (d) thoroughness ('well done?'). The first two aspects refer to a study's 'substance': i.e., its research problem, its foundations, the identified gap, its justification, and its possible effects on a field. The study's logic is associated not only with building a theoretical framework but also with choosing a research design. The last question involves the rigor and ethical aspects associated with the conduct of the research. Another perspective based on quantitative social research (Given, 2008) involves a judgment of the quality of research in two stages, as follows:

> For quantitative social researchers, given their empiricist and realist philosophical dispositions, there has been a general agreement that there is a two-stage process appropriate for judging the quality of research. First, judgments about good versus bad research are based on whether or not the researcher employed the proper methods; this judgment is then followed by a judgment as to the value of the findings in a practical and/or theoretical sense. (p. 752)

The concept of attribute also has a variety of meanings, including the following: (1) that which is proper and peculiar to someone or something; and (2) a condition, property, or quality of something. Grammatically, an attribute modifies a noun by expressing its quality or determination (Michaelis, 2012). In modern philosophy, attribute is used in the logical-grammatical sense predicate (Abbagnano, 2007).

Finally, the research process is defined as "the identification, combination, and use of elements and relations of the conceptual, methodological, and substantive domains" that is divided into three stages with different perspectives on validity (Brinberg & McGrath, 1985, p. 16). Another definition of the research process in social sciences considers philosophical aspects, notably, theoretical perspective and epistemology. Additionally, this definition includes practical elements that embrace both methods (techniques and procedures) and methodology (strategy) (Crotty, 1998). Furthermore, the research process may be described based on the following steps (or stages): findings, design, sample, data collection, analysis, and reporting (Mays & Pope, 2006; Spencer et al., 2003).

Given these definitions, the quality attributes in the research process described in this study include a set of characteristics that demonstrate links between the object of interest, the evidence, the explanatory theory and the researcher's methodological choices. Quality can

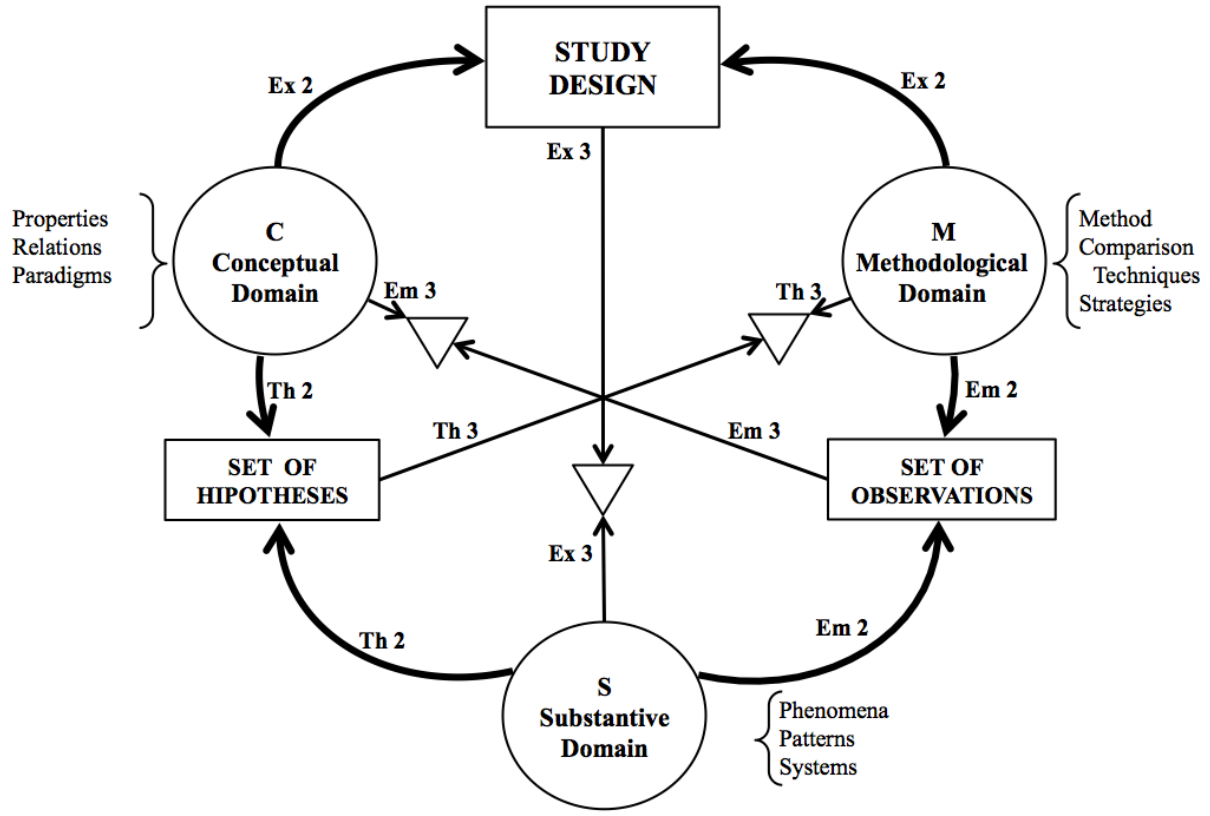
occur in various stages of research planning and execution, permeated by the underlying research tradition.

2.1.1 The integration of the research process and scientific investigation domains

A detailed understanding of the research process may have a positive effect on researchers' practices and help them develop better studies; delimiting the domains of scientific investigation represents a strategy to achieve this goal. As discussed above, Brinberg and McGrath (1985) describe research as a study of relations and present an analytical approach to the research process supported by the concept of validity in various stages of research. According the authors, "validity is not a commodity that can be purchased" but instead represents an ideal state. It is associated with "integrity, character, or quality, to be assessed relative to purposes and circumstances". This approach (or system) is called the Validity Network Schema (VNS). It is focused on the behavioral and social sciences and is organized into domains, levels, stages, and paths, with the following assumptions about each:

> (1) **Domains.** Research involves three interrelated but analytically distinct domains: the conceptual, the methodological and the substantive.
> (2) **Levels.** With respect to each of these domains, research involves relations between elements within an embedding system. The nature of the elements, relations, and embedding system differs among the domains.
> (3) **Stages.** The complete research process involves three major stages, each with several paths.
> (4) **Validity.** The concept of validity takes on fundamentally different meanings within each of the three stages.
> (5) **Paths.** There are three alternative paths for carrying out the central stage of the research process. These paths reflect different styles of doing research and encounter different validity issues. (Brinberg & McGrath, 1985, p. 15-16)

Brinberg and McGrath illustrate the VNS system in a scheme that involves the connections between the domains in research design and their relationships with the sets of observations and hypotheses (Figure 1).



EXPERIMENTAL PATH (Ex):	Building a design, and implementing it by using on a set of substantive events.
THEORETICAL PATH:	Building a set of hypotheses, and testing them by evaluating them with an appropriate set of methods.
EMPIRICAL PATH:	Building a set of observations, and explaining them by construing them in terms of a set of meaningful concepts.

Figure 1: The VNS System: domains, levels, and paths.
Font: Brinberg and McGrath (1985, p. 22). (Adapted).

In this integrated approach, the scientific investigation contains elements and relations of three basic domains (Brinberg, 1982):

(a) The conceptual domain, which includes concepts and relations considered in an abstract way;

(b) The methodological domain, which encompasses tools and techniques for obtaining observations and correlating the set of observations; and

(c) The substantive domain, which embraces events, processes, and phenomena in the real world.

According to the authors, any research design should contain elements and relations of each one of these domains, which vary based on their central object, as summarized in Table 1.

Table 1:
Domains and levels of the VNS

Domain	Levels		
	Element	**Relation**	**Embedding System**
Substantive	Phenomena – states and actions of entities (agents and objects, human or nonhuman).	Patterns of phenomena – relation between two or more states and actions of entities; the process of interest.	Substantive systems – temporal, locational and situational context within which entities are embedded.
Conceptual	Attributes or properties of phenomena – ideas and concepts to describe and explain a phenomenon.	Conceptual relations – temporal, logical, and causal pattern specified among two or more concepts.	Conceptual paradigm – the set of philosophical assumptions within witch the concepts and their relations are embedded.
Methodological	Methods or modes of treatment – methods for measuring, manipulating, and controlling variables (properties of phenomena) to gather information about the phenomena.	Comparison techniques – to make comparison or assess covariation among the variables.	Research strategies in which the elements and relations are embedded.

Source: Adapted from Brinberg and McGrath (1985).

Each domain has its own perspective. Whereas the phenomenon is the central object in the substantive domain, the conceptual domain involves identifying concepts and attributes that can contribute to an explanation of the phenomenon studied. In other words, the purpose of the conceptual domain is to identify a theory (or theoretical lens) that can underpin the analysis of the observations regarding a phenomenon. The methodological domain presents and describes the means for accomplishing the study. In essence, the connection between the object of interest (represented by the substantive domain) and the conceptual relationship, denoted by its properties (theory) leads to a set of hypotheses being generated. As a result, the set of hypotheses might explain the phenomenon. For its part, the connection between the object and the methodological domain relates to the strategies used for the systematic collection of a set of observations that are supported by the evidence and support the findings.

Brinberg and McGrath (1985) argue that the relations level is the main focus of the research process. In support of this position, they argue as follows: relations in the conceptual domain are symbolic representations (abstractions); relations in the methodological domain are procedures; and relations in the substantive domain are processes of interest. In this sense, the materials in each domain serve different functions: the conceptual domain specifies levels; the methodological domain distinguishes among levels of features of the relations; and the substantive domain displays whichever level of features of relations can occur in the "real-world" substantive system.

The VNS system presents the research process organized into three stages, and for each, the validity is addressed from a distinct perspective (Table 2). Stage One, also called the

"Pre-study or Preparatory Stage," considers validity as value (or worth) and "involves generation, development, clarification, and evaluation of elements and relations within each one of the three domains". This stage should be performed before the two following stages and reflects the generative or constructionist paradigm in the research process.

Table 2:
The VNS system: validities and stages of the research process.

Stage One: Validity as Value

(Central tasks of the stage one: Identification, development; and clarification of elements, relations, and embedding systems, for each of the three domains).

Domain	Criteria for Evaluating Elements, Relations
Conceptual (C)	Parsimony, internal consistency, subsumptive power, testability, etc.
Methodological (M)	Efficiency, power, unbiasedness, explicitness, reproducibility, etc.
Substantive (S)	System effectiveness, cost/benefit, feasibility, etc.

Stage Two: Validity as Correspondence

(Central tasks of stage two: selection, combination, and use of elements and relations from all three domains to produce a set of empirical findings).

Paths	*Step 2*	*Step 3*	*Product*
Experimental (Ex)	Study design	Implementation	A set of empirical findings
Theoretical (Th)	Set of hypotheses	Test of hypotheses	
Empirical (Em)	Set of observations	Interpretation	

Stage Three: validity as Robustness

(Central tasks of stage three: verification extension; and delineation of particular stage two findings.)

Replication:	Are the (stage two) findings reproduced when all facets of C, M and S are kept the same?
Convergence Analysis:	Over what range (of values of all the facets of C, M and S) do the (stage two) findings hold?
Boundary search:	Beyond what range (of values of all the facets of C, M and S) do the (stage two) findings fail to hold?

Source: Brinberg and McGrath (1985, p. 23).

According to Brinberg and McGrath (1985) Stage Two or the Central Stage consists of conducting the research, which "involves developing a set of empirical findings by combining subsets of elements and relationships of each of the three domains." It reflects the logical empiricist paradigm or hypothetical-deductive. They divide this stage into three interrelated steps:

1. Choosing the elements and relations of one of the domains, the central element of the researcher's interest;
2. Combining the elements and relations of this domain with the elements and relations of a second domain to form an intermediate or instrumental structure; and
3. Introducing the elements and relations from a third domain inside this structure to generate a set of empirical findings.

When a researcher chooses the main domain from his/her perspective to begin a study and adds a second domain to establish the instrumental structure, he/she is choosing a particular path through which to conduct the study. For instance, there are three paths in the VNS system: experimental, empirical, and theoretical. The "experimental path" is employed when the researcher combines the methodological and conceptual domains (Step 1), which results in an instrumental structure known as study design (Step 2). In the next step, the researcher chooses a substantive phenomenon that enables the research design to be applied to produce empirical findings (implementation). Likewise, the "theoretical path" presents the combination between the conceptual and substantive domains. In turn, the "empirical path" mixes the methodological and substantive domains in the first step. Both add the third domain to reach a set of empirical findings. In summary, this stage refers to the research itself, in which the three domains are combined and the options (paths) are defined in the research design in search of results (Sakamoto, 2011), i.e., the research implementation.

Stage Three involves monitoring the Stage Two findings. Brinberg and McGrath (1985) note that this occurs "by replication and by a systematic search for both the range and the boundaries of those findings" and reflects the paradigm of credibility or generalization. The relevance of this stage is to identify the boundaries in which the study can expand knowledge – in other words, what is the study's effective contribution?

Another perspective on realizing the research process is presented in Spencer, Ritchie, Lewis, and Dillon (2003) and Mays and Pope (2006), whose framework is oriented to evaluate evidence toward nine key features and processes:

(1) findings – this characteristic is associated with judgment on elements as the credibility of the findings, how knowledge is expanded by the study, detailing the scope of inference, and the fulfillment of aims and original purposes;
(2) design – justification for research design (methodological choices);
(3) sample – criteria evaluation for the design and selection of the sample, as well as inclusion and exclusion cases;
(4) data collection – evaluation of the conduction of data collection;
(5) analysis –factors such as depth and complexity of data, choice of approach and formulation analysis, context of data sources, and diversity of perspectives;
(6) reporting – assessment of links between data, interpretations, and conclusions, as well as the overall coherence of the report;
(7) reflexivity and neutrality – clarity on the assumptions, theoretical perspectives, and values that guide the study, and care of errors and biases;

(8) ethics – assessment of how the researcher and his team deal with ethical issues;
(9) auditability – a judgment on compliance with formal procedures and documentation of the study conducting process for future inspections.

This perspective is based on a qualitative inquiry and might be understood as another way to examine the research based on large elements (or aspects). Although findings are not described as a step of the research process, the authors argue that they should be assessed beforehand. In the next step, the authors group elements 2-6 as the various stages of the research process. Finally, although the key features do not belong to step groups, the last three key features represent the general characteristics of the conduct of research.

The research process contains three mutually incompatible desiderata: generalizability, precision and realism (Brinberg & McGrath, 1985). For these authors, generalizability is related to the population to which the information applies; precision is related to the measurement and control of the variables involved; and realism "has to do with the contexts, or concrete behaviors system, to which that information is intended to apply." The authors highlighted the combination of these elements, producing limitations that do not emanate from the choices, preferences, or insufficient resources but that are "inherent in the research process itself." Conversely, steps and key features are judged by criteria and quality questions that will be treated as their own topic.

2.1.2 Relevance, validity, rigor, and integrity as elements of research quality

Research conducted in the university environment must simultaneously exhibit quality and relevance; otherwise it will not be justified. To be relevant, research must have quality; however, the converse is not true because relevance may mean different things to different people (Schwartzman, 1988). Schwartzman argues that the quality of research depends not only on a thorough knowledge of the scientific field but also on an innovative contribution. In practice, the scientific community can only measure this quality by means of its various action forums (journals, conferences, evaluation committees, etc.). Therefore, in practice, there is no quality research without adequate peer-review mechanisms. However, the relevance of research depends not only on its quality but also on its eventual applications to scientific, educational or applied purposes. Finally, Schwartzmann adds that evaluating a study's scientific relevance depends on the scientists themselves; reviews of other forms of relevance require the participation of other interested sectors. One aspect to consider is the importance of peer review as an element of research quality. In a sense, a study judged and

accepted by evaluators who dominate the topic becomes validated in academia. For this reason, the creation of more objective criteria or principles to guide the peer review process is an important task for the scientific community.

Validity might assume a variety of concepts and meanings in the social sciences, including "convergence, correspondence, differentiation, equivalence, generality, repeatability, and some others" (Brinberg & McGrath, 1985). However, these meanings make sense in different steps (or stages) of the research process. In the first stage, which involves planning, validity is associated with the concept of value as merit, "value," or relevance. In the second stage, when the research is executed, validity is associated with correspondence, i.e., with the internal validity of designs (for example, the internal validity proposed in experimental studies). Finally, in the third stage, whose purpose is to observe the findings, validity is considered in terms of the evidence related to the robustness of the findings. In other words, its purpose is to check, extend and delimit the earlier stage results, preceding the action proposed as part of the existing body of knowledge. In this phase, validity as robustness is consistent with external validity in experimental studies (Brinberg & McGrath, 1985; Brinberg, 1982; McGrath & Brinberg, 1983; Sakamoto, 2011)

Certain key elements may adequately represent rigor, which has particular importance in the following situations: (a) when the results to be evaluated are complex, difficult to observe or consist of many elements that react in various ways; (b) when the decisions that follow are significant and expensive; and (c) when evidence is required to convince others about the validity of one's conclusions (Weiss, 1997). Other characteristics involve full and faithful recording as an element of data quality and a clear account of the analytical procedure. Additionally, a study should be presented as a careful documentation of the research process that enables both auditability (audit trail) and an evaluation of the researcher's skills, which is particularly important in qualitative research (Spencer et al., 2003). To sum up, it involves good practices in data collection, analysis, and transparency. However, the perception of rigor is diverse, non-consensual, and may depend on the research method adopted by the researcher. For instance, Davies and Dodd (2002) argue as follows:

> If rigour is understood only in terms of a structured, measurable, systemized, ordered, uniform and neutral approach, then other research methods, that allow flexibility, contradictions, incompleteness, or values will always appear 'sloppy', epitomizing everything that is 'nonrigour' and therefore lacking credibility (Davies & Dodd, 2002, p. 280).

A good example to illustrate how the absence of rigor may compromise the quality of a study is presented by André (2001), who focuses on the education field; her work is based

on previous studies in which she cites the lack of knowledge about assumptions and research methods and techniques. She notes problems in that field's research process, such as methodological weaknesses of studies and surveys resulting from their use of very reduced portions of reality. Additionally, she mentions the limited number of observations and subjects, the use of poor instruments in opinion surveys, poorly founded analyses, and interpretations without theoretical support.

Research integrity reveals a set of principles that direct researchers' actions. These principles present rules of conduct or codes of practice that people and organizations involved in scientific research should follow: (a) honesty in communication; (b) reliability in performing research; (c) objectivity; (d) impartiality and independence; (e) openness and accessibility; (f) the duty of care; (g) fairness in providing references and giving credit; and (h) responsibility to future scientists and researchers (European Science Foundation, 2011). Integrity is an important element of quality because research conducted without high ethical standards mars the quality of the findings and poses a risk to those involved. For instance, the Belmont Report (National Institutes of Health, 1979) argues that research involving human subjects must guarantee three basic principles: respect for persons, beneficence, and justice. Another perspective with respect to integrity is associated with responsive research, whose principles are honesty, accuracy, efficiency, and objectivity (Steneck, 2007). In summary, if research practices do not observe these core principles, then the research itself does not comply with the minimum requirements of honesty and quality. In a practical sense, failure to observe the principles can be evidence of research misconduct.

2.1.3 Good research practices

Given its subjective and evaluative character, the definition of good research is neither easy to define nor consensual. Shamoo and Resnik (2003) argue that people in leadership positions in research can play a key role in developing a culture in which ethical attitudes and good research practices prevail. These authors argue that if principal investigators, managers, companies and government agencies demonstrate and tolerate unethical attitudes and poor research practices, then those attitudes and practices will prevail. For Shamoo and Resnik, the research culture (attitudes and behaviors) sets the tone for the importance of the objectivity, quality and integrity of one's data and results. Their perception of good research practices is that such practices represent rules that researchers can follow to help ensure the quality, objectivity, and integrity of their data.

Another view that might be helpful in comprehending the elements or attributes of good research is described as a set of managerial dimensions by Cooper and Schindler (2003): (1) clearly defined purpose; (2) detailed research process; (3) full research planning; (4) the application of high ethical standards; (5) frankly admitting evidence limitations; (6) analysis adequate to decision-makers' needs; (7) results presented in a non-ambiguous way; (8) justified conclusions; and (9) honest reflections on the research experience.

The expression "horses for courses" is used to illustrate the meaning of good social research (Denscombe, 2010). In general terms, it can be understood as the most appropriate course/path; in other words, some strategies are more suitable than others to answer certain research questions. Denscombe presents three elements to consider:

(1) Approaches are selected because they are appropriate for particular types of investigations and problems. Depending on the nature of the research problems, there are paths to be chosen that reveal themselves as more appropriate for a particular study. Such paths allow the collection of evidence that is more robust than the alternatives.

(2) "Strategic" decisions aim to place the social researcher in the best possible position to achieve the inquiry's best outcomes. The choice of the path best suited to analyze the phenomenon or object of study is intended to provide the researcher with an insider's view to obtain more appropriate results.

(3) Good research choices are (a) reasonable and (b) are an explicit part of every research report. Good research requests disclose the researcher's decisions related to paths and object analysis techniques, and those decisions are both clearly and duly justified.

The notion of good research is associated with rules that investigators can follow to help them guarantee data quality, objectivity, and integrity. To do so, it is necessary to include reasonable and explicit choices as part of every research component (Denscombe, 2010; Shamoo & Resnik, 2003). In this context, rigor and quality exhibit a strong relationship because it is not possible to obtain high-quality research if the process is poorly executed. Thus, the Sao Paulo Research Foundation (FAPESP)[1] is a governmental development institution that has published its Code of Good Scientific Practice (FAPESP, 2012a, 2012b) in order to guide researchers and institutions in the conduct of activity and scientific research, characterized as (FAPESP, 2012a):

[1] FAPESP is the development institution of the State of Sao Paulo, Brazil. It is one of the most significant development agencies in Brazil.

> Scientific activity is herein understood to be any activity that directly aims to conceive and conduct scientific research, communicate the results thereof, encourage scientific interaction among researchers and mentor or supervise researchers in training processes.
>
> Scientific research is herein understood to be any original investigation that aims to contribute to the constitution of a science. A science is understood to be any body of rationally systematized and justified knowledge obtained through the methodical use of observation, experimentation, and reasoning. This broad definition applies to the exact, natural and human sciences, technological disciplines and disciplines usually included among the so-called humanities (p. 9).

The Scientific Director's letter presenting good scientific practices to the Foundation's board of directors (FAPESP, 2011), indicates an intention to form a solid, deep-seated culture embodying ethical integrity in research. This best-practices policy aims to achieve this goal through a series of sustained action strategies based on three interdependent foundations: (1) education; (2) prevention; and (3) fair and strict sanction and investigation. As a result, the Foundation has published an ordinance (FAPESP, 2013) to establish criteria for publicizing cases that have violated good research practices. The summaries of the first nine cases are publicly available on the Foundation's website (http://www.fapesp.br/8577).

2.2 How Should We Evaluate Quality in the Research Process?

People are always evaluating something. The theoretical basis of evaluation helps us understand the effects of evaluation, along with the focus that involves its various purposes. For instance, a government policy, a government program, or a nonprofit entity can be the object of evaluation. Weiss (1997) presents the concept of evaluation as follows:

> Evaluation is the *systematic assessment* of the *operation* and/or the *outcomes* of a program or policy, compared to a set of *explicit* or *implicit standards*, as a means of contributing to the *improvement* of the program or policy.

Weiss (1997) adopts a general view that calls for attention to evaluation as a social practice involving two foci: process and outcome. Process describes how the program/policy is implemented. The outcome is described as an end (or result) for its audience and in this context, the concept of impact as the net effect of a program or policy on the audience is important. In an outcome evaluation, the central aspect is variance. Another method treats foci as models, e.g., the process model and the variance model (Van de Ven & Poole, 2005; Webster & Watson, 2002), with meanings similar to those presented by Weiss.

These highlighted keywords in the evaluation concept refer to main aspects that cover a broad comprehension of evaluation. First, a systematic assessment is related to the step-by-step building of a description of the program or policy, e.g., whether it is analyzed with a

process or a variance evaluation. Weiss, (1997) argues that the evaluator needs must consider whether there is a constant interaction among questions, methods, and data sources.

2.2.1 Influence of research traditions on the research process

The literature addresses building knowledge – including scientific knowledge – based on various views. A study exploring educational research raised issues about the scientific research that may guide questions about this production based on three major points. First, what characterizes a scientific study? What is the relationship between scientific and other types of knowledge? These are questions concerning research purposes and the nature of the knowledge produced. Second, how can we judge what is good research? Who defines such criteria? These questions involve the criteria for assessing the quality of scientific work. Third, what procedures must be followed to maintain rigor in the collection and analysis of data? These questions relate to the assumptions of research methods and techniques, both in situations focusing on local issues and in situations addressing a large number of observations (André, 2001). In discussing research quality, the second and third questions are particularly important because they explore two concepts analyzed in that paper: rigor and quality of the knowledge produced.

Nevertheless, judging research quality requires that the particular assumptions and perspectives of each type of research should be considered. Thus, in a study on implementing and evaluating critical interpretive research, Pozzebon (2004) argues that the existence of different paradigms (or research traditions) leads to different ways of perceiving scientific evidence. As a result, this author posits that it is inappropriate to use the criteria of one research tradition for evaluating research quality driven by another paradigm. Nevertheless, there is no consensus on this issue, and criteria and indicators are analyzed as a stand-alone topic below.

The literature presents distinct modes of fitting studies into different paradigms (or research traditions) based on their idiosyncratic characteristics. One of the most classical modes presents four paradigms that compete in qualitative research: positivism, post-positivism, critical theory and constructivism (Guba & Lincoln, 1994). The authors of this latter study conceptualize the paradigm as a basic belief system or worldview that guides the investigator – not only in the choice of the method but also in fundamental ontological and epistemological paths. These authors emphasize that methodological questions are secondary

to the choice of paradigm; in other words, the paradigmatic view, including its ontology and epistemology, defines the method to be chosen to answer the research questions.

To underline the different views in each research tradition in the business field, Gephart (1999) presents another classification using three groups: positivism, interpretivism, and critical theory/postmodernism. A comparison of these groups considers the following elements: (a) assumptions; (b) key focus or ideas; (c) key theories in the paradigm; (d) key figures (main authors); (e) the goal of paradigms; (f) the nature of knowledge or the form of theory; (g) criteria for assessing research; (h) the unit of analysis; and (i) the research methods and type(s) of analysis. The study by Gephart (1999) is important because it offers criteria for judging research within each paradigm. For positivist studies, the author notes that the main aspect in this tradition is prediction = explanation, and the main criteria are rigor, internal/external validity, and reliability. Conversely, interpretivism focuses on trustworthiness and authenticity. Finally, judgment in the context of critical theory/postmodernist research is based on theoretical consistency, and the criteria include historical insights, transcendent interpretations, and change potential and mobilization.

Attempting to outline these various classifications, Rynes and Gephart (2004) summarize the research traditions didactically based on views of different elements (Table 3).

Table 3:
Research traditions

Tradition	Positivism and Postpositivism	Interpretive Research	Critical Postmodernism
Assumptions about reality	Realism: Objective reality that can be understood by mirror of science: definitive/probabilistic	Relativism: Local intersubjective realities composed from subjective and objective meanings: represented with concepts of actors	Historical realism: Material/symbolic reality shaped by values and crystallizes over time
Goal	Discover truth	Describe meanings, understanding	Uncover hidden interests and contradictions: critique, transformation, and emancipation
Tasks	Undertake explanation and control of variables: discern verified hypotheses or nonfalsified hypotheses	Produce descriptions of members' meanings and definitions of situation: understand reality construction	Develop structural or historical insights that reveal contradictions and allow emancipation, spaces for silenced voices
Unit of analysis	Variable	Verbal or nonverbal action	Contradictions, critical incidents, signs and symbols
Methods focus	Uncover facts, compare these to hypotheses or propositions	Recover and understand situated meanings, systematic divergences in meaning	Understand historical evolution of meanings, material practices, contradictions, inequalities

Source: Rynes and Gephart (2004). *Note:* According to the authors, this table is based on Gephart (1999), Guba and Lincoln (1994), and Lincoln and Guba (2000).

The assumptions shown in Table 3 represent the worldview of the researcher and are consistent with his/her adopted paradigm (or research tradition). The definition of paradigm is "… an integrated set of assumptions about the nature of the social world, about the character of the knowledge we can have about the social world, and about what is important to know." (Greene, 2007, p. 15). Each paradigm or research tradition brings both its respective worldview and its ontological and epistemological assumptions (Pozzebon, 2004). For that reason, the focus of research varies, as summarized by Rynes and Gephart (2004) in Table 3. For instance, in the positivist and post-positivist traditions, the variable is the main unit of analysis, and comprehension occurs when hypotheses are measured and tested (positive view), or trying to falsify them. Consequently, the likelihood obtained in the relationship between independent and dependent variables is considered an adequate measure to explain the phenomena, as it considers quantitative features, and allows generalization. Conversely, the tasks of interpretative research are directed to identify meaning and definitions (and sometimes occurrence patterns); however, generalization is less important in this worldview.

Although both positivism and post-positivism share common features, they have conceptual differences. Gephart (1999) argues that post-positivism is a recent evolution of positivism. Although post-positivism is consistent with positivism in assuming that the objective world exists, post-positivism nevertheless assumes that "the world might not be readily apprehended and that variable relations or facts might be only probabilistic, not deterministic". In this case, post-positivism focuses on falsification instead of verification, given the complexity of real-world phenomena. Whereas positivism uses experimental and quantitative methods, post-positivism attempts to include qualitative elements "to gather broader information outside of readily measured variables".

Finally, Spencer et al. (2003) compare the paradigms that consider quantitative and qualitative approaches (Table 4), although they recognize that these are labels that represent not only paradigms but also research approaches, some types of research methods, and philosophical assumptions.

Table 4:
'Quantitative' and 'qualitative' research paradigms

Quantitative	**Qualitative**
Scientific	Naturalistic
Positivist	Interpretivist/hermeneutic
Realist	Idealistic/relativist/constructivist
Objectivist/materialist	Subjectivist
Foundational	Fallibilistic/anti-foundational
Experimental	Ethnographic

Source: Spencer et al. (2003, p. 45.)

Given the characteristics of current accounting research in Brazil, it might be effectively argued that positivist and post-positivist traditions predominate. There are many theoretical and empirical studies involving accounting and market variables. Additionally, graduate business programs' educational model is powerfully influenced by the American model of business and accounting research, which focuses on market issues.

2.2.2 Stakeholders and their interests

By studying quality from the perspective of qualitative research, Flick (2009) argues it might be argued that the question about research quality (qualitative) can be addressed at four different levels involving four different groups of actors:

1. Researchers' interest in knowing why research is good or bad;
2. Development institutions' interest in assessing what should be or previously has been funded;
3. Academic journal editors' interest in deciding what and what not to publish; and
4. Readers of research guidelines' interest in which research to trust and which research not to trust.

Based on a literature review, one study has listed four stakeholder groups, which are referred to as parts of publishing process (Adler & Liyanarachchi, 2011; Moizer, 2009): (1) editors; (2) reviewers; (3) authors; and (4) development institutions (bureaucrats). Although not explicitly mentioned, in Brazil, the graduate regulator (CAPES) might also be included in the category of bureaucrats because its policies influence research in various fields of knowledge (including administration, accounting, and tourism).

Stakeholders interested in knowledge have different aims when seeking to identify what constitutes the meritorious aspects of research, and their views are not restricted to the qualitative perspective. Beyond the particular use of each group presented by Flick, it is important to consider that stakeholders' interests are not limited to particular types of or approaches to research. However, each group may have different expectations about what to evaluate in an investigation, policy, or program (Adler & Liyanarachchi, 2011; André, 2001; J. G. Greene, 1988; Moizer, 2009; Rossi, Lipsey, & Freeman, 2004; Stufflebeam, 2000; Weiss, 1997; Yarbrough, Shulha, Hopson, & Caruthers, 2011). As a result, regardless of the paradigm or predominant approach, stakeholder groups may be large and share common interests among members that belong to the above-mentioned groups (Flick, 2009; Moizer,

2009). Even if each group's focus on judging quality is different, shared concern is possible based on assumptions adopted by these distinct worldviews (or research traditions).

Independent of the stakeholders of a study, rigor is essential to conduct research carefully to offer the best evidence that can fulfill stakeholders' particular needs, particularly concerning qualitative inquiries, in which the researcher/agent plays a key role. Research in education, for instance, has received some criticism involving the space between theory and practice (i.e., actors and researchers' roles, which are difficult to reconcile in science or intervention policy) and rigor helps address these issues (André, 2001). For André,[2] there are many relevant questions that have not yet been formulated and other issues that must be discovered and/or resolved; therefore, there are many opportunities for all types of research. However, rigor requires careful systematization and data control; properly planned research work; data collection through strict procedures; and dense and well-founded analysis. The report clearly describes the methods and the results obtained. Rigor cannot be disregarded in any type of research or evaluation to diminish those risks.

With respect to its bureaucratic role as the regulatory agent, CAPES establishes policies that delineate requirements and rules to evaluate graduate programs in Brazil. For example, a minimum number of scientific production "points" (based on a scale) are required for the faculty members of graduate programs. Additionally, these policies set standards to classify faculty members' scientific production (i.e., published papers and books) (CAPES, 2012, 2013) and to compare and evaluate the quality of graduate programs. Until 2012, this evaluation was carried out every three years, and the next evaluation will cover four years.

Two other interested groups deserve attention in the evaluation of the research process: the first consists of graduate programs, which are responsible for training new researchers and that care about assessing the quality of what is produced by its graduates and faculty. Evaluating research is one way that this group helps create high-level researchers that generate studies with high academic and social impact. The second group involves the development institutions, which are interested in determining which studies should be financed. These institutions play an important role because they provide scholarships, funds for the acquisition of research infrastructure, and resources to facilitate partnerships among institutions to increase the scope and relevance of their studies. Their role, particularly in the

[2] São tantas as perguntas relevantes que ainda não foram formuladas, tantas as problemáticas que ainda precisamos conhecer, que sobram espaços para todo tipo de investigação, desde que se cuide da sistematização e controle dos dados. Que o trabalho de pesquisa seja devidamente planejado, que os dados sejam coletados mediante procedimentos rigorosos, que a análise seja densa e fundamentada e que o relatório descreva claramente o processo seguido e os resultados alcançados. (André, 2001, p. 57)

Brazilian context, is important both because there is not as much governmental research funding available and because there is competition among many research fields.

In summary, research stakeholders have different interests based on their position or the group to which they belong. Just as in the evaluation practices, it is up to the researcher not only to conduct its investigations with the care necessary to offer these stakeholders consistent results but also to consider the audience during planning, executing and disseminating the results of research. These steps summarize a conception of what might be discussed when evaluating research based on an epistemological view, i.e., in monitoring its quality.

2.2.3 Judging the research quality of journals: perception studies based on surveys and citation-based studies

Evaluating the various stages of research is a task involving multiple stakeholders. On the one hand, development institutions are (typically) initially interested in the results of the planning phase (the project itself) that identifies criteria such as relevance, impact, and feasibility to define what should be funded. On the other, the researcher and his team are concerned with observing the quality attributes that enable them to systematically and rigorously lead the study, i.e., the execution of the research to obtain the best evidence. Finally, other stakeholders focus on evaluating the quality of the 'research product', i.e., scientific publications, which select the best studies. The academic community has used surveys of scientific community members (e.g., referees, editors, researchers) to identify their quality perceptions of certain journals and measures based on the citations of these studies (the resonance of scientific production among the peers).

Surveys that evaluate research quality based on perception assess a journal by considering the opinion of the groups that form the scientific community, which may be undertaken in many fields. For example, surveys are used to judge quality in fields such as the humanities (Hug, Ochsner, & Daniel, 2013) and accounting (Brinn et al., 2001; Lowe & Locke, 2005; Lowensohn & Samelson, 2006; Northcott & Linacre, 2010). Another method is to judge the quality of evidence using surveys as the technique for the studies' data collection (Van der Stede, Young, & Chen, 2005). Additionally, some contextual factors affect perceptions about quality, such as the researcher's geographic origin, research orientation, and the respondent's affiliation with a journal (Ballas & Theoharakis, 2003). Geographic origin is a factor that can explain respondents' preferences for a journal (or a set of journals) based on

the region in which they live. As a result, American researchers have a list of "best quality" journals that is different from the list made by European researchers of their own popular periodicals. Similar results can be observed among periodicals with different research orientations or approaches (focusing on qualitative or quantitative studies, for example) or that are either specialized or generic. Finally, affiliation is relevant because researchers that publish in a journal (or who act as reviewers or members of that journal's editorial board) tend to assess that journal as having the best quality. These types of survey-based evaluations are also referred to as peer reviews.

Lowe and Locke (2005) note that citations-based studies consider the number of citations received by a journal's papers during a given period. The quantity of citations generates an Impact Factor (IF) (Amin & Mabe, 2007; Garfield, 1972, 1999). The concept of the IF has received criticism because citations in some fields cannot be fairly measured just two years after publication, which is the time used (duration) to calculate IF (Le Pair, 1995). Nevertheless, perception studies based on surveys consider quality criteria chosen by researchers; these criteria can vary according to the research tradition in which the research is conducted. In a practical sense, it is inappropriate to use quality criteria for a specific tradition (positivism, for instance) to judge scientific knowledge constructed in a different tradition (like interpretativism) because the assumptions of the two traditions are different (Pozzebon, 2004).

Scientometrics is a field of study that uses measures and indexes to evaluate journals. It has two main applications: (1) evaluating the scientific literature for the distribution of research funds and (2) developing quality criteria to guide readers in selecting the best scientific evidence (Pinto, 2008). The first application has been widely used by development agencies and similar agencies that seek objective criteria (measures) to define which projects should be supported insofar as their financing activities are concerned. The second aims to identify the most scientifically relevant publications based on certain criteria, such as the journals' IF, published annually by the Institute for Scientific Information (ISI).

The ISI has created the Science Citation Index (SCI), which was originally presented in a study published in the early 1970s based on an analysis of approximately one million citations received by articles published in 2,200 global multidisciplinary journals in 1969 (Garfield, 1972). Thus, SCI began in order to provide periodic evaluation metrics according to their IF, which prompted researchers to submit the results of their studies to be published in the indexed journals considered the most relevant by the scientific community. Journals indexed by SCI began to have their IF calculated and reported by the ISI in the Journal

Citation Reports (JCR) (Garfield, 1999). The graphical representation in Figure 2 summarizes the ISI IF model (Amin & Mabe, 2007):

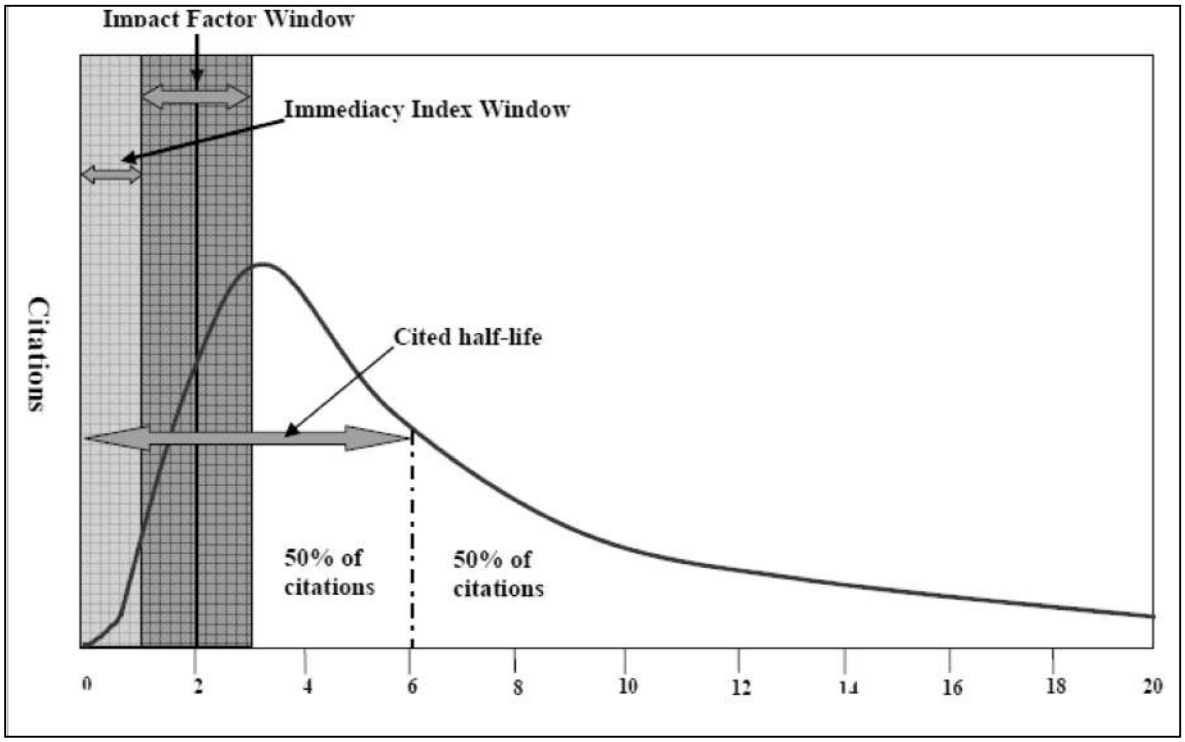


Figure 2: Generalized citation curve.
Source: Amin and Mabe (2007).

The immediacy index window is equivalent to the first year of an article's publication and is discarded in estimating the IF because it was considered too recent for the scientific community to have had access to and to use the text in subsequent studies. The Impact Factor Window considers citations received in the two years following publication; according to Amin and Mabe, this is the window during which articles receive the highest number of citations. According to the IF model, 50% of the total citations received by an article occur in the first six years after the publication, and the peak of the curve occurs at the end of the third year. Thus Garfield (1972) calculates the impact of a journal, as shown in Figure 3:

FI (Year 3) =	$\frac{\text{Total of citations received in Year 3 \# papers published in Years 1 and 2}}{\text{Total of papers published in Years 1 and 2}}$

Figure 3: Model for calculating IF.
Adapted from Garfield (1972).

Nevertheless, even presenting objective criteria for judgment, the IF model has been criticized for its lack of comparability and its failure to adequately portray the characteristics of articles in some fields. For example, it is inadequate to compare journals from different

fields because papers in areas of larger interest – which attract more researchers and readers – tend to receive more citations; this does not necessarily mean that a study in such a field is of higher quality than studies in fields with few investigators and/or studies whose audience is limited (Pinto, 2008). Alternatively, Le Pair (1995) argues that although the SCI tool is interesting, the view expressed by peers as the most relevant texts differ considerably from the citation count when analyzing articles in the fields of thermonuclear research and materials science, for example. Moreover, there is a gap in the citations of applied scientific works, such as technological works, which indicates a lack of bibliometric recognition because indicators based on citations are linked to a more conventional notion of books and scientific papers as the main sources of research results. Texts produced in fields with new discoveries, such as technological research, are thus frequently not recognized by the IF. Moreover, Le Pair (1995) claims to have serious doubts about the general validity of the IF method as applied to all areas. Finally, two additional criticisms are merit highlighting: by analogy, similar situations may occur with studies considered seminal in other fields of knowledge in which most citations tend to occur after the first six years of the publication. Such works would not be adequately recognized in the FI model. Another aspect occurs when the IF is calculated only for English-language journals based on Thomson Reuters databases. In effect, there are relevant non-English language journals, which will result in similar limitations to the calculation of impact. However, the parameters used in the index do not lose their usefulness as a result of evaluation criteria based on citations by peers.

Another mechanism that assesses research quality by considering citations is the H-Index. This measure is primarily used to evaluate researchers and reflects both numbers of publications and citations (Pinto, 2008). For example, a scientist with an H-index of 8 has published eight articles, each of which has been cited at least eight times (like a matrix). The H-index has become popular since the advent of the tool that allows us to create a researcher profile on the Google Scholar platform. Each researcher may create his/her personal profile, validate it with an institutional e-mail account, and register all of his/her publications. After it is created, the profile can be made public. The platform will periodically identify the citations of those studies and will show the H-index of the researcher in the profile automatically. The primary advantage of this platform is that it identifies citations regardless of their language. As a result, there is a quality measure with a greater range than the IF when considering the analysis of researcher productivity (Amara, Landry, & Halilem, 2015; Amara & Landry, 2012; Aragão et al., 2014).

To summarize, both the IF and the H-Index are quantitative measures that are also representations of quality by presenting scientific production's resonance among peers.

2.2.4 Criteria and indicators

According to the definition of the Joint Committee for Standards for Educational Evaluation, a criterion is "a standard by which something can be judged." (Yarbrough et al., 2011). Another definition involves the dimension of merit, i.e., if something is valuable or not or if it is good or bad (Davidson, 2005). A third concept describes criteria as a set of standards that define acceptability (Brinberg & McGrath, 1985, p. 41). Finally, another possibility is to refer to a 'characterizing trait' or 'a standard against which to make a judgment' (Smith, 1984, p. 383) , although with certain features, these conceptions generally designate criteria as standards. However, each evaluator's judgment tends to be different based on his or her purpose. According to Weiss (1997), evaluators use different gauges in the evaluation process, and although the word "evaluation" covers judgment based on broad perspectives it is essentially focused on merit. For the author, yardsticks depend on their purpose and can be explicit or implicit, have an aesthetic approach, or can be concerned with concepts like effectiveness or efficiency. Other measures can address questions about equity and justice, along with the acceptability of community standards.

An indicator is either a gauge or a measure of a variable. In the research context, it "… is a sign of the presence or absence of the concept we're studying." (Babbie, 2010, p. 131). To apply the framework proposed by Spencer et al. for assessing the quality of qualitative research evidence, Mays and Pope (2006) emphasize that it may be possible to use "… a series of quality indicators that point to the kinds of information needed to judge whether or not the quality feature concerned has been achieved". These indicators may be associated with each appraisal question.

Concerns about criteria for evaluating scientific production arose during the nineteenth century, when social scientists began to question the traditional model of research. This model was based on classic scientific criteria such as validity, reliability and generalization (André, 2001), but social investigators believed that traditional standards could not address new research approaches, particularly because qualitative studies sought to understand human and social phenomena. Discussing quality and credibility focused on qualitative research, Patton (2002, p. 542) argues as follows: "Judging quality requires criteria. Credibility flows from those judgments. Quality and credibility are connected in that judgments of quality constitute

the foundation of perceptions of credibility." This author then presents sets of criteria to judge research under its particular research tradition.[3]

There is no consensus related to which criteria are best for judging research quality. For instance, some authors argue that it is impossible to have only one set of criteria to do this in qualitative and quantitative research because of the differences involved in the conduct of these two types of research (Pozzebon, 2004; Smith, 1984; Whittemore, Chase, & Mandle, 2001). LeCompte and Goetz (1982) defend the possibility of using common criteria for qualitative and quantitative research. To judge research quality, particularly mixed-method social inquiries, Greene (2007) proposes a classification of two challenges to inquiry quality when the researcher must choose a set of criteria to judge an inquiry based on different research traditions: (a) to guarantee the 'quality of method' and data obtained, use criteria to judge the inquiry according to its tradition; (b) to guarantee the 'quality of the inferences', which includes conclusions and data interpretations, use different sets of criteria focused on the available data. Another approach combines evaluation and research, in which evaluation is understood in a broader context to include program evaluation, policy research, and studies of practice (Patton, 2002; Spencer et al., 2003). Considering the final view, which is notably related to a study of practice, this work uses a combination of different sources to show a set of general criteria that was judged as possible to use, after taking the appraisal questions that drive the judgment of scientific evidence into account.

André (2001) argues that some broader criteria and other criteria more specific to each study group might be defined. However, building these standards should be a collective and long-term task. The broader criteria to judge scientific production can be summarized as follows: (a) studies must have scientific and social relevance, i.e., they must be inserted into a theoretical framework and contribute to current knowledge; (b) research should have a well-defined object, aims or issues that are clearly formulated, a methodology that is appropriate for its objectives, and sufficiently described and justified methodological procedures; and (c) studies should contain a dense and grounded analysis that presents statements and conclusions supported by findings and demonstrations of what it adds to current knowledge. Conversely, although the author mentions more specific criteria for ethnographic research and action research, she argues that these criteria depend on each specific approach or research strategy, which is particularly true for qualitative research that is built/developed using different paths and techniques for data collection and analysis.

[3] Details are available in Chapter 9 (Patton, 2002).

Another study that discusses evaluation criteria highlights those that "are most relevant in five Key Evaluation Checklist (KEC) checkpoints," including 'consumers,' 'values,' 'process evaluation,' 'outcome evaluation' and 'comparative cost-effectiveness' (Davidson, 2005). Consumers are related to the audience. Values are the main aspect of the evaluation because they represent the substantive domain or "the essence" in the VNS (Brinberg & McGrath, 1985). Process and outcome are related to implementation and results, respectively (Webster & Watson, 2002; Weiss, 1997). Cost-effectiveness is associated with feasibility.

The purpose of the framework for assessing research evidence (Mays & Pope, 2006; Spencer et al., 2003) is to serve as a reference that helps evaluate the development and implementation of social policy, practices, and programs. Although the latter study is based on frameworks of qualitative research, it notes four central principles guiding proposal content that can be extended to other types of research, considering the concepts previously presented:

- *Contributory* in advancing wider knowledge or understanding about policy, practice, theory or a particular substantive field.
- *Defensible in design* by providing a research strategy that can address the evaluative questions posed.
- *Rigourous in conduct* through the systematic and transparent collection, analysis and interpretation of qualitative data.
- *Credible in claim* through offering well-founded and plausible arguments about the significance of the evidence generated. (Mays & Pope, 2006, p. 93; Spencer et al., 2003).

Contributory and credible aspects can be understood as part of the substantive domain (Brinberg & McGrath, 1985; Brinberg, 1982), whose perspective discusses validity such as value (or worth) in the VNS. Rigor and credibility are associated with the methodological domain of the VNS, when researchers must consider their best choices to achieve the best answers to their research questions. Rigor is an important element in the research process because it can bring scientific credibility through systematic data collection and analysis (André, 2001; Clarkson, 2012; Denscombe, 2010; National Institutes of Health, 1979). Credibility is a scientific criterion associated with peer acceptance (by scientific community members) by means of adequate peer review verifying whether specific knowledge fulfills the requirements for being considered scientific (Schwartzman, 1988).

A match among the concepts presented in the first topic of this chapter and the principles for qualitative evaluation cited by Spencer et al. (2003) might be possible. The contributory characteristic is related to relevance because of its applicability to the reality or solution of practical/theoretical problems. Rigor in conduct is directly related both to rigor

itself and to integrity. Spencer et al. (2003) argue that defensible design includes three aspects: the logic of inquiry (clarity of the research question, a clear theoretical orientation, and a method able to generate that type of knowledge and others), the choice of particular methods and sampling (well and carefully defined). Similar to rigor, the logic of inquiry may also help collect sufficient evidence to support the claims of the research: this is an element of credibility. The framework presents 18 appraisal questions grouped into nine key features, as shown in Table 5.

Table 5:
Framework for assessing qualitative evaluations

Key feature	Appraisal questions
Findings	1. How credible are the findings?
	2. How has knowledge/understanding been extended by the research?
	3. How well does the evaluation address its original aims and purpose?
	4. Scope for drawing wider inference – how well is this explained?
	5. How clear is the basis of evaluative appraisal?
Design	6. How defensible is the research design?
Sample	7. How well defended is the sample design/target collection of cases/documents?
	8. Sample composition/case inclusion – how well is the eventual coverage described?
Data collection	9. How well was the data collection carried out?
Analysis	10. How well has the approach to and formulation of the analysis been conveyed?
	11. Contexts of data sources – how well are they retained portrayed?
	12. How well has diversity of perspective and content been explored?
	13. How well has detailed, depth and complexity (i.e. richness) of the data been conveyed?
Reporting	14. How clear are the links between data, interpretation and conclusions – i.e. how well can the route to any conclusions be seen?
	15. How clear and coherent is the reporting?
Reflexivity and neutrality	16. How clear are the assumptions/theoretical perspectives/values that have shaped the form and output of the evaluation?
Ethics	17. What evidence is there of attention to ethical issues?
Auditability	18. How adequately has the research process been documented?

Source: Spencer et al. (2003).

Although this evaluation framework clearly mentions that its scope is based on certain premises regarding the nature of qualitative research, the quality patterns presented in the 18 central questions – grouped in various stages of the research process – may be useful in any type of research to address the quality indicators used to judge research evidence.

In light of these perspectives, it is important to clarify that the present work adopts the opinion that it is possible to have a set of general standards that include some changes of labels but that judge the same central element. Based on the literature and the various frameworks, a set of more general criteria was chosen to orient the analysis of the research process in accounting, which is summarized in Table 6.

Table 6:
Summary of general scientific criteria and their meanings

Criterion	Meaning/summary	Sources
Contribution / quality of theoretical perspective	The extent to which the target population might directly use the findings; the extent to which the study advances wider knowledge or understanding about the object.	(André, 2001; Mays & Pope, 2006; Meyrick, 2006; Spencer et al., 2003; Webster & Watson, 2002)
External validity or generalization/fittingness	The application of results in other contexts to generalize to other populations, or how well the hypotheses might fit in a different context from that which was generated. Factor: applicability.	(Beck, 1993; Brinberg & McGrath, 1985; Brinberg, 1982; McGrath & Brinberg, 1983; Spencer et al., 2003; Uncles & Kwok, 2013; Valentine, 2009; Yarbrough et al., 2011)
Feasibility	The extent to which resource and other factors enable an evaluation to be conducted in a satisfactory manner. Evaluate whether resources are sufficient to perform the study.	(Campbell, Braspenning, Hutchinson, & Marshall, 2002; Davidson, 2005; Rossi et al., 2004; Weiss, 1997; Yarbrough et al., 2011)
Impact	A change in the target population or social conditions that has been created by the study.	(Carmona, 2006; Rossi et al., 2004; Stufflebeam, 2000; Webster & Watson, 2002; Weiss, 1997)
Integrity	The rigor of research questions, design, conduct, and theorizing.	(Andrade, 2011; Antunes et al., 2011; European Science Foundation, 2011; FAPESP, 2011, 2012a, 2012b; OADS, 2012)
Internal validity/credibility or defensibility	A measure of the significance of the independent variable to the dependent variable, or how well and faithfully represented the phenomenon is. Factor: truth value.	(Beck, 1993; Brinberg & McGrath, 1985; Brinberg, 1982; Libby et al., 2002; Mays & Pope, 2006; Schwartzman, 1988; Spencer et al., 2003)
Relevance	The capacity of the research to help a group of practitioners solve problems. The relevance of the research depends on the potential applications for scientific, educational or applied purposes.	(André, 2001; Mays & Pope, 2006; Reiter & Williams, 2002; Schwartzman, 1988; Weiss, 1997; Yarbrough et al., 2011)
Reliability/auditability	Measure of the instrument's consistency in obtaining similar results or the ability of another investigator to follow the audit trail. Factor: consistency.	(Beck, 1993; LeCompte & Goetz, 1982; Mays & Pope, 2006; Rossi et al., 2004; Yarbrough et al., 2011)
Rigor/thoroughness	Full and faithful recording. Rich, detailed, and complex data.	(André, 2001; Denscombe, 2010; Evans et al., 2015; Mays & Pope, 1995; National Institutes of Health, 1979; Schwartzman, 1988; Van der Stede et al., 2005; Webster & Watson, 2002; Williams, 2014)
Suitability	Choosing a strategy that will likely bring success to achieve the research objectives and that clearly and explicitly justifies its choice.	(Creswell, Klassen, Clark, & Smith, 2011; Denscombe, 2010; Healy & Perry, 2000)

An understanding of the meaning of each of the general criteria selected for this study was formed from the various sources used. As a result, some criteria have previously been addressed in Topic 2.1.1, whereas others are presented in a brief overview as follows:

- Contribution or the quality of the theoretical perspective: this criterion is

characterized by the identification of knowledge gaps and description of the study contribution. In addition, as highlighted by André (2001) and Brinberg and McGrath (1985), this criterion can expand knowledge in the field and may offer solutions to the target audience's problems;

- External validity/generalization (or fittingness): whereas external validity is associated with the application of results in other contexts to generalize to other populations, in a qualitative approach, fittingness is related to the possibility of perceiving how well the hypotheses can fit into a different context from that in which they were generated. In both situations, the factor is applicability (Beck, 1993). In a general sense (Spencer et al., 2003), the validity concept is linked to the idea of "accurate correspondence with reality";
- Feasibility: this criterion can be defined as "the extent to which resource and other factors allow an evaluation to be conducted in a satisfactory manner." (Yarbrough et al., 2011, p. 288). To evaluate whether a study is feasible, it is important to consider data access, available resources, the research team's skills, time constraints, and other factors that may hinder the study. As highlighted by Davidson (2005), this criterion also contains a judgment about cost-effectiveness;
- Impact: in the program evaluation, impact may be defined as "a change in the target population or social conditions that has been brought about by the program, that is, a change that would not have occurred had the program been absent." (Rossi et al., 2004, p. 232). Similarly, in the case of scientific research, impact refers to the changes that occur because of or were induced by the study's results;
- Integrity: this criterion was discussed in Topic 2.1.2;
- Internal validity (or credibility/defensibility): in quantitative studies, internal validity "measures whether or not the manipulation of the independent variable really makes a significant difference in the dependent variable." The parallel in qualitative studies is credibility, which "… measures how vivid and faithful the description of the phenomenon is" (Beck, 1993, p. 264); both have truth value as their main factor. LeCompte and Goetz (1982, p. 31) argue that "the value of scientific research is partially dependent on the ability of individual researchers to demonstrate the credibility of their findings." In a general sense, validity is defined as "the extent to which a measure actually measures what it is intended to measure" (Rossi et al., 2004, p. 232);

- Relevance: this criterion was discussed in Topic 2.1.2;
- Reliability or auditability: reliability in quantitative research is the measure of the instrument's consistency in obtaining similar results. As a counterpoint, auditability has arisen as a measure in qualitative studies. Auditability is described as "… the ability of another investigator to follow the decision of audit trail." (Beck, 1993, p. 264). In other words, reliability is "the extent to which a measure produces the same results when used repeatedly to measure the same thing" (Rossi et al., 2004, p. 433). Another, more detailed definition is "… consistency and precision, freedom from random error" (Yarbrough et al., 2011). In these definitions, therefore, consistency is the primary factor;
- Rigor (or thoroughness): this criterion was discussed in Topic 2.1.2;
- Suitability: according to Denscombe (2010), whether a research strategy is good or bad – or right or wrong – is not a matter of definition. Suitability involves not only identifying and choosing a strategy that might bring success in achieving one's research objectives but also clearly and explicitly justifying one's strategic choices. In practice, suitability concerns how useful and how appropriate they are to solve the proposed research problem. As noted by André (2001), the central notion is to use an appropriate methodology for one's research purposes.

Another criterion is objectivity, which "requires facts capable of proof, and transparency in the handling of data" (European Science Foundation, 2011). In practice, this is part of integrity, however, there is no consensus regarding the possibility of application in different contexts and types of research. According to Spencer et al. (2003), although objectivity has been described as difficult to achieve (or as unachievable), it is reasonable. The issue addresses bias caused by one's methodology (selectivity of data, for example) or epistemology, particularly regarding certain types of qualitative inquiries that require the researcher to be immersed in the context to analyze the phenomenon, which involves a subjectivity that does not allow the separation of researcher and context. As a result, another aspect that appears is neutrality, which cannot be achieved in these types of research because of the researcher's influence.

2.3 Quality in Accounting Research

Both the international literature and the Brazilian scientific literature have presented a set of studies involving features related to the quality of accounting research. As is the case with other fields of knowledge, these studies have focused on parts (or stages) of the research process, along with an analysis of the performance and productivity of university business departments. The studies that were found are also mostly based on outcomes and involve the use of metrics to evaluate quality. In Brazil, some accounting studies have focused on the methodological paths described in theses, dissertations, and other studies published in leading journals and conferences. Others have addressed the ethical issues involved in research in the area. More recently, there have been studies discussing journals' IF, along with certain aspects of research practices in graduate programs.

2.3.1 Prior studies of research quality in business and accounting

To identify accounting studies that address quality attributes and research practices in the field, a search was performed in databases and journals, along with the use of the Google Scholar online tool. This resource reveals important evidence about the average performance of business scholars, such as the number of contributions, citations, and the h-index is much higher when performances are assessed using Web of Science (Amara & Landry, 2012), so, it can be considered useful to identify the best sources. The focus of the selected set of studies covers citation analysis (Aragão et al., 2014; Brown & Gardner, 1985; Dunbar & Weber, 2014); perceptions of the quality of accounting journals, also called peer reviews (Ballas & Theoharakis, 2003; Brinn et al., 2001; Brown & Huefner, 1994; Lowe & Locke, 2005; Lowensohn & Samelson, 2006; Taylor, 2011); productivity and quality in university business departments (Chan et al., 2012; Doyle & Arthurs, 1995; M.J. Jones et al., 1996; Martins & Lucena, 2014); quality criteria, such as validity (Libby et al., 2002); rigor (Evans et al., 2015; Williams, 2014); relevance (Reiter & Williams, 2002); impact (Carmona, 2006); and integrity (Andrade, 2011; Antunes et al., 2011); and research practices and features (Martins & Lucena, 2014; Mendonça Neto et al., 2009; Miranda et al., 2011; Taylor, 2011; Theóphilo & Iudícibus, 2005).

One investigation has treated business schools from the United Kingdom as a research-quality case study (Doyle & Arthurs, 1995). These authors explored citation analysis to judge aspects such as quality, excellence, and influence. Despite the good results obtained

by the departments of these schools in the 1992 governmental evaluation known as the UK Research Assessment Exercise, it is not the same as good results based on other judgments. For instance, seven of the departments studied achieved the highest score (on a scale of 1-5), equivalent to a level of international excellence. However, their applications for research grants have met with little success. Thus, Professor George Bain of the London School of Economics was appointed to chair a commission created by the Council for Economic and Social Research. The commission's goal was to analyze the reasons for the schools' low success rate, and its report concluded that "much research has lacked the rigor and critical reflection more common in other social science disciplines. Studies have tended to be atheoretical and non-comparative" (p. 258).

Doyle and Arthurs (1995) argue that judging research quality necessarily involves differentiating research quality from the concepts of influence and international excellence. For those authors, influence "may be measured by the extent to which work is cited by others, particularly where it is cited by people from other countries" (p. 259). However, a work's international excellence might be judged based on measurements by peers from outside the country. Finally, the authors propose "to use publication in journals (and in particular, in the top journals) as a measure of international influence, where 'top' is defined by the historic tendency of such journals to be cited, so that instead of working with actual citations, we work with expected citations." To assess the performance of business schools, Doyle and Arthurs (1995) suggest two approaches: the use of experts in the field (typically departmental chairs) to rate other departments, and the use of publicly available data to measure performance. Although this suggestion is driven to evaluate performance, it might easily be applied to judge research quality, and a parallel with the types of studies used to judge journal quality (i.e., citations and surveys about perceptions) can be made.

To understand different perspectives on perceiving and analyzing the quality of accounting research, select previous studies are summarized in Table 7.

Table 7:
Summary of previous studies of research quality in accounting

Reference	Approach	Source	Data	Purpose and main results (fragments)
Brown and Gardner (1985)	Theoretical-empirical	Paper in journal	Citations of articles from CAR published between 1976 and 1982 in other four journals.	Purpose: to assess the overall impact of major research journals on contemporary accounting research (CAR), and to identify the specific articles with the greatest impact on CAR. Citation analysis was applied to assess the contribution of four accounting journals to CAR, and those accounting articles with the most influence in this regard were identified. The Journal of Accounting Research (JAR) exhibited the highest impact in the three periods studied: 1963-1975, 1976-1978, and 1979-1982.
Brown and Huefner (1994)	Theoretical-empirical	Paper in journal	Survey (367 seniors faculty)	Purpose: to determine the familiarity of senior faculty at Business Week's "best 40 MBA programs" with and quality perceptions of 44 accounting journals. The findings revealed that five journals were nearly universally known, and 15 had wide recognition. Financial, managerial, and auditing faculty exhibited similar familiarity patterns, whereas tax faculty exhibited a somewhat different pattern. With respect to quality perceptions, relatively few journals were considered high quality. There was, however, general consensus across the various subject-area faculty members as to the top journals.
Jones et al. (1996)	Theoretical-empirical	Paper in journal	Documentary (comparing with peer review studies)	Purpose: to contribute to and develop the discussion initiated by Doyle and Arthurs (1995) on the publishing patterns of UK academics. The authors found that the fundamental argument of Doyle and Arthurs – that UK business-school academics do not publish in US journals – is confirmed by peer-reviewed studies in accounting. The main reasons for this are as follows: first, British academics are networked into the UK academic community; and second, British academics work mainly with British data that are not of primary interest to Americans.
Brinn et al. (2001)	Empirical	Paper in journal	Survey (134 respondents)	Purpose: there are two major issues: to determine how familiar are UK academics with the top US journals and to determine what UK academics' perceptions are of the reasons for their collective failure to publish in US journals? This paper presents eight possible reasons for UK accounting faculty's failure to publish in US journals: general explanations include (a) they do not believe they will succeed (b) they believe publishing in UK is easier; specific explanations include: (c) networking (non-US network) (d) content (non-US data); (e) existence of 'gatekeepers'; (f) methodology; (g) research quality; and (h) English and grammar.
Libby et al. (2002)	Theoretical	Paper in journal	None	Purpose: to present a predictive validity framework that accounts for methodological considerations regarding the efficiency and effectiveness of the experimental research design in accounting. A discussion of the link between theory and hypotheses, the operationalization of dependent and independent variables, the choice of levels of the dependent variables and other aspects that can help the researcher choose the path of experimental research.

(table continues)

Reference	Approach	Source	Data	Purpose and main results (fragments)
Reiter and Williams (2002)	Critical-theoretical	Paper in journal	None	Purpose: to study the accounting research based on the progress view (defined as innovation and relevance). The authors critically assessed that structural barriers lead to inadequate transformative criticism, which contributes to the lack of progress in accounting research programs.
Ballas and Theoharakis (2003)	Theoretical-empirical	Paper in journal	Survey (1,230 respondents).	Purpose: to examine how contextual factors such as a researcher's location and research orientation may influence perceptions regarding journal quality and readership patterns. Perceived quality was measured using four metrics: journal familiarity, average rank position, the percentage of respondents who classify a journal as top-tier, and readership. The results showed that there is significant variation in journal quality perceptions based on a researcher's geographic origin, research orientation, and journal affiliations.
Lowe and Locke (2005)	Theoretical-empirical	Paper in journal	Web-based survey (149 respondents)	Purpose: to elicit the views of academics in British accounting and finance departments regarding journal paradigm and quality. This paper used only two paradigms for classification: functionalist/positivist and critical/interpretative. Differences in the rankings of well-known accounting journals by capital markets and finance researchers compared to researchers from all other areas were statistically significant. Additionally, there was a broad similarity in the perceived quality of the top five journals and even the top ten journals. According to the respondents' perceptions, the top five journals consist of the following: Accounting, Organizations and Society; The Accounting Review; Journal of Accounting Research; Journal of Accounting and Economics; and Contemporary Accounting Research.
Theóphilo and Iudícibus (2005)	Theoretical-empirical	Paper in journal	Documentary study (238 articles, dissertations, and theses).	Purpose: to identify and critically assess the epistemological, theoretical, methodological, and technical dimensions, manifest and latent, observed in the scientific accounting literature. There is an assessment of empirical studies in terms of the closest approach to a scientific format. Conversely, studies have revealed difficulties in the use of certain methodological instruments, along with limited use of various methodological alternatives.
Carmona (2006)	Theoretical-empirical	Paper in journal	Documentary (410 papers published in the period 1990-1999)	Purpose: to address the roles played by publications, such as journal articles and research monographs, in disseminating accounting research. The study concentrates on accounting history. According to the author, the data shown in this paper indicate the superiority of generalist over specialist journals in the diffusion of accounting research, question the use of journal rankings, and suggest that books and research monographs have substantial impact on the diffusion of accounting research.

(table continues)

Reference	Approach	Source	Data	Purpose and main results (fragments)
Lowensohn and Samelson (2006)	Theoretical-empirical	Paper in journal	Survey (517 faculty members of 5 areas of the American Accounting Association – AAA)	Purpose: to identify top-quality research publication outlets in five specialized areas of accounting, as perceived by accounting faculty familiar with these areas. The five areas chosen include: accounting, behavior and organizations; American taxation association; government and nonprofit; information systems; and management accounting. Results show that although there are multiple outlets for high-quality research in specialized fields, many of these outlets are relatively new and have been overlooked in prior studies. Additionally, accounting academics in at least two specialized areas of accounting research (government and nonprofit and information systems) may experience difficulties substantiating the quality of their research.
Mendonça Neto et al. (2009)	Theoretical-empirical	Paper in journal	405 papers published in the conference proceedings (1996-2005)	Purpose: to analyze the works presented in the 1996-2005 editions of the Meeting of the National Association of Graduate Programs in Administration (ENANPAD). There was a predominance of positive research in 82% of the studies, and a normative approach was applied exclusively in 13% of them. The share of positive approaches proved to be increasing over the period, and the data showed a low productivity of the authors based on the coefficients of the bibliometric model of the generalized Lotka's Law.
Andrade (2011)	Theoretical-empirical	Dissertation (Doctorate)	Survey (85 respondents)	Purpose: to examine the attitudes of accounting researchers related to misconduct in accounting research. There is evidence of accounting researchers' involvement in inappropriate practices, such as sharing study authorship with a person who did not contribute in exchange for participation in another paper to which the first author will not contribute. Additionally, authors extend section references with citations to sources that they have not read but have merely seen mentioned in other articles to bolster their credibility.
Miranda et al. (2011)	Theoretical-empirical	Paper in journal	Content analysis (50 dissertations)	Purpose: to analyze the thesis statements of dissertations defended at the Graduate Program in Controllership and Accounting at the University of Sao Paulo (USP) between 2004 and 2008. There is a predominance of theoretical and empirical studies that are characterized as "quasi-experiments." The dissertations presented statements of research problems and/or aims and although the hypotheses to be investigated were largely enunciated, they fall short of the desired levels of scientific criteria. It is only since 2007 that authors have begun to clearly state the thesis (statement) of their research.
Antunes et al. (2011)	Theoretical-empirical	Conference proceedings	Survey (56 respondents)	Purpose: to understand how accounting academics realize the ethical aspects of research in the following dimensions: a) the research process; b) the guidance of students and c) relationships with peers and knowledge dissemination. Although almost all the ethical issues were considered important by the respondents, the paired analysis of the differences in responses between the importance of an ethical aspect and the practice shows that revealed belief is not always consistent with the practice adopted. These results suggest that some values considered important are not practiced with the same intensity; whereas other values regarded as extremely important are intensely practiced.

(table continues)

Reference	Approach	Source	Data	Purpose and main results (fragments)
Taylor (2011)	Theoretical-empirical	Paper in journal	Outcomes of 2008 Research Assessment Exercises (RAEs)	Purpose: to investigate the extent to which the outcomes of the 2008 Research Assessment Exercise in the UK, as determined by peer review, can be explained by a set of quantitative indicators. The study tested whether three cognate units of assessment – business and management (B&M), economics and econometrics (E&E) and accounting and finance (A&F) – are correlated with a set of quantitative indicators related to research activity. The main finding is that each of the three components of research activity (namely, research output, prestige, and research environment) is highly correlated with various quantitative indicators.
Chan et al. (2012)	Theoretical-empirical	Paper in journal	29,384 articles (48 journals in the period 1991-2010).	Purpose: to assess the research productivity of the accounting and finance community in Australian and New Zealand Higher Education Institutions (HEIs) using high-quality accounting and finance journals. Australian and New Zealand HEIs publish approximately 1,658 and 299 weighted articles, representing 5.6 per cent and 1.0 per cent of the total, respectively. The HEIs' research output steadily increased during the 20-year period. It is a challenge for academics to publish multiple articles. If an author manages five total appearances over a 20-year period, he/she is in the top 15 percent of all authors.
Aragão et al. (2014)	Theoretical-empirical	Paper in journal	Citation analysis (1,675 citations related to 577 articles).	Purpose: to identify which characteristics of papers published in Brazilian accounting journals are associated with their resonance in scientific production. Associations were found between the journal, year of publication, language, and citations received per citation vehicle; associations were also found between the language of the paper and the country of the citation. The resonance of the scientific production analyzed was considered low; data variability is related to the characteristics of the articles/journals; and the distance among the journals' IFs has decreased.
Dunbar and Weber (2014)	Theoretical-empirical	Paper in journal	Citation analysis (165,314 citations related to 3,538 papers).	Purpose: to identify individual research that have had a strong influence on contemporary accounting research. The authors present and discuss lists of the individual works that are most heavily cited in each category (audit, financial, managerial, tax, and other) and methodological approach (archival, experimental, theoretical, and other). Papers from nine journals were analyzed.
Martins and Lucena (2014)	Theoretical-empirical	Paper in journal	Survey (113 respondents)	Purpose: to identify the profile and scientific practices of professors of postgraduate programs in accounting. There is evidence that respondents tend to separate the results of their research into different publishable parts, and when rejected from the highest-impact journals, the research is submitted to lower-impact journals until it is accepted for publication. Respondents do not, on average, name others as authors who have not participated effectively in the study.

(table continues)

Reference	Approach	Source	Data	Purpose and main results (fragments)
Williams (2014)	Theoretical	Paper in journal	None	Purpose: to describe two significant flaws in the claim that the prevailing form of rigorous accounting research is actually as rigorous as claimed, when considered only in terms of how rigor is conceived within that prevailing form. The author submits a severe criticism of the supposed rigor adopted in accounting research, considering that some studies apply rigor only as a rhetorical device, given the strength of the adjective 'rigorous'. Most of the accounting research is predicated on the putative quantitative nature of accounting data. However, accounting numbers are not particularly precise because they are operational numbers, not quantities, and the likely intractability of the data problems suggests that this research methodology is not the only path to accounting understanding.
Evans et al. (2015)	Theoretical	Paper in journal	None	Purpose: to help accounting researchers, particularly those who are less experienced, improve the quality of their research and more clearly communicate how they address key issues. The article presents a list of Points to Consider (PTCs) when conducting empirical accounting research. PTCs are organized into the following five issues: research question, theory, contribution, research design and analysis, and interpretation of results and conclusions.

A study on perceptions conducted a survey of 134 UK accounting and finance academics on the topic of why such scholars do not publish in top American journals (Brinn et al., 2001). That study thus addresses two major questions: (1) how familiar are UK academics with the top US journals; and (2) what are UK academics' perceptions of the reasons for their collective failure to publish in US journals. The findings revealed eight answers, which were classified by the authors as two general explanations and six specific explanations, as shown in Table 8.

Table 8:
Reasons that English articles are not published in leading American accounting journals

Reason	Perception
General	
Researchers believe they will not succeed achieving publication	The perception among these authors is that their articles will not be published, and they do not submit papers as a result. As a consequence, it becomes a self-fulfilling prophecy because if the article is not submitted, it will certainly never be published.
It is easier to publish in the United Kingdom	British academics prefer to submit their papers to journals in the UK, where they believe the chance of success is greater than in the US.
Specific	
Networking (outside the American network)	UK academics believe that being outside of the American academic network is a major impediment to publishing in American journals.
Content	American journals are only interested in US data.
The existence of gatekeeping	There is a perception of gatekeeping, defined as American faculty's tendency to favor submissions from the US to the detriment of authors who are not from the US.
Methodology	American journals adopt highly empirical methodologies.
Research quality	American journals have a higher quality differential. They are more demanding and their key criterion is contribution to knowledge. Additionally, Americans' academic background and innovative capacity are, in general, better.
English and grammar	English and grammar are a major contributory factor to the lack of UK success.

Source: Brinn et al. (2001). Adapted.

Some of these explanations are particularly interesting to a non-native speaker of English. Although UK researchers tend to be native English speakers, the English language constitutes a barrier for publishing in the US because the language that English academics use in their articles may not resonate with American English speakers, reflecting some characteristics of regional or national idiomatic usage. Another issue is that UK academics believe that American researchers' have better academic backgrounds than their own, which is a factor that contributes to the higher quality of research published in the US. To break some of these barriers, based on the results, the authors argue that some respondents "recommended collaboration with a US author as a way to circumvent the US network issue" (Brinn et al., 2001, p. 230).

2.3.2 Research quality in accounting graduate programs under the regulatory view in Brazil

The CAPES Foundation, which regulates and is responsible for evaluating Brazil's graduate programs, uses metrics to judge and classify journals per strata using the Qualis system, which is the standard for assessing scientific journals in various fields of knowledge in Brazil. In the area of administration, accounting, and tourism, this process has involved the use of a set of criteria that defines the stratum in which a journal should be classified, as explained in documents pertaining to each area of study.

The content of the document from the Board of Assessment that relates to the area of administration, accounting and tourism (CAPES, 2012, p. 3) was applied to the electronic version of Qualis/CAPES (entitled "Webqualis") for journals published during the 2007-2009 period, along with Coleta CAPES data[4] related to articles published in 2010. The document emphasizes that there will be discussions in the area of study to establish new quality indicators focused on journals' content; those indicators will be added to existing shape criteria. This position indicates that area members of the regulators of Brazilian graduate programs recognize the need to define evaluation criteria for paper content. Judgment is based on considerations in addition to the shape of journals, which was done in previous Qualis assessment processes.

The Area Document for administration, accounting and tourism for 2010-2012 (CAPES, 2013) demonstrates the need for greater acceptance of scholarly production. CAPES sets up projects funded by agencies or organizations that promote scientific research, and projects selected for funding through competitive processes that use peer review are particularly valued. Accordingly, CAPES indicates that the existence of projects funded by development institutions is a quality indicator for a graduate program. The resulting scientific production of funded projects tends to be understood as higher quality because the proposals for such projects were previously judged by the funding evaluators. Consequently, funding is an important criterion that indicates a program's quality.

With respect to quality criteria, the 2013 Area Document (CAPES, 2013) stressed that its evaluation of journals used the same criteria as the previous assessment, plus quality references created for the area based on previous discussions with editors. The regulator recognizes an evolution in journals during the period, but once more, it highlighted that the

[4] Coleta CAPES was an online platform used by graduate programs to send data about these programs to the regulator. It was replaced by the Sucupira Platform.

shape's homogeneity ceased to be a differentiating criterion of quality for classifying periodicals in the Qualis/CAPES strata, as noted in 2012. According to CAPES, this evolution may be credited both to the induction process generated by previous reviews and to the National Association of Graduate and Research in Administration's (ANPAD) discussions with editors about good publication practices.

These criteria detailed in the Area Document published in 2013 were primarily based on the shape elements of the journal[5] and are not focused on assessing the content of what is published. Nonetheless, some of the strata criteria for quality elements relate to content, such as peer review and IFs such as H-Scopus and JCR, which are taken into account in the classification of the superior levels (A1, A2, and B1). In this case, the number of Brazilian accounting journals in these layers is small: most such journals do not have a JCR IF because they are published in Portuguese. The area documents have indicated that content criteria should be included in the assessment process of the next three years, given the difficulty of differentiating the periodicals. Therefore, in the next quadrennial evaluation, new criteria will likely have to be defined to judge the quality of the graduate programs' scientific production.

In 2015 CAPES conducted a classification process of the area journals for the quadrennium 2013-2016, and the report was published in May of that year (CAPES, 2015b). To do so, a preliminary assessment of scientific production of graduate programs concerning 2013 and 2014 was carried out, based on data reported by the programs through Sucupira Platform. This assessment allows a projection for the quadrennium 2013-2016 from the experience of its first two years and the previous triennial evaluation. The report shows that about two-thirds of journals with Brazilian production in the area of administration, accounting, and tourism in 2013-2014 – already classified and that have impact factor – are not on the list of journals that belong to the areas of administration, business, accounting, tourism and hospitality, and public administration on the Thomson Reuters SCImago, SciELO, and Redalyc databases. Based on data, the report itself points out that it is likely that the participation of published journals outside Brazil grow in the 2013-2016 assessment, but not necessarily in journals belonging to the core of the area. This evidence suggests that

[5] According to the Report (CAPES, 2015b), previous evaluations were based on a list of formal parameters of quality: description of mission and focus; having defined and reported periodicity on its website; having peer review; showing submission standards; reporting name and affiliation of the editor; informing name and affiliation of the members of the editorial board; the editorial board composition should be diverse regarding the membership of its members; annually disclose of the reviewers' nomination; having at least two issues per year; reporting complete data of articles in the article itself; reporting affiliation of the authors; informing address of at least one of the authors; reporting the formalities assessment/approval; reporting the evaluation process (responsible editor, date of receipt of the article and the phases of the evaluation process and acceptance); presenting the bibliographical legend of the journal in each article; chief editor cannot be author.

Brazilian researchers in the field may be having difficulty inserting their scientific production in international journals recognized as the core of the area and, therefore, choose to submit their manuscripts to journals from other areas, which tend to have less weight on the definition of quality scientific production in administration, accounting, and tourism.

The set of current evaluation criteria from Qualis/CAPES is shown in Table 9.

Table 9:
Evaluation criteria to judge journals by the Qualis/CAPES in quadrennial 2013-2016 (2015).

Stratum	Criteria to be classified in the stratum
A1	• ISSN; • Having at least two issues/year; • JCR >1.4 (67%); • H-Scopus > 24 (75%); • Journals within the limits above but which were not listed as belonging to the area, according to the Impact Factor calculation bases were classified in stratum A2.
A2	• ISSN; • Having at least two issues/year; • 1.4 >= JCR > 0.7 (33%); • 24 >= H-Scopus > 9 (50%); • Journals within the limits above but which were not listed as belonging to the area, according to the Impact Factor calculation bases were classified in stratum B1.
B1	• ISSN; • Having at least two issues/year; • Scielo with IF > 0.01 and belonging to the area by Qualis/CAPES base criterion, or • 0.7 >= JCR > 0; • 9 >= H-Scopus > 0; • Journals within the limits above but which were not listed as belonging to the area, according to the Impact Factor calculation bases were classified in stratum B2.
B2	• ISSN; • Having at least two issues/year; • Being in Redalyc or to be edited by the Publishers described on the Area Document (Sage, Elsevier, Emerald, Springer, Inderscience, Pergamo, Wiley, Routledge, and Taylor & Francis); • Or Scielo FI < 0.01, or Scielo FI >0.01, but, belonging to the another area by Qualis/CAPES base criterion.
B3	• ISSN; • Having at least two issues/year; • Delay-Index <=0.5; • Having 3 or more years of existence; • Belonging to at least one of the index databases defined in the Area Document: Ebsco, Doaj, Gale, Clase, Hapi, ICAP, IBSS – confirmed indexing.
B4	• ISSN; • Having at least two issues/year; • Delay-Index <=0.5; • Having 2 or more years of existence.
B5	• ISSN; • Having at least two issues/year; • Maximum delay of issue = one year.

Source: CAPES (2015b). The content of this table is a free translation from the original in Portuguese.

In the Report 2015, CAPES took an important step to establish that, given the homogeneity of the compliance of formal parameters of previous evaluations, discrimination

passes the focus on criteria that may indicate the impact and relevance of journals and - supposedly - indirectly articles published by them. The upper strata were occupied by journals with impact factor calculated by one of the bases considered by the area. The classification in one of the lower strata was due to issues related to periodic management and age. Thus, only the journals with impact factor were classified as B1, a requirement that was not present in the previous evaluations. In addition, the CAPES took an action that discourages the practice of publishing in journals from other areas to establish that these journals that meet the requirements of a particular stratum will be classified at a lower level than those of the area itself. It is an important incentive mechanism for researchers to publish in their field journals to increase the visibility and relevance of the Brazilian production, as well as it improves productivity indicators to inhibit publications on low-impact journals to the area.

Unlike the triennial evaluation of graduate programs, the strata that Qualis/CAPES assigns to journals in each area have been updated at different moments: during the PPGs' assessment and, at least, once more at some time over the three/four-year period. As a result, the quality of the graduate program is influenced by the publication vehicle for the scientific production of faculty and students in the PPG and its respective stratum, after considering the publication year and the channel by which such scientific dissemination occurred.

After judging and defining the stratum for each journal, such information is used for two purposes in the evaluation: to calculate the productivity rate for each faculty member of the graduate program and to measure the quality of each graduate program, based on the journals in which the papers considered in the evaluation were published. There is a yardstick to measure the value of an article: A1 = 100 dots; A2 = 80 dots; B1 = 60 dots; B2 = 50 dots; B3 = 30 dots; B4 = 20 dots; B5 = 10 dots; and C = 0 dots. Additionally, if a paper has more than one author who is also a faculty member (co-authoring among faculty members of the same program), the dots for that article in the program evaluation are distributed evenly among the co-authors to judge the quality of the program. For example, if a paper was published in a journal whose stratum is A1 (100 dots) and the paper has three faculty members listed as co-authors, each member will be awarded 1/3 of the credit (i.e., 33 dots). Unlike the proportional distribution of the dots of co-authorship papers made to qualify the program, dots are fully allocated to assess each co-author's productivity – faculty member – individually. During the period of the last triennial evaluation, each faculty member had to earn at least 150 dots.

In conclusion, it is important to note that on September 21, 2015, the CAPES Evaluation Board (DAV/CAPES) released Circular Letter nº. 022/2015 (CAPES, 2015a),

whose subject was the publication of the updated Qualis for journals for 2013 and 2014. The results are available for public access on the Sucupira platform.[6] However, the next Area Document was not published until December 31; consequently, it is understood that the standards established by the previous official paper remained valid (CAPES, 2013). Notably, this is only one dimension used by the triennial/quadrennial evaluation to define program quality, in addition to other elements such as an evaluation of books published by faculty members and the quality of the program's proposal.

2.3.3 Quality criteria and research practices in the steps of the research process (evaluation)

After defining the list of general criteria and considering the elements presented in the Framework for Assessing Research Evidence, a matrix was built to guide the structuring of collection instruments.

First, the matrix included the list of 10 general scientific criteria selected based on the literature, as shown in Table 6. Next, the nine key features presented in the evaluation framework were added: design, sample, data collection, analysis, findings, reporting, reflexivity and neutrality, ethics, and auditability. In this dissertation, these characteristics are treated as synonyms of phases (or stages) of the research process. In the next step, the appraisal questions from the framework, as proposed by Spencer et al. (2003) and Mays and Pope (2006), have been incorporated. The key features and appraisal questions were presented in Table 5. To adequately describe the main characteristics of this study, the 'sample' label was replaced by 'research subjects' because the intention is to offer a comprehensive proposal that can help assess the various designs of and approaches to the accounting research process.

Relationships and attributes were identified in the literature to help guide the assessment at each stage of the research process, and they were then added to the array. Although the main sources to do this have been Spencer et al. (2003) and Mays and Pope (2006), a dozen other sources were considered, such as the validity schema network (Brinberg & McGrath, 1985), points to consider (Evans et al., 2015), principles of evaluation (AEA, 2004; Yarbrough et al., 2011), and legislation regarding ethics in research using human subjects (Brasil, 2012, 2013; National Institutes of Health, 1979). Additionally, the focus of each relationship/attribute was indicated. Although the definitions of the indicators were not

[6] Available at: http://sucupira.capes.gov.br/sucupira/public/consultas/index_consultas.jsf.

used because of their measuring characteristics, some attributes are clearly associated with those indicators.

In summary, Table 10 shows the matrix with these elements. However, an entire matrix, including VNS domains, the stages of the research process, and statements for Delphi, is presented in Appendix A (English version) and Appendix B (Portuguese version).

Table 10:
Matrix of general criteria and their attributes/relationships

Criterion	Key feature	Focus	Appraisal question (Spencer et al.)	Attribute/relationship
Contribution/quality of theoretical perspective	Findings	Advancement of knowledge	How has knowledge/understanding been extended by the research?	Insights for thinking about the field
		New areas for future studies	How has knowledge/understanding been extended by the research?	New areas based on the findings
	Reporting	Boundaries	How has knowledge/understanding been extended by the research?	Boundaries of the study
		Limitations	How has knowledge/understanding been extended by the research?	Disclosure of limitations
External validity or generalization/fittingness	Findings	Applicability	How credible are the findings?	Comparison of results with those of other studies
	Reporting	Scope	Scope for drawing wider inference – how well is this explained?	Explicit possibility of generalization
Feasibility	Design	Access to data sources	How defensible is the research design?	Access to the data
		Audience	How defensible is the research design?	Research strategy x target audience
		Time constraints	How defensible is the research design?	Time constraints
Impact	Reporting	Advancement of knowledge	How has knowledge/understanding been extended by the research?	Discussion of impact on knowledge
Integrity	Data collection	Ethics	What evidence is there of attention to ethical issues?	Formal agreement from humans
		Ethics	What evidence is there of attention to ethical issues?	Voluntary participation of subjects
	Analysis	Logic of inquiry	How well has the approach to and formulation of the analysis been conveyed?	Description of nature and form of data
	Reporting	Limitations	How well does the evaluation address its original aims and purpose?	Reasons for limitations
	Reflexivity and neutrality	Robustness	How clear are the assumptions/ theoretical perspectives/values that have shaped the form and output of the evaluation?	How to address errors and biases
		Robustness	How clear are the assumptions/ theoretical perspectives/values that have shaped the form and output of the evaluation?	Impact of team participation on the research
	Ethics	Code of Ethics	What evidence is there of attention to ethical issues?	Use of the principles of an ethics code
		Confidentiality	What evidence is there of attention to ethical issues?	Confidentiality of participants' data
		Respect for human subjects	What evidence is there of attention to ethical issues?	Formal respect for human subjects
			What evidence is there of attention to ethical issues?	Strategy to mitigate possible harm
Internal validity/ credibility or defensibility	Research subjects	Logical coherence	Sample composition/case inclusion – how well is the eventual coverage described?	Representativeness of subjects
		Representativeness	How well defended is the sample design/target collection of cases/documents?	Criteria to design/select subjects
	Findings	Logical coherence	How clear is the basis of evaluative appraisal?	Impact from the nature of any divergences

(table continues)

Criterion	Key feature	Focus	Appraisal question (Spencer et al.)	Attribute/relationship
			How credible are the findings?	Path to achieve conclusions
			How well does the evaluation address its original aims and purpose?	Check links of findings x purpose
		Support for data and evidence	How credible are the findings?	Link between findings x evidence
	Reporting	Logical coherence	How clear and coherent is the reporting?	Literature review x main concepts
			How clear and coherent is the reporting?	Theory to support propositions
			How clear are the links between data, interpretation and conclusions – i.e. how well can the route to any conclusions be seen?	Appreciation of drafts by colleagues
			How well does the evaluation address its original aims and purpose?	Conclusions x aim
Relevance	Analysis	Significance	How well has detailed, depth and complexity (i.e. richness) of the data been conveyed?	Significance of data to reaching goals
	Findings	Gap	How well does the evaluation address its original aims and purpose?	Previous findings x hypothesis
Reliability/auditability	Analysis	Consistency	Contexts of data sources – how well are they retained and portrayed?	Context x impact on data analysis
			How well has diversity of perspective and content been explored?	Other views provided to explain context
	Findings	Robustness	Scope for drawing wider inference – how well is this explained?	Does the context allow replication
	Auditability	Consistency	How adequately has the research process been documented?	Records of the reasons for changes
			How adequately has the research process been documented?	Safeguard databases for checks
	Auditability	Methodological rigor	How adequately has the research process been documented?	Guard documents to reduce risks
			How adequately has the research process been documented?	Records of design changes
Rigor/thoroughness	Data collection	Methodological rigor	How well was the data collection carried out?	Notes for divergent events
			How well was the data collection carried out?	Notes for each research step
	Analysis	Logic of inquiry	How well has detailed, depth and complexity (i.e. richness) of the data been conveyed?	Implicit/explicit links - findings x aims
		Methodological rigor	How well has the approach to and formulation of the analysis been conveyed?	Description of tools and procedures
	Ethics	Ethics	What evidence is there of attention to ethical issues?	Compulsory submission to the CEP
			What evidence is there of attention to ethical issues?	Registration of the study with the Ethics Committee (CEP)
			What evidence is there of attention to ethical issues?	Formal submission to the CEP
			What evidence is there of attention to ethical issues?	Rules for registering studies with the CEP
Suitability	Design	Appropriateness	How defensible is the research design?	Clear overview that guides the study
			How defensible is the research design?	Goal/problem shown precisely
			How defensible is the research design?	Overview x strategy
			How defensible is the research design?	Overview x theory
		Usefulness	How defensible is the research design?	Arguments for technique choices
			How defensible is the research design?	Useful strategy for purpose

Note: The table is ordered by criterion, key feature (or stage of the research process), and focus.

CHAPTER 3 –
RESEARCH DESIGN

This chapter presents the research design and discusses the tools and techniques used both to obtain the observations and to relate them to one another. Moreover, this chapter helps identify the attributes of good research and the collection of primary data to build consensus about the predicates of research practices. Furthermore, the criteria from the literature are compared with criteria determined as the consensus on research practices in Brazilian accounting research, after considering not only factors that demonstrate the links among the substantive, conceptual, and methodological domains but also key features used to evaluate the research process. Finally, an approach is presented that discusses the attributes of good research to judge the quality of scientific production in accounting at various stages of the research process. Based on that perspective, the chapter addresses the act of conducting research to obtain a set of evidence that supports its arguments carefully and logically.

3.1 Method Choice and Steps of Research Development

The study is based on the inductive perspective with the use of mixed methods. First, the research is quantitative, with data collection conducted by applying questionnaires using iterative queries in rounds of the Delphi Technique. The research is also qualitative and is focused on building a theoretical and conceptual approach to obtaining a more detailed understanding of quality in the process of developing good academic research based on best practices. The study is designed in the following four stages:

1. Bibliographical research to identify the attributes of good research as described in the literature;
2. Gathering primary data from accounting researchers using the Modified Delphi Technique in two rounds to build a consensus about the attributes of good research that are present in the respondents' research practices;
3. Comparing the attributes identified in the literature with those resulting from the Delphi consensus about practices in the research process, considering factors that demonstrate linking among the substantive, conceptual, and methodological domains, along with key features to evaluate the research process; and

4. Presenting an approach that discusses the attributes of good research to help judge the quality of scientific production in accounting at various stages of the research process.

Based on the Validity Schema Network (Brinberg & McGrath, 1985), the research design of this study is summarized in Figure 4:

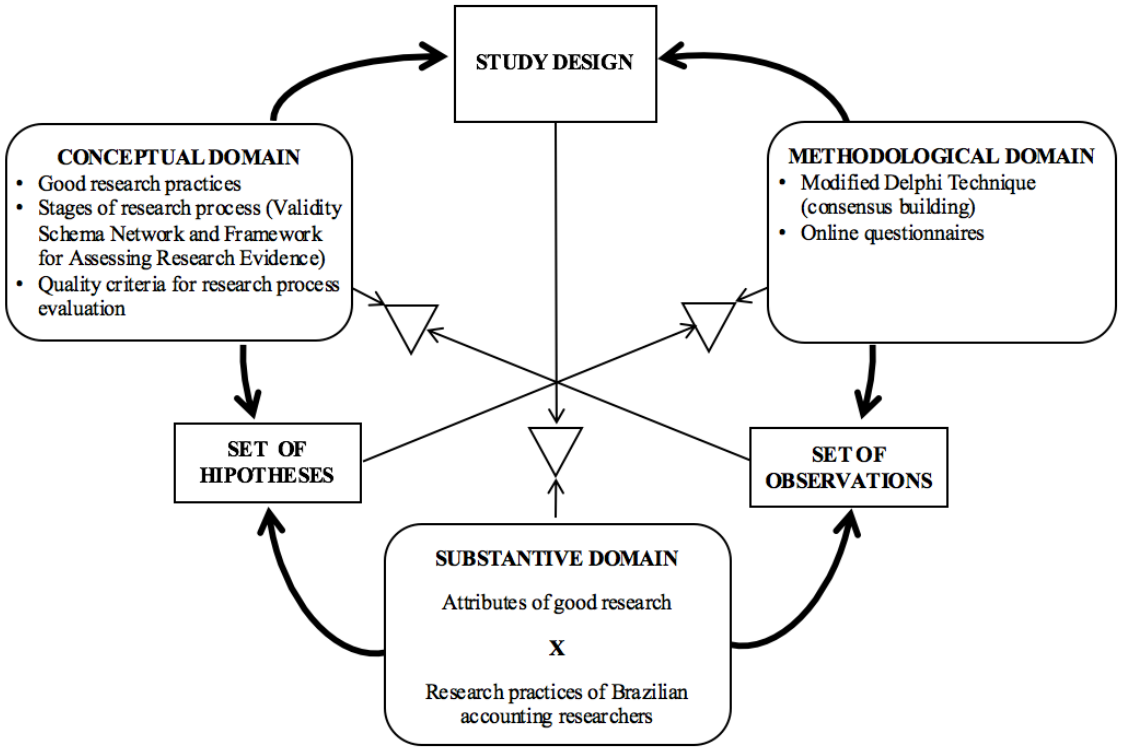


Figure 4: Summary of research design.

The research design was submitted to the Ethics Committee for Research with Human Beings at the State University of Feira de Santana (CEP/UEFS) through Plataforma Brasil (http://plataformabrasil.saude.gov.br/)[7] in September 2015, and the CEP/UEFS Committee approved the design (Appendix C). The summarized version of the design and the informed consent were prepared using Resolution CNS n. 466/2012 (Brasil, 2012) and Operational Standard CNS n. 001/2013 (Brasil, 2013) complementary to that resolution; both the resolution and the standard were obtained from the Brazilian National Health Council (CNS). These legal instruments are Brazil's primary regulatory basis for research involving humans.

[7] Plataforma Brasil is a national platform created by the Brazilian federal government to unify the registration of research involving human beings. Research designs are submitted and evaluated through it, and a summary of each study is available for public consultation over the Internet. It covers studies registered at the local or institutional Research Ethics Committees (CEP) levels, along with at the National Research Ethics Commission (CONEP) level. The CEP/CONEP system is under the jurisdiction of the National Health Council, which is linked to the Brazilian Ministry of Health.

The informed consent was written in the form of an invitation letter and presented all the information required by the Ethics Committee and current legislation regarding the study's goals, risks and benefits, guarantees of confidentiality and privacy, and other legal issues.

3.2 Stage 1: Identifying the Attributes of Good Research in the Literature

At this stage, the attributes of good research have been identified in the literature. Elwyn et al., (2006) have built a framework for quality criteria to assess instruments for helping patients make decisions. The authors used an initial list of quality areas proposed by an earlier study. The list was submitted for consideration by the delegates of the International Shared Decision Making Conference, which was held in Swansea (Wales) in 2003. One hundred and eighty-one delegates discussed the validity and created a list of 12 broad "quality domains," which the authors used as the basis for their study, grouping the statements from a data collection instrument designed to apply the Delphi methodology. According to the researchers, 80 criteria were included for evaluation by respondents. That paper served as the reference for developing the data collection instrument.

The purpose of this stage was to find elements to create the instrument to apply in the first round of the Delphi. The focus was an attempt to identify relationships that might be useful to understanding respondents' perceptions regarding the quality attributes of research practices. To do so, an orientation matrix for building the instrument (Appendixes A and B) was made based on the theoretical platform. Initially, the stages of the research process and the domains of scientific research (Brinberg & McGrath, 1985) were inserted into the matrix. Next, matching was performed for the nine key characteristics (here regarded as synonymous of stages) of the research process (Mays & Pope, 2006; Spencer et al., 2003). The nine key features include findings, design, sample, data collection, analysis, reporting, reflexibility and neutrality, ethics, and auditability. In this manner, the appraisal questions for assessing research evidence were introduced, and each issue was split into its respective quality indicators (Spencer et al., 2003) and the quality criteria (or attributes) identified in the literature. Wider criteria were associated with relationships and more specific criteria to give the study more of a practical sense. Finally, based on these indicators, criteria, and relations, the propositions were elaborated to compose the data collection instrument.

The data collection tool was written in Portuguese and the propositions were grouped into those nine essential features, with some modifications, considering the nature of this

study. The findings were changed to the fifth position, and the 'sample' label was replaced by 'subjects of research.'

3.3 Stage 2: Data Collection with the Delphi Technique for Consensus-Building

The Delphi technique is defined as a set of iterative procedures applied to a non-attendance group to obtain the consensus opinion on an issue or a set of issues for which there are insufficient or contradictory data (Justo, 2005). The Delphi technique for consensus-building is used to achieve a convergence of opinion. In the medical field, the Delphi technique is used to build frameworks and guidelines in studies to choose the best clinic criteria for a particular disease (Graham, Regehr, & Wright, 2003), or to identify risk factors associated with a particular illness (Deckers et al., 2015). Additionally, it is used to build evaluation frameworks for judging quality criteria (Elwyn et al., 2006). It is also used in other fields such as the humanities (Hug et al., 2013), food (Kim, O'Bryan, Crandall, Ricke, & Neal, 2013) and accounting (Michael John Jones & Xiao, 2004).

Graham et al., (2003) summarize some of the advantages of using the Delphi technique for consensus-building that are described in the literature:

(1) Not having to gather participants in person increases the feasibility and significantly reduces the cost;
(2) There are less likely to be constraints on the group's size or composition because participants can be recruited from various geographical locations;
(3) The group's consensus reliability for the question being analyzed improves as the number of panel members increases. It is easier to achieve an appropriate panel size for the issue at hand because of the inherent feasibility of the Delphi process;
(4) "The anonymous nature of the exercise ensures that a single influential participant will not have a disproportionate impact on the outcome of the group, as can occur with other group processes." (p. 1.150).

By contrast, Graham et al., (2003) also present criticisms of the use of Delphi technique from the literature:

(1) It is "subject to bias because the investigator limits the scope of the issue assessed by the panelists." Consequently, the issue in question is at least partially controlled by the researcher, and the consensus can thus be somewhat distorted;
(2) Panelists do not meet, as they do in other group processes that depend on interaction among participants as a source of new insights into a problem;

(3) "[D]ue to the nature of Delphi, no discussion takes place, and any consensus that the group appears to have developed can only derive from information provided to it by the investigator," and the consensus reached in other types of group processes may thus be significantly different; and

(4) Criteria have not been established to determine whether group consensus was reached (p. 1,151).

To mitigate the risks of possible biases in applying the technique, this study's strategy for drawing up a data collection instrument involved an optional open field for suggestions and comments on the propositions and on the tool itself. While participants were responding, they could comment on attributes described in the literature. Additionally, they were invited to highlight points considered relevant in research practice that either were not mentioned or were unclear during the first-round statements.

Tetzlaff, Moher, and Chan, (2012), from the Clinical Epidemiology Program of the Ottawa Hospital Research Institute, drew up a consensus-forming survey using the Delphi Technique to develop a guideline for creating a clinical study protocol. The panel consisted of 96 specialists, including trial investigators, methodologists, research ethics board members, funders, industry, regulators, and journal editors. The response rate was between 88 and 93% in each round.

Skulmoski, Graham, and Krahn, (2007) present a flowchart of planning and conducting research using the Delphi Technique over three rounds (Figure 5).

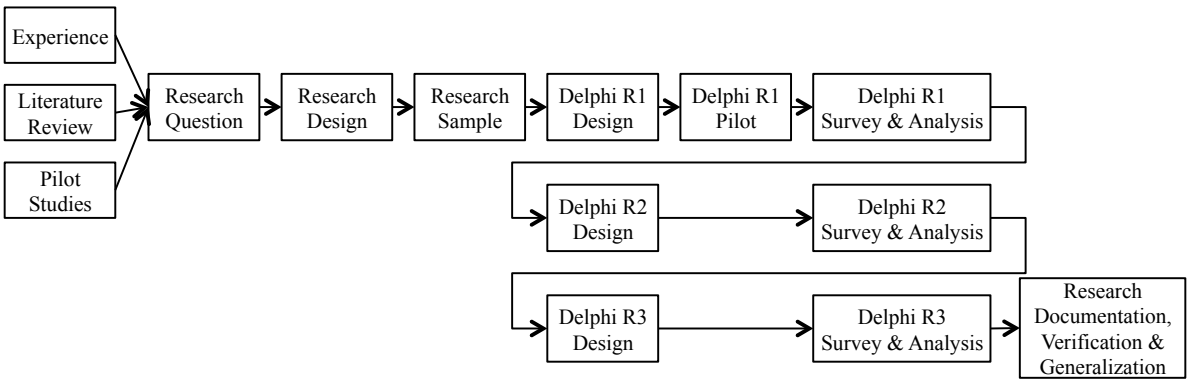


Figure 5:
Three round Delphi process
Source: Skulmoski, Hartman & Krahn (2007, p. 3).

The flowchart (Figure 5) helped define the steps taken in this work. Kelbaugh (2003) uses the modified Delphi Technique to obtain consensus on identifying the success factors of teamwork in extension programs. According to Kelbaugh, in the classic Delphi, the first

round of the research is open and allows each participant to provide a single entry. The ideas and items so generated are then used to develop the next phase of the research. The modified technique is so named because the researcher creates the first-round propositions based on an exhaustive literature review and his or her experience in the subject area. To do this, Kelbaugh (2003) uses an intentional sample. The expert panel featured 25 extension professors from American universities financed from governmental land concessions (called "land grant" institutions). The experts took part in three rounds. The areas of the studied extension program include agriculture and natural resources, family and consumer sciences, youth and community and economic development. Kelbaugh (2003) also indicates that the modified Delphi was used in the first round of statements and the survey used a 6-point Likert scale. The study was conducted using a Web-based survey instrument, which enables quick answers and the ability to analyze the data in real time.

The online Delphi Technique is appropriate in the context of this study for enabling the collection of various individuals' attitudes and opinions. The Delphi facilitates the involvement of people from different geographical regions because in-person participation is not required. Moreover, the use of the Internet makes the Delphi Technique low-cost and feasible. These features further allow the sample size to be increased without many resources.

3.3.1 Number of rounds, composition, and selection of panel members

In the literature, there is no consensus on the numbers of rounds and members needed to form a Delphi. To analyze the diversity of sample size and the number of rounds in the application of the technique, Skulmoski et al., (2007) identify 16 articles published between 1973 and 2005. Nine of those articles are in the field of information systems/technology information, and seven are not. There were between one and three rounds in these studies, and the sample size ranged from 3-171 participants. In turn, Graham et al., (2003) use the technique in two rounds, with 14 participants in the first round and 12 in the second round. Elwyn et al., (2006) construct an assessment framework to evaluate patient decision aids with the use of a consensus-building, two-round Delphi. Overall, 212 members were invited to join the panel, 122 of whom participated in the first round, and 104 of whom participated in the second round. To ensure a balance (equal weight) among the four groups of represented stakeholders, the authors obtained a weighted average for each cluster by calculating the cumulative distribution function of each subset.

The reality in accounting studies is the same, and there is variation in the number of members surveyed. Worrell, Di Gangi, and Bush, (2013) summarize 16 studies in the area of accounting information systems that used the Delphi technique and were published between 1987 and 2011. They were between 9 and 83 participants in those studies and the number of rounds was not analyzed. In Brazil, Cunha (2007) conducted a study with doctors of accounting that included 15 members and lasted three rounds.

Based on the literature (Adler & Liyanarachchi, 2011; Moizer, 2009), there were identified four stakeholder groups, referred to in terms of their roles in the publishing process: (1) publishers; (2) referees; (3) authors; and (4) funding institutions (the bureaucrats). These activities were considered for purposes of identifying the respondents' profiles. Considering that there are researchers involved in all these activities, the experts panel from the faculty of Brazilian graduate programs in accounting was chosen because many of the researchers who belong to these programs frequently perform the above-mentioned activities. Preliminarily, the study was designed to have three rounds.

The information-gathering process to form the list of invited participants included the following steps:

(a) The CAPES website was accessed to identify Graduate Academic Programs (PPGs) in accounting, as recognized/recommended by the regulator. Given their characteristics, professional programs[8] were disregarded. This process identified 23 academic programs in accounting (Appendix D), labeled in Brazil as programs in accounting science, accounting, or controllership (September 2015);

(b) Searches on the Plataforma Sucupira/CAPES (http://sucupira.capes.gov.br/) were undertaken to identify all the faculty members of the selected programs (September 2015). Three hundred and twenty-seven participants in these PPGs, including both permanent and collaborating professors, were identified. Visiting professors were excluded because of the temporality of the bond in the program (sometimes those professors are foreign specialists who teach a short-term class and do not maintain a constant presence in the program). Eight duplicates were found (bonds with more than one PPG), along with the present study's advisor. Both were excluded from the list, paring the research subjects down to 318 specialists;

[8] In Brazil, there are Professional Masters programs, but they are less academic and more technological in nature. For example, although theses in these programs can follow the traditional model, they can also be a process, piece of software, product, etc., with a focus on innovation. In this study, only the academic programs, awarding Master of Science and Doctor of Science (Ph.D.) degrees, were considered.

(c) Each researcher's email address was identified from the official website of the applicable graduate program. Alternatively, when that information was unavailable, an attempt was made to find it on the institutional website of the academic unit (department or college) to which the professor is connected. When these attempts failed, an attempt was made to find the e-mail address in recently published journal articles authored by the researchers, using the Google Scholar tool (http://scholar.google.com).

After identifying the e-mail addresses, an invitation letter (Appendix E) was e-mailed to the 318 researchers. An answer was requested within 11 days, which was extended by another five days via a reminder e-mail. To increase the possibility of securing a larger number of respondents, the coordinator of the University of Sao Paulo's graduate program in accounting kindly sent an email to the managers of the other 22 academic graduate programs in accounting at the advisor's request (Appendix F). The faculty members in those programs were selected to comprise the list of participants, and the purpose of the message was to ask the coordinators to encourage them to participate in the expert Delphi panel. After the messages, 74 people responded positively, confirming their willingness to participate on the experts' panel. Only respondents to this initial message received the second e-mail with instructions for participation in the first round. Considering their experience, it is important to observe that participants may or may not be senior researchers.

3.3.2 Instrumentation: preparation of data collection instruments

Preparation of the data collection instrument for the first round was performed by organizing the propositions in key features presented in the theoretical platform. According to the foundations of the Delphi Technique, the data collection instrument was made available electronically and the answers were controlled by the researcher. Because the first round's questionnaire was built beforehand based on the literature, this study can be considered to have used the modified Delphi Technique.

Studies present various scales of measurement for the data collection instruments, such as a six-point Likert scale (Kelbaugh, 2003) or a 1-10 numerical scale (Rodríguez-Mañas et al., 2013; Tetzlaff et al., 2012). Other studies combine various metrics – such as the percentage of respondents who agreed with a certain element in the first round – with note assignment on a numerical scale in subsequent rounds (Cunha, 2007). Elwyn et al. (2006) use a 9-point scale (1 to 9) for the panelists to assess a set of 80 criteria regarding quality

judgment. For this study, the numeric layover of 10 points (1-10) was chosen in light of the belief that it might be more efficient to assess the propositions quantitatively. The instruments from the other rounds derive from the first round and include new elements that might have emerged from the experts' suggestions.

A pre-test of the tool for the first round was performed to identify possible inconsistencies that might hinder the understanding of the propositions. The pre-test step was performed by submitting a preliminary version to be evaluated by five people: 3 Ph.D. students in Accounting, 1 Ph.D. in Accounting, and 1 Ph.D. in Education, who made many comments to improve the inquiry. The questionnaire was available electronically on the online SurveyMonkey platform (http://www.surveymonkey.com). As a result, the instrument was fixed/adjusted and prepared to receive the respondents' answer online.

The final version of the first round's instrument was written entirely in Portuguese (Appendix G), with 53 statements (distributed in nine key categories) and 14 questions used to characterize the respondents. Additionally, there was one open space for comments and suggestions. The first round was electronically organized into 5 pages as follows:

- First page: informed consent (Appendix H);
- Second page: propositions regarding research design, research subjects, and data collection;
- Third page: propositions regarding data analysis, findings, and reporting;
- Fourth page: propositions regarding reflexivity and neutrality, ethical issues, and auditability; and
- Fifth page: questions that characterize the respondents and one free space (open question) for comments and suggestions regarding the research (optional).

Cronbach's alpha was calculated using the statistical package SPSS, version 23, to measure the reliability of the instrument based on the cases and items (propositions) of the first round. The results show that the instrument achieved a Cronbach's alpha level of 92.2%, which is important evidence of reliability. Only the open question had an optional answer in the instrument.

The second instrument was based on the first, but with fewer statements (28 propositions) distributed over 3 pages, including the informed consent and the open space for comments. Although a three-round Delphi Technique has been proposed in the design, the results showed that it was not necessary to apply the third round because a saturation of the standard deviation of the propositions was identified between rounds one and two.

3.3.3 Measurement of consensus and classification of agreement levels of evidence

The literature demonstrates various ways to measure consensus using the Delphi Technique, such as the frequency of answers to the most important items, the score (Deckers et al., 2015), median (Tetzlaff et al., 2012), 'equimedian' rates per expert group, differences between groups (F and p-value) (Elwyn et al., 2006), mean and consensus level (%) (Jones & Xiao, 2004; Kim et al., 2013), and score, mean, median, and variation coefficient (Cunha, 2007). Other examples include when "80% of the ratings fell within two categories on the six-point Likert-type scale" (Kelbaugh, 2003, p. 153), and the percentage of answers rated ≥ 8 or ≤ 3 on a 10-point scale (Rodríguez-Mañas et al., 2013).

There are also various possibilities with respect to the classification of criteria into agreement levels. An example shows that one way to sort the results in terms of agreement or disagreement is to use labels such as "strong agreement (> 80% of answers rated ≥ 8 or ≤ 3), moderate agreement (70% – 80%), low agreement (50% – 70%), and no agreement (< 50%)" (Rodríguez-Mañas et al., 2013, p. 63). Another model considered that "participants 'disagreed' if 30% or more of the ratings were in the lower third (ratings 1-3) and 30% or more of the ratings were in the upper third (ratings 7-9)" on a 9-point scale (Elwyn et al., 2006, p. 3). These authors regarded "quality criteria with an overall equimedian rating of 7 to 9 (without disagreement) as 'important' and included them". Additionally, they considered "criteria rated as 4 to 6 (without disagreement) to be 'equivocal' and criteria rated with an equimedian of 1 to 3 as 'not important'." (p. 3).

Based on these examples, some rules were established first to determine which statements individually could be considered to demonstrate consensus and then to choose the propositions for the next round:

- 75% or more of the respondents scored the proposition with grades of 1-3 or 8-10 (strong agreement);
- 25% or more of the respondents scored the proposition with grades of 1-3 (disagreement);
- Variation of standard deviations ≤ 25% between rounds.

These percentages were chosen because they involve equal intervals in the statistical quartiles, representing a more careful consideration than the use of arbitrarily defined rates.

Considering the criteria for the classification of the results by levels (Rodríguez-Mañas et al., 2013), an adjustment was made in the bands to line up with the 75% figure that

had been chosen to characterize the previously defined consensus. Next, rules were created to group the final results using the percentage of the final score, as shown in Table 11.

Table 11:
Criteria for classifying the statements in accordance with agreement levels.

Rodríguez-Mañas et al. (2013)	Adjusted model	Agreement level
>80% of answers rated ≥ 8 or ≤ 3	>75% of answers rated ≥ 8 or ≤ 3	Strong agreement
70% – 80%	65% – 75%	Moderate agreement
50% – 70%	50% – 65%	Low agreement
<50%	<50%	No agreement

Sources: Rodríguez-Mañas et al. (2013), and adapted by the author.

The percentage of the final score was calculated based on the ratio of the sum of the grades obtained and the maximum possible score (number of responses multiplied by 10). For example, if a proposition had 41 valid answers, the maximum score is 410 because 10 is the highest score on the 10-point scale used by the data collection instrument. Supposing that the total score of that proposition is 380, the percentage obtained is 92.68% (380/410).

3.3.4 Application of the first round

Before beginning data collection, the present study generated a password and 74 customized codes for the participants who had confirmed their availability. Each respondent's code was sent in the same individual e-mail that contained the password (Appendix I). The code definition was generated using an alphanumeric pattern that followed the model "delphiXXX"; the final three digits were a logical numerical combination to mitigate the risks of possible participation by another respondent whose data entry was incorrect (i.e., an interleaved – as opposed to an exact – sequence). In practice, in addition to allowing answers to be linked among the three rounds, the respondent code worked as a new password because it was specific to each participant. In the event of an error that would not permit the linkage of responses to a given member throughout all the stages, that member’s answer was assessed to decide whether to discard it, thus avoiding any compromise of the quality of the Delphi in the final round. Control over the respondent's codes was the personal task of the researcher in charge, who made a restricted-access spreadsheet (with a PIN) to preserve the confidentiality and privacy of that information.

When visiting the Internet address/link provided by e-mail to begin the first round (http://pt.surveymonkey.com/r/QualiPPGCCFEAUSP), the expert was required by the opening screen to complete the password field, as shown in Figure 6:

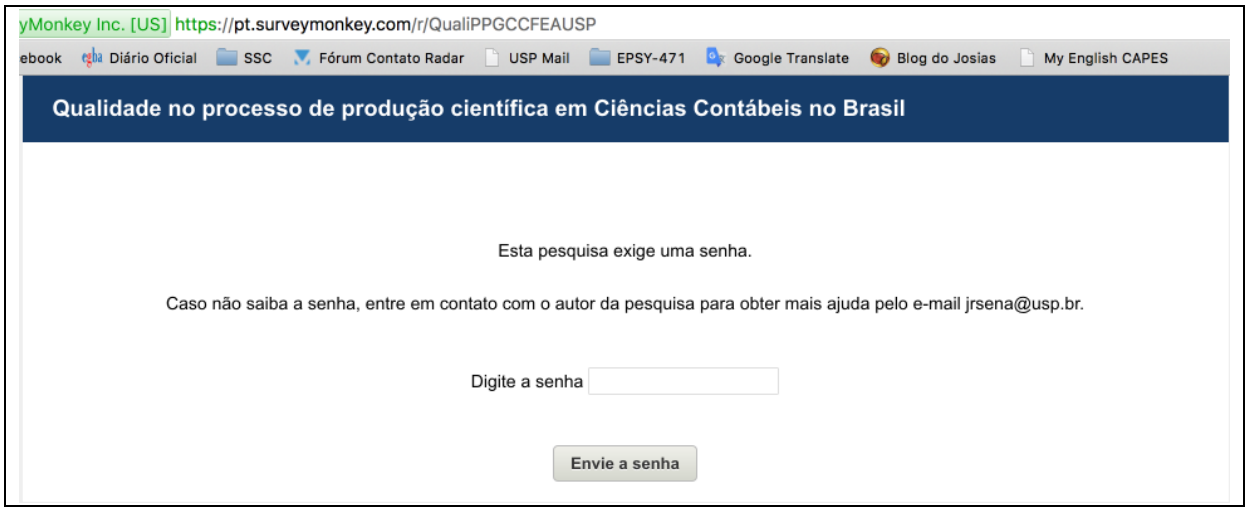


Figure 6: Initial screen to access the data collection instrument.
Source: data collection instrument on the SurveyMonkey.

The password for access, along with the customized respondent's code, had been sent through individual e-mails to each faculty member who responded positively regarding his or her availability to participate on the Delphi expert panel. After inserting the password, the first page of the collection instrument, which contained the informed consent form, was opened (Figure 7).

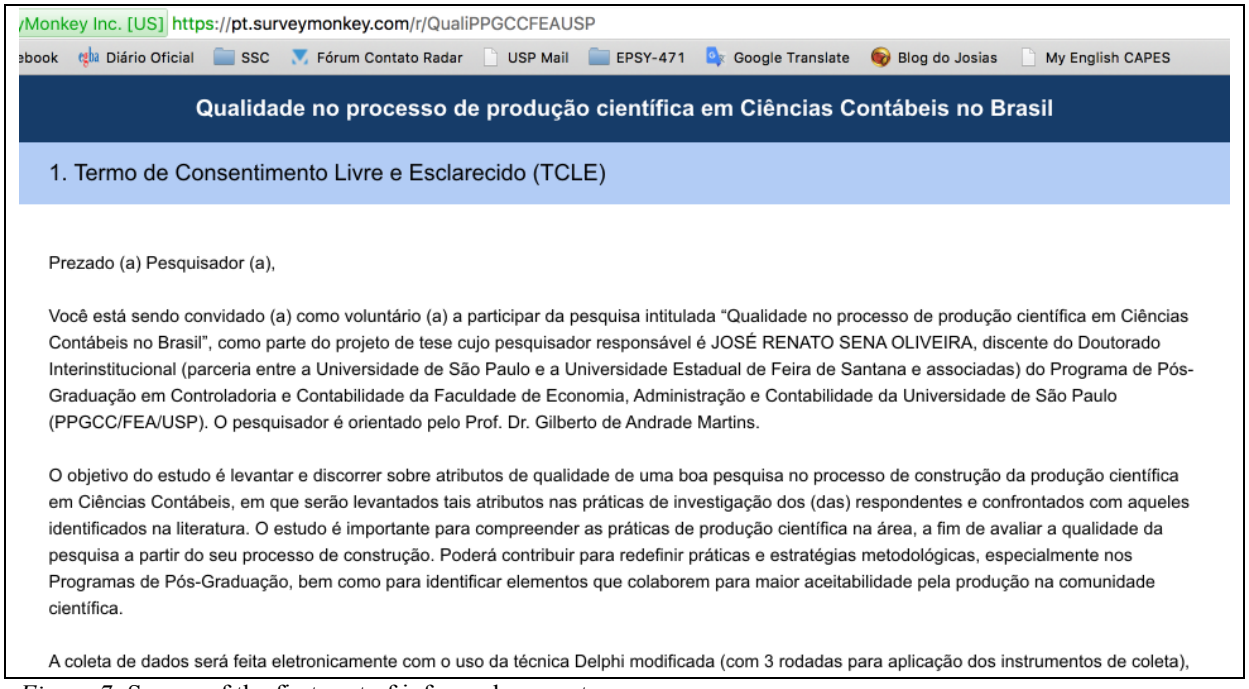


Figure 7: Screen of the first part of informed consent.

Considering the principles of respecting and affording privacy to the respondents, the informed consent was shown on the first page of the online instrument. In addition to the

informed consent, the page contained two mandatory fields in which to insert the participant's code and to express agreement with the informed consent. This procedure is justified by the need to allow connection (matching) between the respondent's answers in the three rounds and to avoid the inclusion of any personally identifiable information in the database. The bottom of the page also included a recommendation that a copy of the informed consent should be saved after completing the first page's fields that contained the explicit agreement and the respondent's code (Figure 8).

Following the informed consent, participants found the statements shown in Figure 9 on the next page; the answers were mandatory. For each statement, the expert had to assign a grade on a 1-10 scale, considering the adhesiveness level of the statement in relation to his/her research practices.

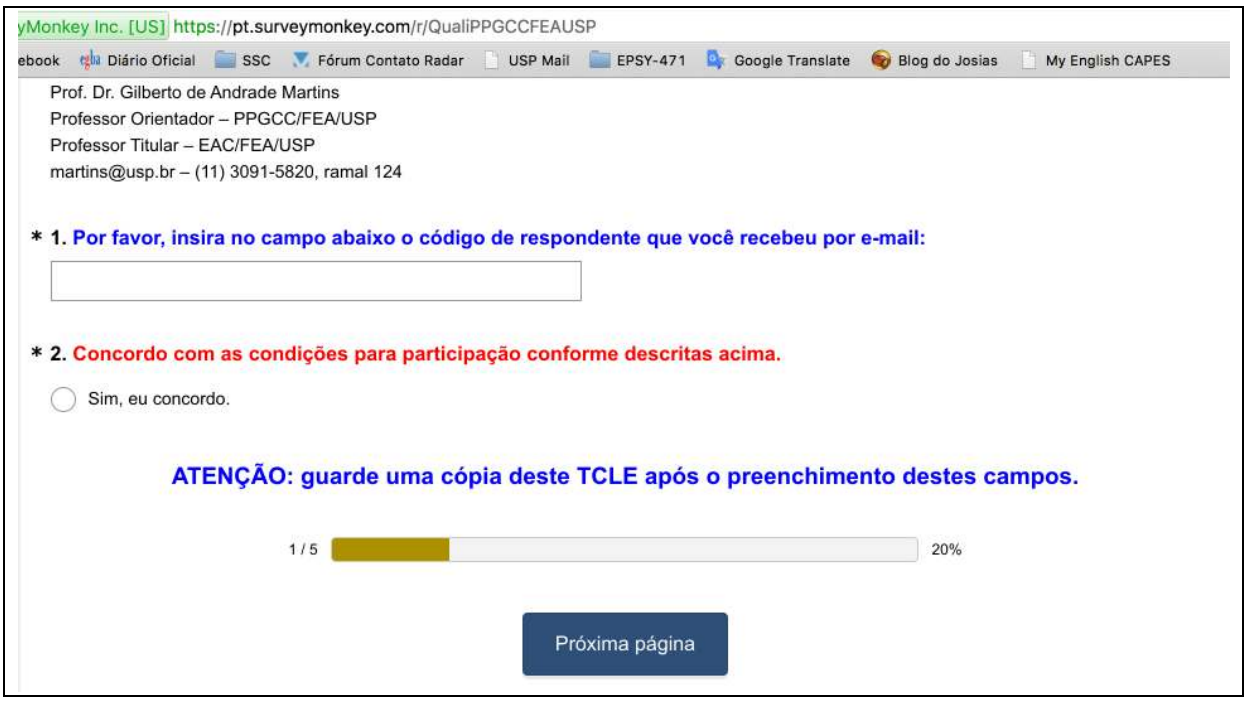


Figure 8: Screen with fields to insert the respondent's code and to declare agreement.

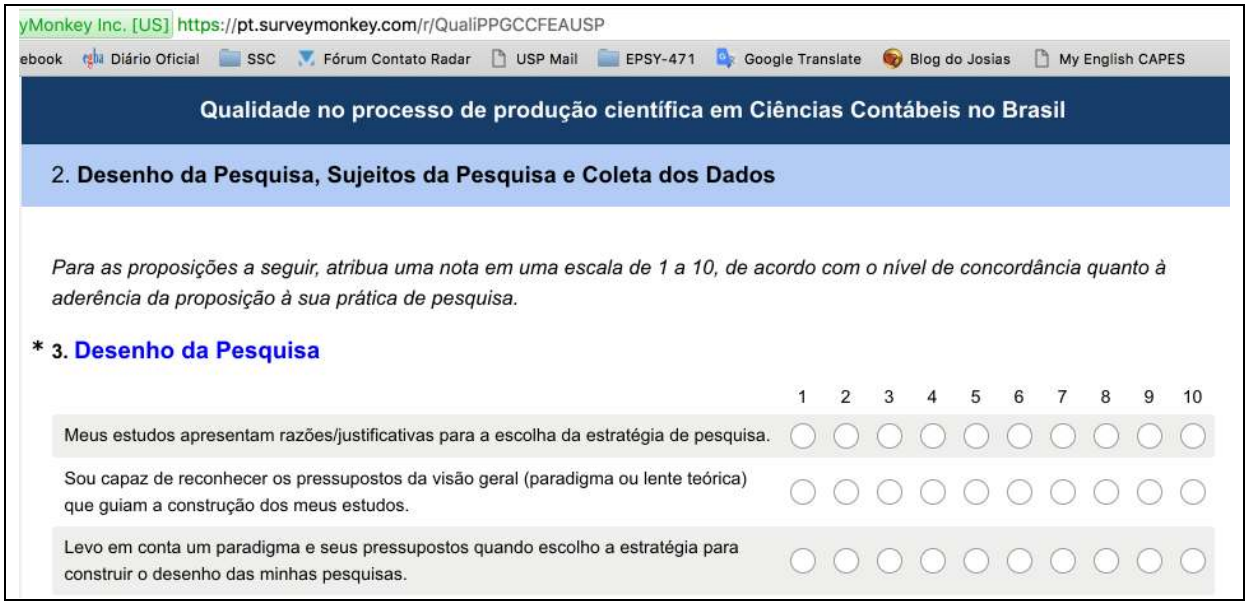


Figure 9: Second page of the first round's tool (research design, research subjects, and data collection).

A customized e-mail with instructions for participating in the first round (Appendix I) was sent to each participant that provided a positive answer to the initial invitation. The deadline to finish this step was 12 days, which was extended for another five days via a reminder e-mail. In this first round, of the 74 experts who were willing to participate, 42 completed the entire questionnaire.

3.3.5 Application of the second round

Before beginning the second round, a customized report for each participant (Appendix J) was prepared that included the new instrument's propositions. To facilitate participation in the second round, the report also included the grades that the participant awarded to those statements in the previous round along with the median given by the group of respondents. The reports were generated as .pdf documents.

As mentioned in the section regarding the preparation of the data collection instruments, the second-round tool was organized into 3 pages, which contained informed consent (Figure 10), 28 statements organized by key features (Figure 11), and an open question. These items are holdovers from the first instrument. In addition, new elements were included. First, at the suggestion of some participants, the second-round tool included the option "never did/do not do research involving human beings" in the groups pertaining to data collection, ethical aspects, and auditability (Figure 12 and Appendix K). The purpose of including this option was to decrease possible biases in the analysis of certain propositions

because there are experts whose research practice does not involve human beings as research subjects. This suggestion was also taken into account in the analysis of the data collected during the first round. Second, the concept of "research involving human beings" displayed in Resolution CNS n. 466/12 (Brasil, 2012) of the National Health Council was included. The electronic version of the second instrument was available from a SurveyMonkey link/address (http://pt.surveymonkey.com/r/DelphiQuali2).

To begin this round, the respondent's code and password were re-sent by e-mail (Appendix L). Additionally, the customized report was attached. Only the 42 respondents who participated in the first round received the message in this step. The instructions were similar to those of the first round, although the second-round instrument was shorter, so the deadline was only eight days later. However, access remained open for seven more days.

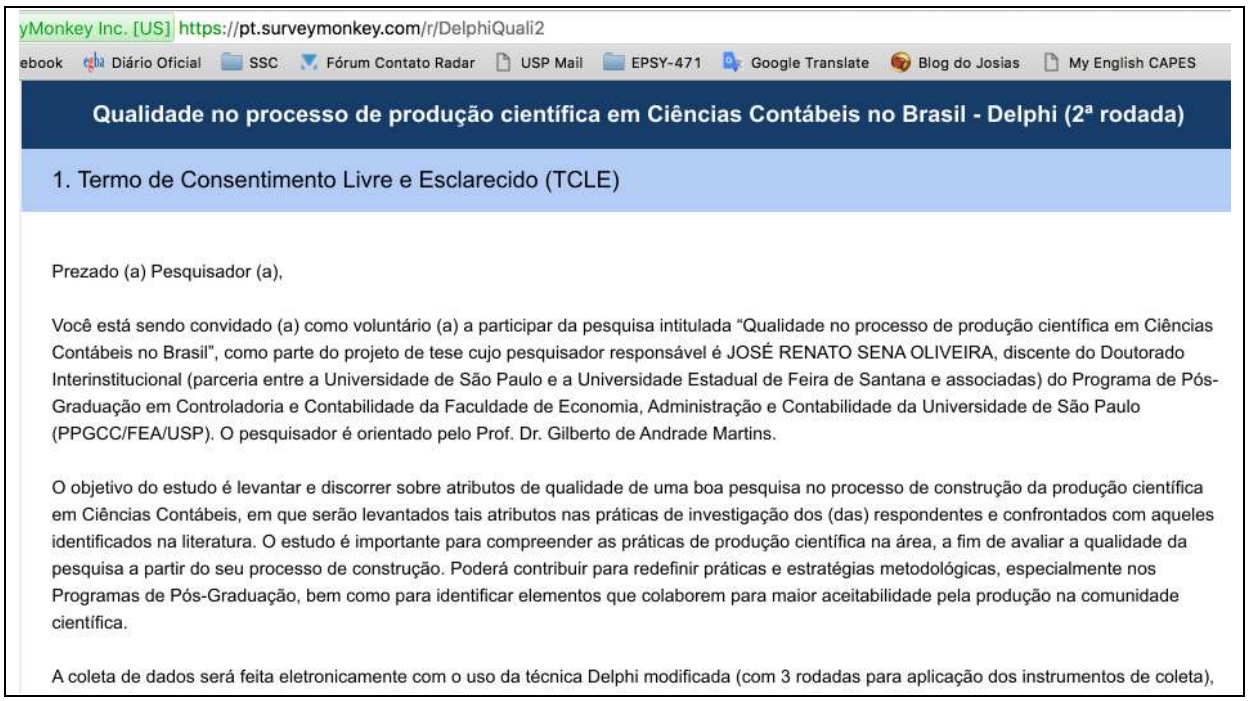


Figure 10: Second round's informed consent.

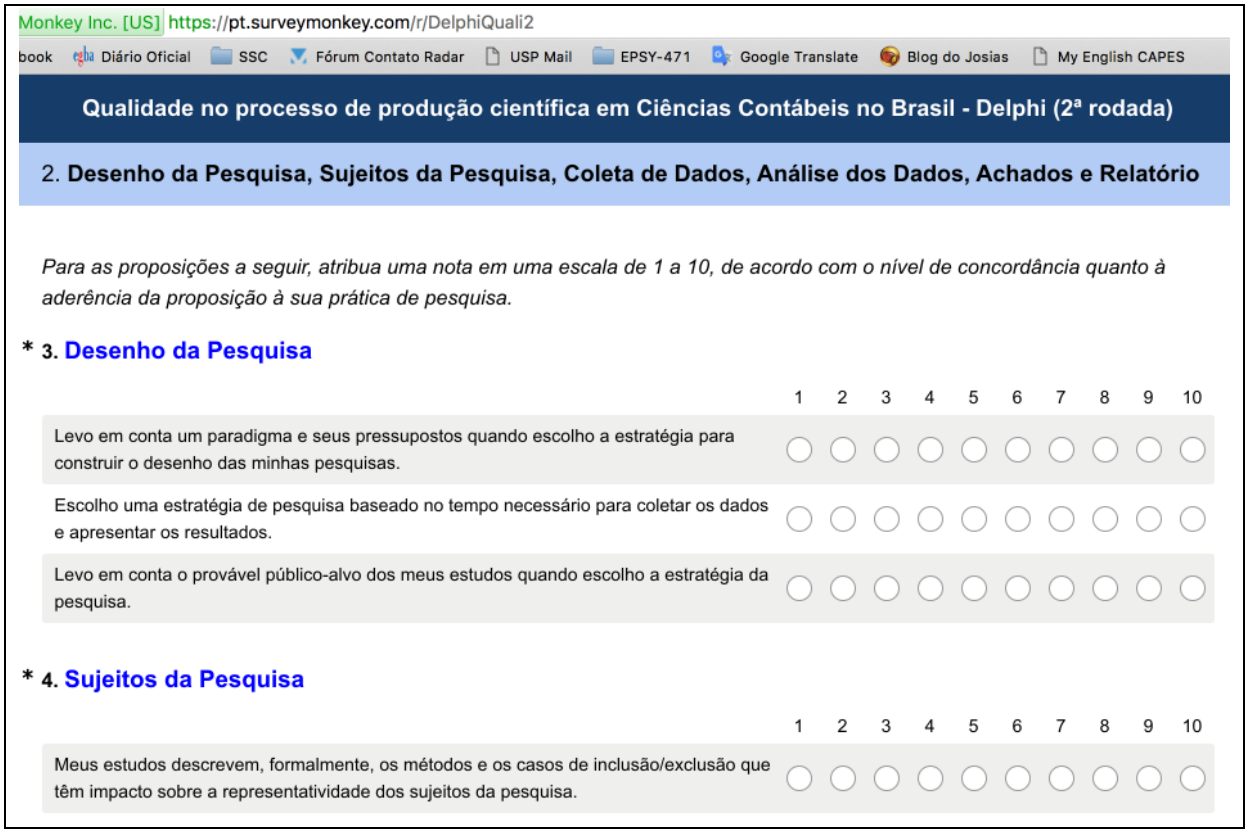


Figure 11: Second page of the online data collection instrument (second round).

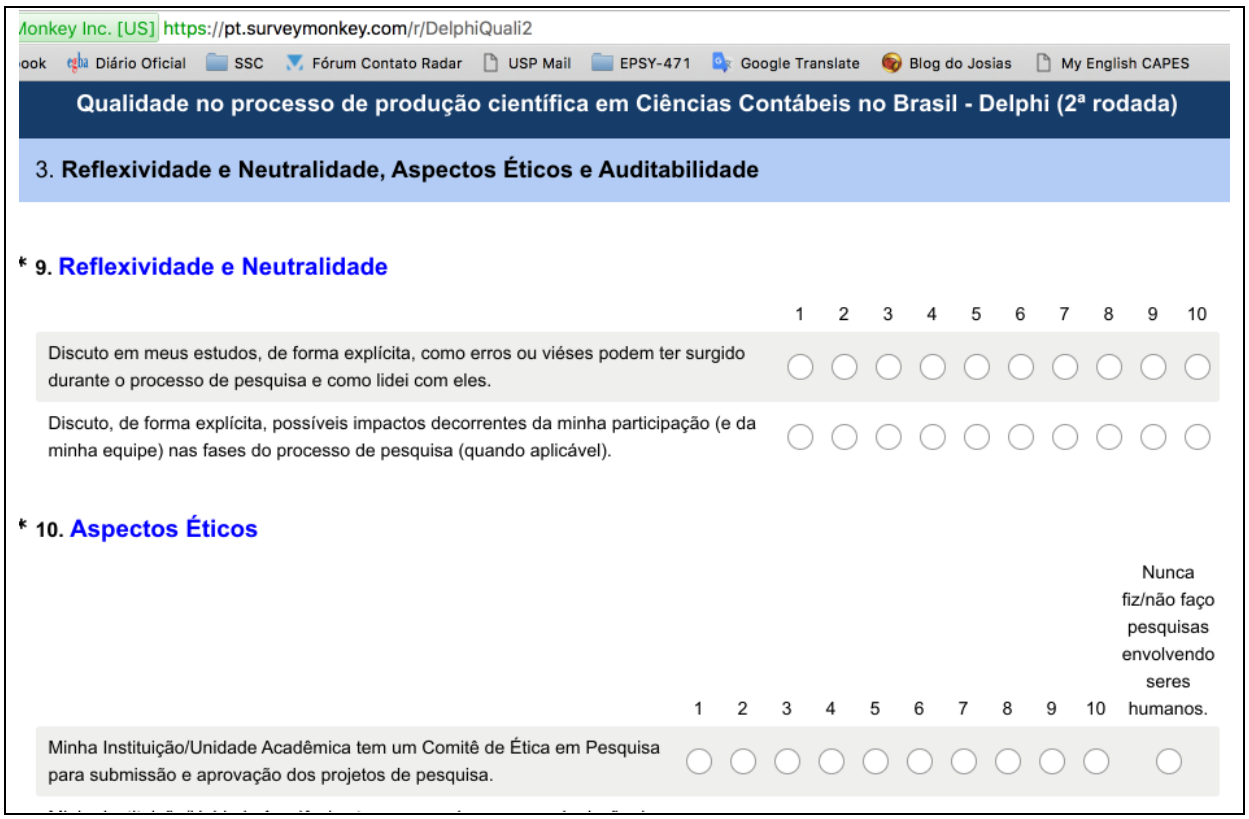


Figure 12: Third page of the second round's tool with the new alternative of response.

As with the first round, the results of the analysis consider the dispersion measures frequently used in the literature to evaluate the results (mean, median, and standard deviation). Additionally, to evaluate the consensus, the measures used are those detailed in Topic 3.3.3. There was one dropout during the second round. For ethical reasons and upon request, the data related to that expert's first-round answers were excluded from the database, leaving 41 participants in that step. In the end, 37 experts participated in the final step.

After the second round, few statements failed to meet the consensus criteria used in the study. No third round was convened because the number of propositions for which consensus was pending represented less than 15% of the analyzed propositions, and these divergences were discussed in the work. The data collection was completed in December 2015.

Considering the set of general criteria that have been chosen to analyze the research practices in this study, the final list of statements was distributed, as represented in Figure 13.

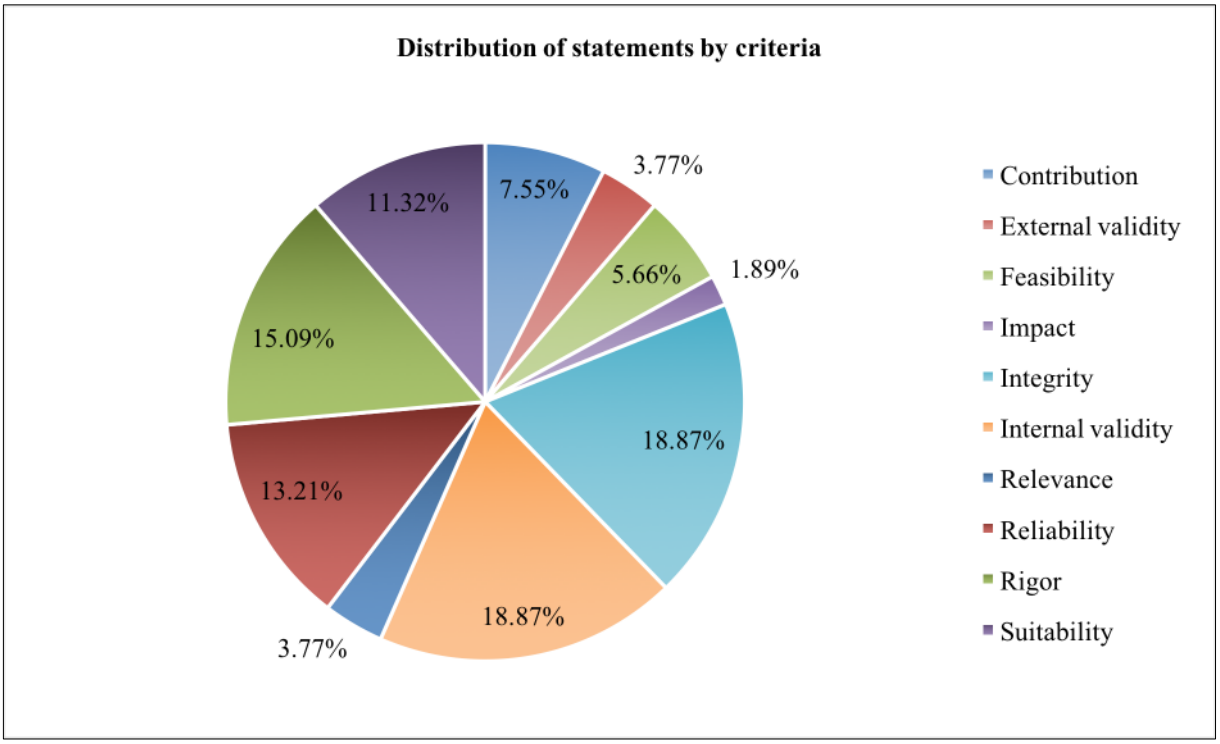


Figure 13: Distribution of statements by criteria.

Because the study's focus was trained on research practices, most of the propositions are primarily aimed at analyzing criteria such as integrity and internal validity/credibility or defensibility.

3.4 Stage 3: Comparing the Literature with the Findings

This step compares the literature review on research quality with the study findings. After identifying the attributes of good research in the literature (Step 1) and generating the outcomes of the respondents' consensus with the Delphi Technique (Step 2), the evidence was compared to identify possible variations (Step 3).

Divergences and convergences were assessed against the criteria used to measure both consensus in the data collection instruments and their agreement level, after accounting for their adherence to the respondents' research practices. To complete this stage, the propositions were classified into groups based on degree of agreement, as explained in section 3.3.3. Additionally, the divergences that arose in some statements for which consensus was not formed after the second round were discussed.

The purpose was to attempt to determine similarities and differences in the researchers' practices to provide support for building an approach to judge research quality.

3.5 Stage 4: An Approach to Judge the Quality of the Research Process in Accounting

The development of the approach to judging research quality sought to include an analysis of the building process of scientific production, considering both the literature and the elements that emerged from the Delphi consensus. This approach focuses on issues that can address researchers' practices to mitigate problems in developing field studies and is an attempt to present a set of quality criteria and create an association between those criteria and various stages of the research process. To do so, domains beyond those of the investigation (Brinberg & McGrath, 1985), key features (Mays & Pope, 2006; Spencer et al., 2003) were considered.

To achieve this intent, the logic model used in the evaluation process, projects and programs (Taylor-Powell & Henert, 2008) was selected to systematize the approach. This technique is based on a logical structure driven by a summarized model with inputs, outputs and processes that can be used to evaluate both outcomes and processes. Evaluation of processes is consistent with this study because it focuses on analyzing the scientific production process in the accounting field. The structure of this logic model is shown in Figure 14:

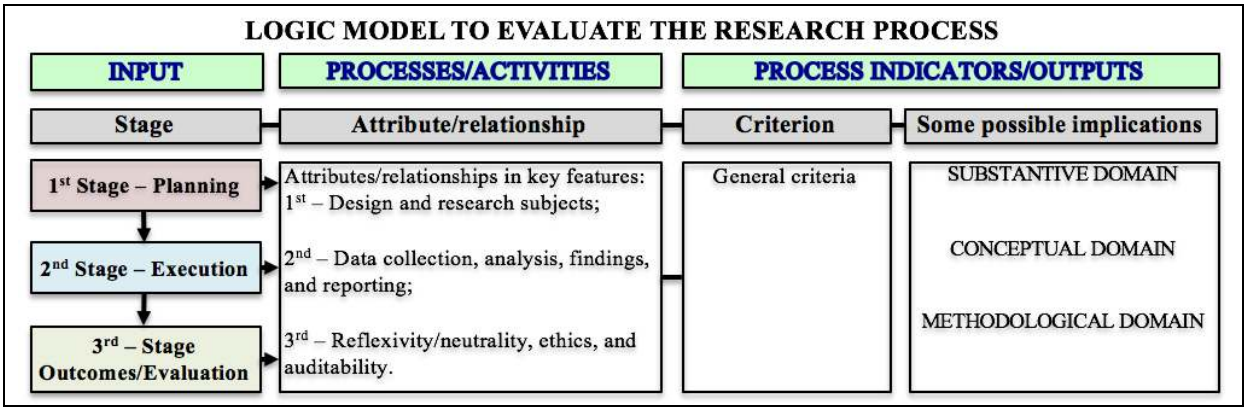


Figure 14: Summarized logic model to evaluate the research process.
This logic model was created by adapting examples provided by Prof. Thomas Schwandt during EPSY-470 - Introduction to Evaluation Theory lectures in the graduate program of the Department of Educational Psychology at the College of Education at the University of Illinois at Urbana-Champaign (UIUC), in September 2014. At that time, the author was auditing the graduate program as a visiting scholar at UIUC.

According to Morell (2010), a logic model of evaluation aims to understand the relationships among a program's activities, its consequences, and its environment. From the perspective of a formative evaluation, this model seeks to analyze what a program needs so that it can work as expected. Research practice is closely linked both to the graduate programs and to the tripod of university action (teaching-research-extension).

CHAPTER 4 –
RESULTS AND DISCUSSION

This chapter presents the results and discussion of data collection. With respect to the first topic, it shows the characteristics of respondents and the graduate programs in which they are faculty members. Next, the results of the implementation of two rounds of modified Delphi are displayed, with attributes/relationships linked to research quality criteria according to the key feature (or stage of research process) to which they refer. In sequence, attributes/relationships are grouped based on the respondents' level of agreement as to the final result. Finally, an approach based on attributes of a good research is presented to discuss the application of quality criteria to research practices in accounting.

4.1 Characterization of Respondents

To understand the profile of the 41 respondents, they were asked about some of the characteristics of their professional qualifications and experience. The findings revealed that 65.84% have a doctorate (Ph.D.) in accounting and that 70.73% of those respondents earned their doctorates in a different national institution from the one in which they work. Additionally, it was found that 85.36% of respondents did not participate in exchange programs (as visiting scholars) during graduate school, and 78.05% had not performed post-doctorate work. The data also revealed that 9.76% of respondents were involved in foreign-exchange programs during their doctorates, and 14.63% did their postdoctoral studies in an international institution. These results suggest that the respondents have little experience in international activities, which may be reflected in a low level of insertion into international research networks.

Another important indicator is the endogeneity analysis in training faculty members. The findings revealed that only 9.75% of respondents obtained their doctoral degrees in the same program in which they work. According to the criteria used by CAPES in the triennial evaluation, programs with low endogeneity are considered higher quality. Additionally, programs with faculty members whose educational tracks are highly diversified tend to have researchers with different skills and worldviews, which may contribute to the emergence of more comprehensive, innovative studies that involve the combination of those researchers' various experiences. Finally, the length of the respondents' experience in graduate programs

was analyzed. The findings showed that most of the respondents (34.14%) have up to two years of experience as faculty members in this type of program. However, 26.83% have between 5-7 years of experience and 19.51% have more than ten years, thus revealing heterogeneity in the distribution of respondents related to this feature. The details are shown in Table 12.

Table 12:
Characteristics of the respondents' professional qualifications and experience.

Aspect	Category	*n*	%
Field of his/her doctorate (or Ph.D.).	Accounting	27	65.84
	Business/Finance	5	12.20
	Economics	3	7.32
	Education and technology	1	2.44
	Industrial engineering	5	12.20
	Total	41	100.00
Where did he/she earn the doctorate (Ph.D.).	In the same program in which I work	4	9.75
	At the same institution, but in a different program from the one in which I work	5	12.20
	At another national institution	29	70.73
	At another international institution	3	7.32
	Total	41	100.00
Exchange program (visiting scholar) during the doctoral degree	Yes, at another national institution	2	4.88
	Yes, at an international institution	4	9.76
	No, I did not	35	85.36
	Total	41	100.00
Postdoctoral work	Yes, at another national institution	3	7.32
	Yes, at an international institution	6	14.63
	No, I did not	32	78.05
	Total	41	100.00
Time in graduate program (in years)	0-2 years	14	34.14
	3-4 years	4	9.76
	5-7 years	11	26.83
	8-10 years	4	9.76
	More than 10 years	8	19.51
	Total	41	100.00

Given that the respondents are members of an expert panel to analyze research practices, it is relevant to learn about their experience in these activities. The first two points analyzed concern the ability to raise funding for research. The findings revealed that 56.1% of respondents currently have projects funded by development institutions, and 24.39% have research productivity grants, as shown in Table 13. The receipt of research funding and productivity grants are indicators of the researcher's expertise in the relevant knowledge production in his field because funding opportunities are offered through public calls, and the researcher must be able to compete with other researchers both from other areas and from different institutions.

The following items, also shown in Table 13, are related to the respondents' experience as part of the scientific publishing process. The data revealed that 43.9% of the respondents either work or have worked as editor of a journal ranked in the Qualis/CAPES strata (except Qualis C). Moreover, 63.41% are or have been a member of the editorial board of a journal classified in these same strata. Finally, 100% of the respondents act or have acted as referees for a journal that meets that same requirement. These data reveal that the panel participants have extensive experience related to the publishing process.

Table 13:
Respondents' research experience

Category	Funding for research		Productivity grant		Journal editor		Member of editorial board		Scientific journal referee	
	n	%	*n*	%	*n*	%	*n*	%	*n*	%
Yes	23	56.10	10	24.39	18	43.90	26	63.41	41	100.00
No	18	43.90	31	75.61	23	56.10	15	36.59	0	0.00
Total	41	100.00	41	100.00	41	100.00	41	100.00	41	100.00

To complete the analysis of the features related to research experience, a specific item asked respondents about their productivity in publishing journal articles that are considered high impact by Qualis/CAPES (called superior strata: A1, A2, and B1) in the previous triennial assessment (2010-2012). The results showed that 34.15% of respondents have published between one and two papers in journals evaluated in the upper strata; another 34.15% have published between three and five papers at these same levels. Finally, 29.27% have published more than five articles during the period, as shown in Figure 15. Only one respondent (i.e., 2.44%) has had no article published during the triennial assessment.

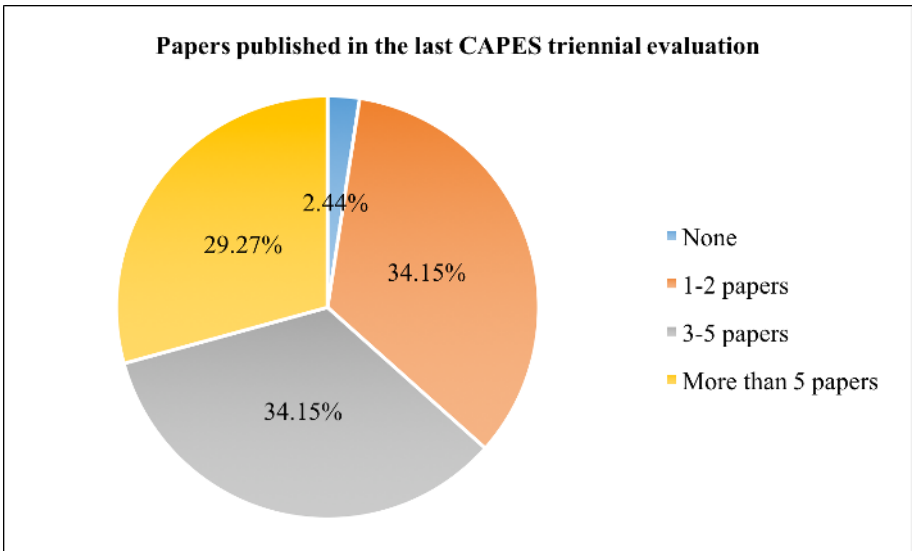


Figure 15: Distribution of the respondents by number of papers published in the last CAPES triennial evaluation.

With respect to the characteristics of graduate programs to which respondents are attached, Figure 16 shows that 53.66% are linked to programs that offer master's and doctoral courses (M.Sc. and Ph.D.), and another 46.34% are faculty members in programs that have only master's programs. Additionally, 19 of Brazil's 23 graduate programs in accounting are represented in the Delphi expert panel; ten of those programs offer both degrees, as revealed in Figure 17. Finally, Figure 18 shows that most respondents (17 experts) perform their activities in graduate programs rated as Grade 4 in the CAPES evaluation, which is equivalent to 41.46% of the respondents. Grade 4 is considered the minimum for a graduate program to be accredited to offer a doctoral degree.

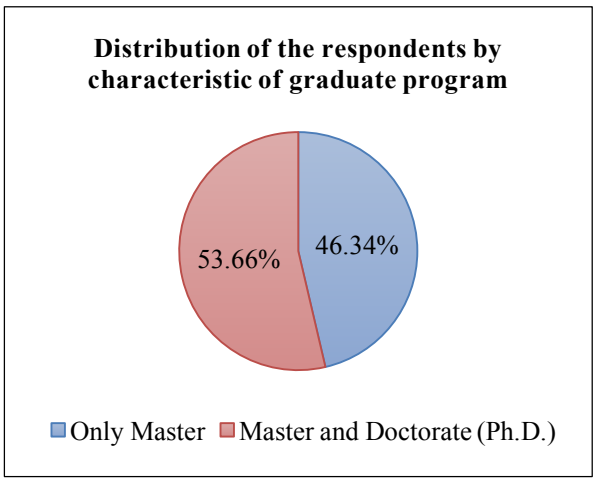


Figure 16: Distribution of the respondents by the characteristics of their graduate programs.

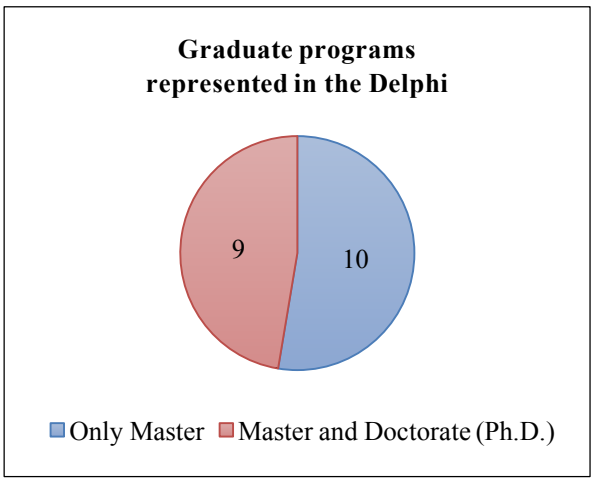


Figure 17: Distribution of the graduate programs represented in the Delphi by level.

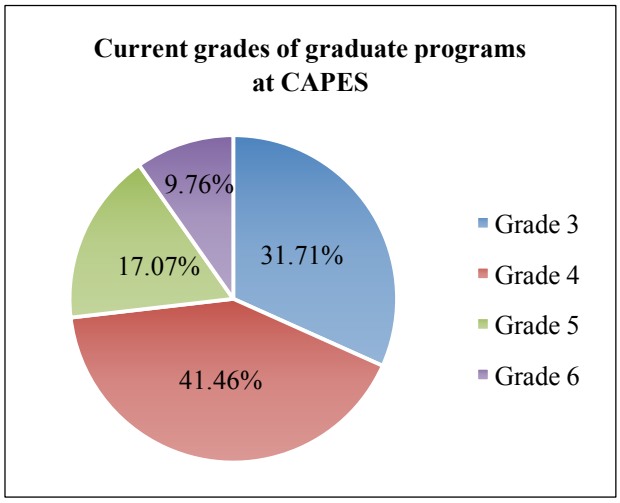


Figure 18: Distribution of the graduate programs by their current grades (CAPES).

Source: research data, 2016.

Another important finding was the low concentration of the respondents in the same program: the highest rate was equivalent to 12.20% (5 experts).

4.2 Application of the Delphi Technique

As detailed in Chapter 3 (Methods), the technique used for data collection was the modified Delphi, characterized by the submission of preliminary statements based on the literature during the first round instead of entries based on the respondents' suggestions. The source of the findings presented in Tables 14-36 is the inquiry undertaken in two rounds of the Delphi. The tables are ordered based on nine key features (or stages of the research process) to which they relate and are shown in the same order as in the data collection instrument. Tables 14-32 are sorted in descending order based on the highest percentage of grades equal to or exceeding 8 (the sixth column of the tables). The inquiry consisted of 53 statements distributed in nine key features. Forty-one valid responses were obtained in the first round and 37 were obtained in the second round.

The codification of the propositions in the tables uses seven characters: the first two characters are related to the Delphi round, the next three characters are the first three letters of the word that defines the key feature, and the final two numbers refer to the position of the statement in the data collection instruments. For example, the code R1Des04 (the first of Table 14) means that the proposition refers to the first round, the research design is the feature, and the proposition is the fourth one in the questionnaire.

4.2.1 Results of the first round

The initial data for the first round of the Delphi are disclosed in Table 14, which presents the results of statements related to the research design and primarily involves the general criteria of feasibility and suitability.

Table 14:
First round – research design (descending order of percentage of grades ≥8)

Propos.	Attribute/relationship	Grade ≤3		Grade ≥8		Min	Max	Mean	Median	SD	Consensus
		n	%	n	%						
R1Des04	Goal/problem shown precisely	0	0.00	41	100.00	8	10	9.63	10.00	0.54	Yes
R1Des01	Useful strategy for purpose	1	2.44	39	95.12	1	10	9.00	9.00	1.50	Yes
R1Des02	Clear overview that guides the study	1	2.44	36	87.80	3	10	8.73	9.00	1.61	Yes
R1Des06	Arguments for technique choices	1	2.44	36	87.80	3	10	8.93	9.00	1.47	Yes
R1Des05	Overview x theory	2	4.88	34	82.93	1	10	8.51	9.00	1.93	Yes
R1Des07	Access to the data	0	0.00	33	80.49	5	10	8.56	9.00	1.40	Yes
R1Des03	Overview x strategy	3	7.32	29	70.73	1	10	7.88	9.00	2.32	No
R1Des08	Time constraints	8	19.51	21	51.22	1	10	6.44	8.00	2.92	No
R1Des09	Research strategy x target audience	6	14.63	19	46.34	1	10	6.76	7.00	2.55	No

Note: in this round only the first two rules for building consensus presented in Topic 3.3.3 were considered, given that the third rule applies only to the comparison of the standard deviations between different rounds.

The findings revealed that ‘goal/problem shown precisely’ was the item with the highest percentage, reaching 100% of answers with a score equal or greater than 8. However, five of the nine propositions had at least one answer with a minimum score of 1, showing a dispersion over the maximum range. Six items are also verified to reached consensus based on the criteria explained in Topic 3.3.3. The results of the statements related to research subjects and data collection are displayed in Tables 15 and 16.

Table 15:
First round – research subjects (descending order of percentage of grades ≥8)

Propos.	Attribute/relationship	Grade ≤3		Grade ≥8		Min	Max	Mean	Median	SD	Consensus
		n	%	n	%						
R1Sub01	Criteria to design/select subjects	1	2.44	32	78.05	3	10	8.59	9.00	1.53	Yes
R1Sub02	Representativeness of subjects	2	4.88	30	73.17	1	10	8.15	9.00	1.99	No

Table 16:
First round – data collection (descending order of percentage of grades ≥8)

Propos.	Attribute/relationship	Grade ≤3		Grade ≥8		Min	Max	Mean	Median	SD	Consensus
		n	%	n	%						
R1Col03	Voluntary participation of subjects	1	2.44	38	92.68	1	10	9.37	10.00	1.65	Yes*
R1Col01	Notes for each research step	0	0.00	33	80.49	6	10	8.80	9.00	1.25	Yes
R1Col02	Notes for divergent events	0	0.00	30	73.17	4	10	8.44	9.00	1.57	No
R1Col04	Formal agreement from humans	8	19.51	21	51.22	1	10	6.90	8.00	3.28	No

*This was included in the second round because this proposition is linked with ethical aspects.

The issues included in Tables 15 and 16 are associated with the criteria of internal validity/credibility or defensibility, rigor/thoroughness, and integrity. The item "representation of the subject" did not form consensus in the group "research subjects". With respect to data collection items, two of them reached agreement; however, the "voluntary participation of subjects" was included in the second round because of the connection to ethical issues. Table 17, in turn, presents the statements regarding the data analysis that are linked to the criteria of integrity, relevance, reliability/auditability, and rigor/thoroughness.

Table 17:
First round – analysis (descending order of percentage of grades ≥8)

Propos.	Attribute/relationship	Grade ≤3		Grade ≥8		Min	Max	Mean	Median	SD	Consensus
		n	%	n	%						
R1Ana05	Significance of data to reaching aims	1	2.44	36	87.80	1	10	8.56	9.00	1.55	Yes
R1Ana02	Description of tools and procedures	0	0.00	35	85.37	6	10	8.73	9.00	1.12	Yes
R1Ana03	Context x impact on data analysis	0	0.00	34	82.93	4	10	8.32	8.00	1.35	Yes
R1Ana01	Description of nature & form of data	0	0.00	33	80.49	7	10	8.56	9.00	1.07	Yes
R1Ana06	Implicit/explicit links - find. x aims	0	0.00	33	80.49	5	10	8.51	9.00	1.27	Yes
R1Ana04	Other views to explain context	3	7.32	18	43.90	1	10	7.07	7.00	1.95	No

The findings revealed that most of the attributes and relationships of the data analysis stage obtained high acceptance rates, except for the item 'other views to know the context,' which is linked to the criterion of reliability/auditability and did not achieve consensus in the first round. The propositions included in Table 18, which relates to the findings, are associated with the criteria of contribution/quality of theoretical perspective, external validity or generalization/fittingness, internal validity/credibility or defensibility, relevance, and reliability/auditability.

Table 18:
First round – findings (descending order of percentage of grades ≥8)

Propos.	Attribute/relationship	Grade ≤3		Grade ≥8		Min	Max	Mean	Median	SD	Consensus
		n	%	n	%						
R1Fin01	Path to achieve conclusions	0	0.00	39	95.12	7	10	8.76	9.00	0.86	Yes
R1Fin02	Link between findings x evidence	0	0.00	39	95.12	6	10	8.85	9.00	0.94	Yes
R1Fin07	Check links of findings x purpose	0	0.00	39	95.12	5	10	9.10	9.00	1.04	Yes
R1Fin06	Previous findings x hypothesis	0	0.00	36	87.80	7	10	9.00	9.00	1.12	Yes
R1Fin04	New areas based on the findings	0	0.00	35	85.37	7	10	8.80	9.00	1.10	Yes
R1Fin03	Comparison results x other studies	0	0.00	34	82.93	5	10	9.10	10.00	1.37	Yes
R1Fin08	Does the context allow replication	0	0.00	34	82.93	6	10	8.56	9.00	1.25	Yes
R1Fin05	Insights for thinking about the field	0	0.00	30	73.17	6	10	8.27	8.00	1.10	No
R1Fin09	Impact from the nature of divergences	2	4.88	25	60.98	1	10	7.44	8.00	1.83	No

For these propositions, the findings showed high levels of acceptance in that three of the attributes and relationships associated with the internal validity criterion resulted in a higher percentage than 90%. Another detail to note is that eight of nine statements had minimal scores equal to or greater than 5, and only two items did not achieve consensus. Moreover, the item 'compare results with other studies' reached a median equal to 10, which shows the respondents' confidence regarding the applicability of that relationship in their research practice. Table 19 presents statements on reporting that are associated with the criteria of the contribution/quality of a theoretical perspective, external validity or generalization/fittingness, impact, internal validity/credibility or defensibility, relevance, and reliability/auditability. Table 20 also shares the integrity criterion and addresses issues of reflexivity/neutrality.

Table 19:
First round – reporting (descending order of percentage of grades ≥8)

Propos.	Attribute/relationship	Grade ≤3		Grade ≥8		Min	Max	Mean	Median	SD	Consensus
		n	%	n	%						
R1Rep06	Theory to support propositions	0	0.00	39	95.12	7	10	9.05	9.00	0.86	Yes
R1Rep05	Literature review x main concepts	0	0.00	38	92.68	6	10	9.10	9.00	0.94	Yes
R1Rep07	Conclusions x aim	0	0.00	38	92.68	6	10	9.20	10.00	1.10	Yes
R1Rep08	Explicit possibility of generalization	1	2.44	31	75.61	1	10	8.41	9.00	2.00	Yes
R1Rep02	Disclosure of limitations	0	0.00	29	70.73	5	10	8.29	9.00	1.33	No
R1Rep01	Discussion of impact on knowledge	0	0.00	28	68.29	4	10	7.90	8.00	1.14	No
R1Rep04	Boundaries of the study	1	2.44	24	58.54	1	10	7.49	8.00	1.61	No
R1Rep03	Reasons for limitations	0	0.00	22	53.66	4	10	7.46	8.00	1.53	No
R1Rep09	Appreciation of drafts by colleagues	6	14.63	16	39.02	1	10	6.54	7.00	2.67	No

Table 20:
First round – reflexivity and neutrality (descending order of percentage of grades ≥8)

Propos.	Attribute/relationship	Grade ≤3		Grade ≥8		Min	Max	Mean	Median	SD	Consensus
		n	%	n	%						
R1Neu01	How to address errors and biases	4	9.76	18	43.90	1	10	6.54	7.00	2.44	No
R1Neu02	Impact of team participation	5	12.20	15	36.59	1	10	6.22	7.00	2.33	No

The results shown in Table 19 demonstrate that three of the attributes/relationships that focused on the evaluation of the internal validity/credibility or defensibility criterion resulted in acceptance rates greater than 90%. In contrast, the statement 'Appreciation of drafts by colleagues,' which was associated with the same standard, presented an acceptance percentage of only 39.02%. Notably, five of the nine attributes in Table 19, along with the propositions related to reflexivity/neutrality in Table 20, also did not achieve consensus.

The next group of statements in the data collection instrument concerns Ethics. Although a portion of the items reached consensus, the respondents' comments on the final question revealed that some do not perform research that involves human beings; the propositions did not include a 'not applicable' option. Some researchers work with studies based on archival, for instance, for which characteristics such as submission to an Ethics Committee in Research and presentation of informed consent do not make sense. For this reason, the results might produce a biased analysis, so it was decided to resubmit all the propositions concerning ethical issues for assessment in the second round, while the data obtained in the first round were not presented.

The final group of statements, which is presented in Table 21, addresses auditability; these statements are associated with the general criterion of relevance/auditability.

Table 21:
First round – auditability (descending order of percentage of grades ≥8)

Propos.	Attribute/relationship	Grade ≤3		Grade ≥8		Min	Max	Mean	Median	SD	Consensus
		n	%	n	%						
R1Aud03	Safeguard databases for checks	0	0.00	36	87.80	4	10	8.95	9.00	1.40	Yes
R1Aud01	Records of design changes	4	9.76	29	70.73	1	10	7.68	8.00	2.61	No
R1Aud02	Records of the reasons for changes	4	9.76	25	60.98	1	10	7.29	8.00	2.63	No
R1Aud04	Guard documents to reduce risks	7	17.07	23	56.10	1	10	6.98	8.00	3.27	No

The findings demonstrated that only the safekeeping of databases for future checks achieved consensus. The other items were kept for retesting in the second round. Notably, 17.07% of respondents attributed a grade equal to or less than 3 for the item on the storage of study-related documents, which shows a weakness in the proper maintenance of research records for possible subsequent verification of research procedures.

4.2.2 Results of the second round

Statements that did not achieve consensus in the previous round, along with those related to ethical issues, comprised the data collection instrument in the second round. Overall, 28 propositions in nine primary characteristics (or stages of research process) were distributed in Tables 22-30 on this topic. Table 22 refers to items related to the research design and analyzes the criteria of feasibility and suitability. Table 23 shows only a statement regarding research subjects and its primary criteria: internal validity/credibility or defensibility.

Table 22:
Second round – research design (descending order of percentage of grades ≥8)

Propos.	Attribute/relationship	Grade ≤3		Grade ≥8		Min	Max	Mean	Median	SD	Consensus
		n	%	n	%						
R2Des03	Overview x strategy	1	2.70	30	81.08	3	10	8.54	9.00	1.64	Yes
R2Des09	Research strategy x target audience	3	8.11	22	59.46	1	10	7.14	8.00	2.19	No
R2Des08	Time constraints	5	13.51	18	48.65	1	10	7.05	7.00	2.71	No

Note: in this round, only the first two rules for building consensus presented in Topic 3.3.3 were considered given that the third rule applies only to the comparison of the standard deviations between rounds.

Table 23:
Second round – research subjects (descending order of percentage of grades ≥8)

Propos.	Attribute/relationship	Grade ≤3		Grade ≥8		Min	Max	Mean	Median	SD	Consensus
		n	%	n	%						
R2Sub02	Representativeness of subjects	2	5.41	26	70.27	1	10	7.76	8.00	2.09	No

Only the 'overview x strategy' relationship achieved consensus in the two tables set forth above. For those propositions that did not attain agreement, high standard deviations

above 2 remain, which shows that the respondents hold divergent opinions regarding the adequacy of these items for their research practices. Additionally, the 'time constraints' statement achieved a low median and a high standard deviation, which also suggests disagreement among experts related to the applicability of this aspect in their everyday research practice. Research design and research subjects may be associated with the first stage described by Brinberg and McGrath (1985) because they refer to a study's planning stage.

The key features denominated data collection, data analysis, findings, and reporting are associated with stage 2 of the VNS model proposed by Brinberg and McGrath (1985) because they relate to the research development stage. Tables 24 and 25 evaluate attributes/relationships related to data collection and data analysis and are associated with the criteria of integrity, rigor/thoroughness, and reliability/auditability.

Table 24:
Second round – data collection (descending order of percentage of grades ≥8)

Propos.	Attribute/relationship	Grade ≤3		Grade ≥8		Min	Max	Mean	Median	SD	Consensus
		n	%	n	%						
R2Col03	Voluntary participation of subjects	0	0.00	23	100.00	8	10	9.65	10.00	0.65	Yes*
R2Col02	Notes for divergent events	0	0.00	22	78.57	6	10	8.46	8.50	1.23	Yes
R2Col04	Formal agreement from humans	1	3.85	18	69.23	1	10	8.15	8.00	2.05	No

*This was included in the second round because this proposition is linked with ethical aspects.

Table 25:
Second round – analysis (descending order of percentage of grades ≥8)

Propos.	Attribute/relationship	Grade ≤3		Grade ≥8		Min	Max	Mean	Median	SD	Consensus
		n	%	n	%						
R2Ana04	Other views to explain context	0	0.00	19	51.35	4	10	7.51	8.00	1.33	No

To avoid bias in interpreting the data, the 'voluntary participation of subjects' item of Table 24 was included in the second round because of its connection with ethical aspects. Thus, the data collection instrument in the second round provided the 'never did/do not do research involving human subjects' item with the purpose of segregating those respondents who do not conduct that type of research. The results showed that 23 respondents, which equaled 100% of the researchers that conduct research with human beings, assigned equal or superior grades to 8. As a result, this item and the items related to 'Notes for divergent events' reached consensus by a percentage of higher grades.

Tables 26 and 27 highlight the propositions of the last two features related to Stage 2. These statements are linked to the criteria of the contribution/quality of theoretical perspective, impact, integrity, and internal validity/credibility or defensibility.

Table 26:
Second round – findings (descending order of percentage of grades ≥8)

Propos.	Attribute/relationship	Grade ≤3		Grade ≥8		Min	Max	Mean	Median	SD	Consensus
		n	%	n	%						
R2Fin05	Insights for thinking about the field	0	0.00	29	78.38	6	10	8.16	8.00	1.04	Yes
R2Fin09	Impact from the nature of divergences	0	0.00	25	67.57	5	10	7.81	8.00	1.17	No

Table 27:
Second round – reporting (descending order of percentage of grades ≥8)

Propos.	Attribute/relationship	Grade ≤3		Grade ≥8		Min	Max	Mean	Median	SD	Consensus
		n	%	n	%						
R2Rep02	Disclosure of limitations	0	0.00	30	81.08	6	10	8.51	9.00	1.10	Yes
R2Rep01	Discussion of impact on knowledge	0	0.00	27	72.97	6	10	8.14	8.00	1.00	No
R2Rep03	Reasons for limitations	0	0.00	27	72.97	5	10	7.92	8.00	1.09	No
R2Rep04	Boundaries of the study	1	2.70	25	67.57	1	10	7.95	8.00	1.70	No
R2Rep09	Appreciation of drafts by colleagues	8	21.62	14	37.84	1	10	6.19	7.00	2.93	No

The tables discussed above show that only two of the seven attributes/relationships reached consensus in this round. In line with what was observed in the judgment of the proposition regarding time constraints (Table 22), it is noted in Table 27 that 21.62% of respondents attributed grades of equal to or less than 3 to the item 'appreciation of drafts by colleagues.' In addition, this statement achieved a low median and a high standard deviation, which also suggests the experts' disagreement on the applicability of this aspect of their research practices. The submission of drafts for assessment by colleagues is a practice that should be a habit among researchers because it helps identify weaknesses in papers before they are submitted for publication and might thus contribute to increasing acceptance rates.

Tables 28-30 are associated with the third stage of the Brinberg and McGrath (1985) schema and focus on reflexivity and neutrality, ethical issues, and auditability. The propositions that are in these tables are linked to the criteria of integrity, rigor/thoroughness, and reliability/auditability.

Table 28:
Second round – reflexivity and neutrality (descending order of percentage of grades ≥8)

Propos.	Attribute/relationship	Grade ≤3		Grade ≥8		Min	Max	Mean	Median	SD	Consensus
		n	%	n	%						
R2Neu01	How to address errors and biases	2	5.41	21	56.76	1	10	7.19	8.00	2.15	No
R2Neu02	Impact of team participation	3	8.11	16	43.24	1	10	6.73	7.00	2.19	No

Table 29:
Second round – ethical issues (descending order of percentage of grades ≥8)

Propos.	Attribute/relationship	Grade ≤3		Grade ≥8		Min	Max	Mean	Median	SD	Consensus
		n	%	n	%						
R2Eth08	Confidentiality to participants' data	0	0.00	25	96.15	7	10	9.54	10.00	0.81	Yes
R2Eth07	Strategy to mitigate possible harm	2	7.69	21	80.77	1	10	8.15	9.00	2.48	Yes
R2Eth01	Registration of the study/CEP	5	19.23	20	76.92	1	10	7.96	10.00	3.57	Yes
R2Eth06	Formal respect for human subjects	3	12.00	17	68.00	1	10	7.80	9.00	3.01	No
R2Eth05	Use of the principles of a code	5	19.23	17	65.38	1	10	7.19	9.00	3.48	No
R2Eth02	Rules for registering studies/CEP	5	20.00	16	64.00	1	10	7.04	9.00	3.55	No
R2Eth04	Compulsory submission to the CEP	9	36.00	10	40.00	1	10	5.68	7.00	3.88	Yes
R2Eth03	Formal submission to the CEP	7	28.00	9	36.00	1	10	5.76	6.00	3.50	Yes

None of the propositions shown in Table 28 achieved consensus in this second round regarding the first two rules used to define them. Notably, the standard deviations were high: both exceeded 2. With respect to the propositions relating to ethical issues (Table 29), it is important to consider that the total number of considered responses is lower because of their specificity. The findings revealed high medians for issues relating to the confidentiality of participants' data and 'CEP to register the studies'. However, the statements 'compulsory submission to the ethics committee' and 'formal submission to the ethics committee' were the only ones that achieved consensus by more than 25% of the grades equal to or less than 3. Finally, only three of these eight attributes/relationships did not achieve consensus based on the first two rules presented in Topic 3.3.3.

Table 30:
Second round – auditability (descending order of percentage of grades ≥8)

Propos.	Attribute/relationship	Grade ≤3		Grade ≥8		Min	Max	Mean	Median	SD	Consensus
		n	%	n	%						
R2Aud01	Records of the design changes	1	3.45	23	79.31	1	10	8.34	9.00	1.86	Yes
R2Aud02	Records of the reasons for changes	1	3.45	21	72.41	1	10	8.17	8.00	1.91	No
R2Aud04	Guard of documents to reduce risks	2	8.00	16	64.00	1	10	7.76	9.00	2.49	No

To complete the analysis of attributes/relationships evaluated in the second round of Delphi, Table 30 includes the attributes/relationships related to auditability. The findings revealed that the 'record of the design changes' item was the only one that achieved consensus at this stage. Nevertheless, the high standard deviation about the 'guard documents to reduce risks" item is highlighted. Similar to what happened to the propositions regarding data collection and ethics, items related to auditability also included the possibility of the choice of a specific alternative by researchers who do not conduct studies involving human subjects to avoid biased results. Table 31 shows the distribution of responses among researchers who did and did not conduct research involving human subjects; the reason for doing so is to identify

the number of valid cases such that the results revealed in Tables 24, 29, and 30 can be calculated.

Table 31:
Distribution of respondents based on valid cases.

Answer	R2 Col02	R2 Col03	R2 Col04	R2 Eth01	R2 Eth02	R2 Eth03	R2 Eth04	R2 Eth05	R2 Eth06	R2 Eth07	R2 Eth08	R2 Aud01	R2 Aud02	R2 Aud04
Yes	28	23	26	26	25	25	25	26	25	26	26	29	29	25
No	9	14	11	11	12	12	12	11	12	11	11	8	8	12
Total	37	37	37	37	37	37	37	37	37	37	37	37	37	37

Yes = respondent performs research that involves human beings.

The results in Table 31 showed some contradictions because of a change observed in the number of negative responses related to the use of research involving human subjects. First, on some items there was a variation in the number of respondents who declared that they conducted research involving human beings within a range of 8 to 14 respondents. Second, even for those attributes/relationships concerning data collection and auditability that did not refer to features involving studies with humans, some researchers preferred to select the differentiated alternative. As a result, nine respondents used that option in the 'Notes for divergent events' statement, which belongs to the data collection characteristic. Additionally, eight experts used the declarations 'records of design changes' and 'records of the reasons for changes,' both of which were linked to auditability and apply to all types of research.

To determine whether a third round of Delphi would be required, the results of the two previous rounds were summarized and matching was performed. In so doing, only the attributes/relationships that comprised the data collection instrument of the second round were considered. The rules described in Topic 3.3.3 – primarily the comparison of the variations of the standard deviations between the two rounds for those statements that did not achieve consensus – were taken into account, and the results are highlighted in Table 32.

Table 32:
Consensus results after the second round.

Proposition	Key feature	1st round			2nd round			% variation of SD	Variation ≤25%	Final consensus
		Mean	SD	Cons.	Mean	SD	Cons.			
R1Des03	Design	7.88	2.32	No	8.54	1.64	Yes	29.03	No	Formed
R1Des08		6.44	2.92	No	7.05	2.71	No	7.15	Yes	Formed
R1Des09		6.76	2.55	No	7.14	2.19	No	14.11	Yes	Formed
R1Sub02	Subjects	8.15	1.99	No	7.76	2.09	No	-4.64	Yes	Formed
R1Col02	Data coll.	8.44	1.57	No	8.46	1.23	Yes	21.35	Yes	Formed
R1Col03		9.37	1.65	Yes	9.65	0.65	Yes	60.88	No	Formed
R1Col04		6.90	3.28	No	8.15	2.05	No	37.50	No	Unformed
R1Ana04	Analysis	7.07	1.95	No	7.51	1.33	No	32.18	No	Unformed
R1Fin05	Findings	8.27	1.10	No	8.16	1.04	Yes	5.00	Yes	Formed
R1Fin09		7.44	1.83	No	7.81	1.17	No	35.84	No	Unformed
R1Rep01	Reporting	7.90	1.14	No	8.14	1.00	No	11.57	Yes	Formed
R1Rep02		8.29	1.33	No	8.51	1.10	Yes	17.44	Yes	Formed
R1Rep03		7.46	1.53	No	7.92	1.09	No	28.98	No	Unformed
R1Rep04		7.49	1.61	No	7.95	1.70	No	-5.23	Yes	Formed
R1Rep09		6.54	2.67	No	6.19	2.93	No	-9.65	Yes	Formed
R1Neu01	Refl. /Neut.	6.54	2.44	No	7.19	2.15	No	12.09	Yes	Formed
R1Neu02		6.22	2.33	No	6.73	2.19	No	5.81	Yes	Formed
R1Eth01	Ethical Issues	7.49	3.61	No	7.96	3.57	Yes	0.96	Yes	Formed
R1Eth02		6.66	3.81	Yes	7.04	3.55	No	6.79	Yes	Formed
R1Eth03		5.95	3.63	Yes	5.76	3.50	Yes	3.38	Yes	Formed
R1Eth04		5.56	3.91	Yes	5.68	3.88	Yes	0.80	Yes	Formed
R1Eth05		7.02	3.45	No	7.19	3.48	No	-0.87	Yes	Formed
R1Eth06		7.22	3.11	No	7.80	3.01	No	3.11	Yes	Formed
R1Eth07		8.12	2.91	Yes	8.15	2.48	Yes	14.84	Yes	Formed
R1Eth08		9.07	1.63	Yes	9.54	0.81	Yes	50.34	No	Formed
R1Aud01	Auditability	7.68	2.61	No	8.34	1.86	Yes	28.90	No	Formed
R1Aud02		7.29	2.63	No	8.17	1.91	No	27.35	No	Unformed
R1Aud04		6.98	3.27	No	7.76	2.49	No	23.85	Yes	Formed

As shown in the column in which the results of second round's consensus (yes/no) are presented, 11 of the 28 propositions comprising the data collection tool achieved consensus in the second round based on the first two rules specified in Topic 3.3.3. Next, the percentage variation of the standard deviations between the rounds was calculated, and the results are shown in the ninth column. Then, the items with a difference less than or equal to 25% to build consensus were found, as outlined in the rules, and the results are in the "Variation ≤25% column". Finally, items that did not result in consensus based on the parameters have been highlighted. Given these parameters, the results shaded in the last column of Table 32 showed that only five statements did not reach a consensus based on the rules described above. Because that number was less than 10% of all the attributes/relationships assessed, it was decided to eliminate the third Delphi round because the gain would only be marginal and there would be no guarantee that consensus would result for those items.

4.2.3 Final results

Following the results of the two rounds of the Delphi and the determination that convening an additional round would not be relevant for data collection, the final numbers were evaluated. As previously detailed, the evaluation indicated that only five statements had not reached consensus and those statements did not justify the effort of convening another round because all of these not only displayed a high variation in standard deviations but also presented standard deviations greater than 1, even after the second round.

The results the evaluation of the statements in the two rounds were consolidated into one list, and the outcomes obtained were considered in the stage in which consensus was built. These results were organized into a general table with eight columns: first, the key feature with which the proposition is associated; second, the assessed attributes/relationships; third, the number of valid answers; fourth, the round in which the consensus was reached; fifth, the absolute score; sixth, the relative score; seventh, the ranking by the highest relative score; and eighth, the level of agreement.

The absolute score is the sum of the grades received from the attribute/relationship in the round in which consensus was built. The relative score of each item was calculated from the proportion of the absolute score and the maximum score possible, which is the number of valid responses for each statement multiplied by the maximum score of the scale (10). This was considered the relative score, which was the basis for classifying the propositions in the rankings. Additionally, propositions have been classified into four levels based on an adaptation of the scale according to levels of agreement (Rodríguez-Mañas et al., 2013), as detailed in Topic 3.3.3. Finally, the issues that did not achieve consensus were included to provide the full list of the attributes/relationships analyzed. The results are shown in Tables 33-36 based on the level of agreement. Table 33 presents the results that reached the level of strong.

Table 33:
Final classification – attributes/relationships with strong level of agreement.

Key feat.	Attribute/relationship	*n*	Round	Score	%	Ranking	Level
Data coll.	Voluntary participation of subjects	23	2	222	96.52	1º	Strong
Design	Goal/problem shown precisely	41	1	395	96.34	2º	Strong
Ethics	Confidentiality to participants' data	26	2	248	95.38	3º	Strong
Reporting	Conclusions x aim	41	1	377	91.95	4º	Strong
Findings	Comparison results x other studies	41	1	373	90.98	5º	Strong
Findings	Check links of findings x purpose	41	1	373	90.98	5º	Strong
Reporting	Literature review x main concepts	41	1	373	90.98	5º	Strong
Reporting	Theory to support propositions	41	1	371	90.49	8º	Strong
Design	Useful strategy for purpose	41	1	369	90.00	9º	Strong
Findings	Previous findings x hypothesis	41	1	369	90.00	9º	Strong
Auditability	Safeguard databases for checks	41	1	367	89.51	11º	Strong
Design	Arguments for technique choices	41	1	366	89.27	12º	Strong
Findings	Link between findings x evidence	41	1	363	88.54	13º	Strong
Data coll.	Notes for each research step	41	1	361	88.05	14º	Strong
Findings	New areas based on the findings	41	1	361	88.05	14º	Strong
Findings	Path to achieve conclusions	41	1	359	87.56	16º	Strong
Design	Clear overview that guides the study	41	1	358	87.32	17º	Strong
Analysis	Description of tools and procedures	41	1	358	87.32	17º	Strong
Subjects	Criteria to design/select subjects	41	1	352	85.85	19º	Strong
Design	Access to the data	41	1	351	85.61	20º	Strong
Analysis	Description of nature & form of data	41	1	351	85.61	20º	Strong
Analysis	Significance of data to reaching aims	41	1	351	85.61	20º	Strong
Findings	Does the context allow replication	41	1	351	85.61	20º	Strong
Design	Overview x strategy	37	2	316	85.41	24º	Strong
Reporting	Disclosure of limitations	37	2	315	85.14	25º	Strong
Design	Overview x theory	41	1	349	85.12	26º	Strong
Analysis	Implicit/explicit links - find. x aims	41	1	349	85.12	26º	Strong
Data coll.	Notes for divergent events	28	2	237	84.64	28º	Strong
Reporting	Explicit possibility of generalization	41	1	345	84.15	29º	Strong
Auditability	Records of design changes	29	2	242	83.45	30º	Strong
Analysis	Context x impact on data analysis	41	1	341	83.17	31º	Strong
Findings	Insights for thinking about the field	37	2	302	81.62	32º	Strong
Ethics	Strategy to mitigate possible harm	26	2	212	81.54	33º	Strong
Reporting	Discussion of impact on knowledge	37	2	301	81.35	34º	Strong
Ethics	Registration of the study/CEP	26	2	207	79.62	35º	Strong
Reporting	Boundaries of the study	37	2	294	79.46	36º	Strong
Ethics	Formal respect for human subjects	25	2	195	78.00	37º	Strong
Auditability	Guard documents to reduce risks	25	2	194	77.60	38º	Strong
Subjects	Representativeness of subjects	37	2	287	77.57	39º	Strong

The findings revealed that 73.58% of the propositions received the level of strong acceptance. Moreover, it appears that of the ten items that achieved the highest relative scores and those at the top of the ranking, the criteria of internal validity/credibility or defensibility stand out with four statements, and the criteria of integrity and suitability stand out with two statements each. The 'voluntary participation of subjects' attribute, which is linked to the criterion of integrity, achieved the highest relative score. However, it did not achieve the highest absolute score because it achieved consensus in the second round, when the number of answers was lower. The item that earned the highest absolute score was 'goal/problem

shown precisely', which is related to the criterion of suitability and obtained 395 points out of a possible 410.

With respect to the principal characteristics, the findings relating to the ten better-ranked attributes/relationships revealed that 'findings' and 'reporting' had the highest number of items included in the top 10 ranking, each with three statements. With respect to the number of propositions linked to each general criterion, feasibility had the lowest percentage of items with a high level of acceptance, equivalent to 33.33%, which is followed by integrity, with 50% of the items. Conversely, internal validity/credibility or defensibility obtained the highest number of statements with a high level of acceptance, with eight propositions that are equivalent to 80% of the declarations related to this criterion. Table 34 shows the group of propositions that were classified as having a moderate level of agreement.

Table 34:
Final classification – attributes/relationships with moderate level of agreement.

Key feat.	Attribute/relationship	*n*	Round	Score	%	Ranking	Level
Ethics	Use of the principles of a code	26	2	187	71.92	1º	Moderate
Refl./Neutr.	How to address errors and biases	37	2	266	71.89	2º	Moderate
Design	Research strategy x target audience	37	2	264	71.35	3º	Moderate
Design	Time constraints	37	2	261	70.54	4º	Moderate
Ethics	Rules for registering studies/CEP	25	2	176	70.40	5º	Moderate
Refl./Neutr.	Impact of team participation	37	2	249	67.30	6º	Moderate

The findings revealed that six of the items were classified as having a moderate level of agreement and consensus on these statements was achieved in the second round. Moreover, it is noted that three of these items are associated with the integrity criterion and two are associated with feasibility and rigor. These findings suggest that some of these elements have no significant weight in the respondents' research practices, which might be cause for concern because crucial aspects of conducting those studies are included among the propositions with a moderate level of agreement. Examples of this inclusion include time constraints for carrying out the research, some aspects related to ethical issues, and disclosure of how the researcher addressed errors and biases arising during the study. With respect to the key features, the propositions with a moderate level of acceptance refer to ethics, reflexivity/neutrality, and research design. These findings, with lower acceptable levels, may suggest the existence of a possible misalignment between the methodological, substantive, and theoretical domains of the VNS (Brinberg & McGrath, 1985; Brinberg, 1982; McGrath & Brinberg, 1983). Table 35 shows the attributes/relationships that obtained a low level of agreement.

Table 35:
Final classification – attributes/relationships with low level of agreement.

Key feat.	Attribute/relationship	*n*	Round	Score	%	Ranking	Level
Reporting	Appreciation of drafts by colleagues	37	2	229	61.89	1º	Low
Ethics	Formal submission to the CEP	25	2	144	57.60	2º	Low
Ethics	Compulsory submission to the CEP	25	2	142	56.80	3º	Low

Table 35 shows that the three items that have a low level of agreement are related to rigor/thoroughness and internal validity/credibility or defensibility. Moreover, the statement regarding appreciation of preliminary versions (drafts) by peers is associated with the reporting stage, and the other two statements are related to the ethical aspects of formal and compulsory submission of research projects to ethical committees.

The appraisal of preliminary versions of research reports by colleagues plays a significant role in identifying potential problems and suggesting points for improvement. Such appraisal also helps increase a paper's consistency and logical coherence. As a result, it is believed that such an appraisal may enhance the possibility of a manuscript being approved for conferences and publications because it reduces the effects of possible failures that might be identified in the peer review and that might lead to a rejected submission.

Findings relating to the formal and mandatory submission of research projects involving human beings to the Ethics Committee for Research (CEP) may suggest a certain weakness associated with rigor/thoroughness and integrity. As is the case internationally, Brazilian law requires that all studies involving human subjects be submitted for assessment and prior approval by an ethics committee in order to ensure that research complies with general ethical principles such as respect for each participant's dignity and autonomy. Thus, Brazilian legislation requires that the informed consent describes the study's purpose, goals, and procedures, clarify any possible discomfort and risks of participation, and specify the form of monitoring and assistance to which all participants have a right, even after completion of the study (Brasil, 2012, 2013). Additionally, studies must not only ensure that participants are granted full freedom but also assure their confidentiality and privacy during all stages of research. In this manner, the law guarantees not only human subjects' rights but also protection for the researcher and any institutions involved against possible risks. These practices are consistent with international principles of respect for persons, beneficence, and justice (National Institutes of Health, 1979), along with accountability, respect, and integrity (European Science Foundation, 2011; OADS, 2012). One of the respondents, for example, highlighted in the open question that his or her institution has a Research Ethics Committee for the submission and approval of research designs. However, only projects linked to

graduate programs (dissertations and theses) are currently required to submit their research designs to the committee. Other studies involving human beings that do not fit within this rule are typically not submitted for committee approval.

There are at least three practical effects of compliance with ethical standards covering research involving humans:

(a) ensuring that the research complies with internationally accepted standards of integrity;

(b) mitigating risks for the researcher and the institution related to possible court challenges by participants who have felt harmed to some extent as a result of study participation; and

(c) increasing external validity and relevance, which may culminate in greater acceptance by the scientific community and a greater chance of obtaining funding.

Notably, the Operational Norm n. 001/2013 (Brasil, 2013) envisages that in cases of complaints and ethical infraction situations – particularly those that imply risks to research participants – there should be an examination that might involve sanctions, including a demand for an investigation by the Brazilian prosecution system. Table 36 shows the results of those propositions that have not achieved consensus.

Table 36:
Final classification – attributes/relationships with unformed consensus.

Key feat.	Attribute/relationship	*n*	Round	Score	%	Ranking	Level
Auditability	Records of the reasons for changes	29	2	237	81.72	1º	Unformed
Data coll.	Formal agreement from humans	26	2	212	81.54	2º	Unformed
Reporting	Reasons for limitations	37	2	293	79.19	3º	Unformed
Findings	Impact from the nature of divergences	37	2	289	78.11	4º	Unformed
Analysis	Other views to explain context	37	2	278	75.14	5º	Unformed

The consensus definition unformed appearing in Table 36 was based on the failure to comply with the three rules set forth in Topic 3.3.3. The five attributes and relationships that did not achieve consensus are associated with the criteria of integrity and reliability/auditability (with two propositions each) and the criteria of internal validity/credibility or defensibility (with one proposition). Although the final scores are equivalent to a strong level of agreement, none of the items reached 75% for scores equal to or greater than 8, which was one of the criteria required to establish consensus and which was shown in Table 32. Each of the stages of research process referred to is associated with a distinct feature.

The two attributes/relationships related to integrity were 'formal agreement of human beings' concerning the declaration of agreement to participate in the study and 'reasons for limitations' on research. The first attribute must be expressly obtained by means of the participant's signature on the informed consent and is a means of protecting both the researcher and the participant regarding compliance with legal requirements for research involving human subjects. In turn, the item about limitations highlights the need to provide the target audience with a frank explanation of the reasons that contributed to generating those constraints. Furthermore, the exposure of study limitations contributes to increasing a study's external validity because the scope of findings becomes more obvious and researchers will be able to choose ways to mitigate limitations in future studies. The attributes/relationships relating to reliability/auditability are 'records of the reasons for changes' and 'other views to know the context'. The first is related to the proper and accurate representation the methodology followed in developing the study, comparing what was originally planned with what was actually done and recording the reasons for any changes. Therefore, this is a measure that enables checking procedures. The other views on understanding the context contribute to offering alternative solutions to the research problem; adequate characterization of the context is relevant because it allows a comparison of the findings with those of other studies. The final attribute/relationship is 'impact on the nature of divergences' and involves internal validity/credibility or defensibility. The discussion of the nature and source of possible divergences aims to highlight their potential effects on the study's findings. In addition to bringing more logical connection to the study, discussion of divergences augments consistency by showing how to reach the study's conclusions and is also associated with the possibility of identifying alternative approaches to solving the proposed problem in light of the data. Finally, Figure 19 shows the distribution of statements based on levels of agreement.

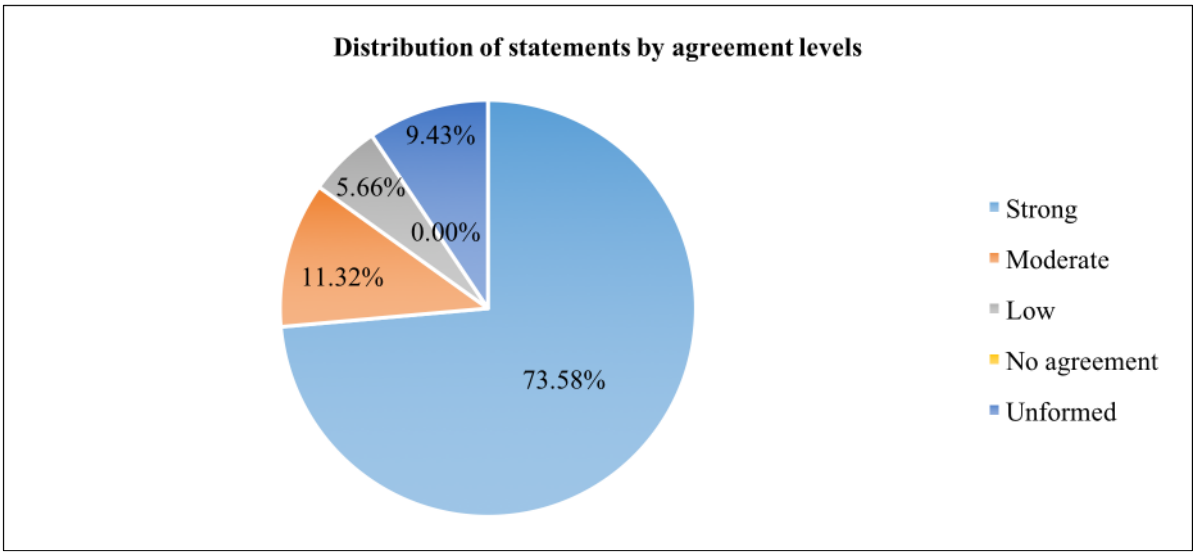


Figure 19: Distribution of statements by agreement levels.

As observed in Figure 19, approximately 3/4 of the statements showed a strong level of agreement. Moreover, considering the model adapted from Rodríguez-Mañas et al. (2013) for a definition of the agreement levels, there was no proposition with a relative score of less than 50%, which would be characterized as no agreement.

4.3 An Approach to Judging the Quality of the Research Process in Accounting

After the presentation of the Delphi results, and after taking into account the importance of the link between research quality criteria and respondents' perceptions of the statements' consistency with their practices, it was decided to complete the analysis by presenting an approach to assess the quality of accounting research. Thus, the present study suggests a set of elements to evaluate the research process in the field from the stages of the research process (key features), attributes/relationships, general criteria for research quality, and domains of scientific investigation.

A visual representation of the distribution of criteria at each level was generated to more clearly highlight representativeness and to assess the respondents' levels of agreement as to the suitability/adherence of statements to their research practices. These results are displayed in Figure 20.

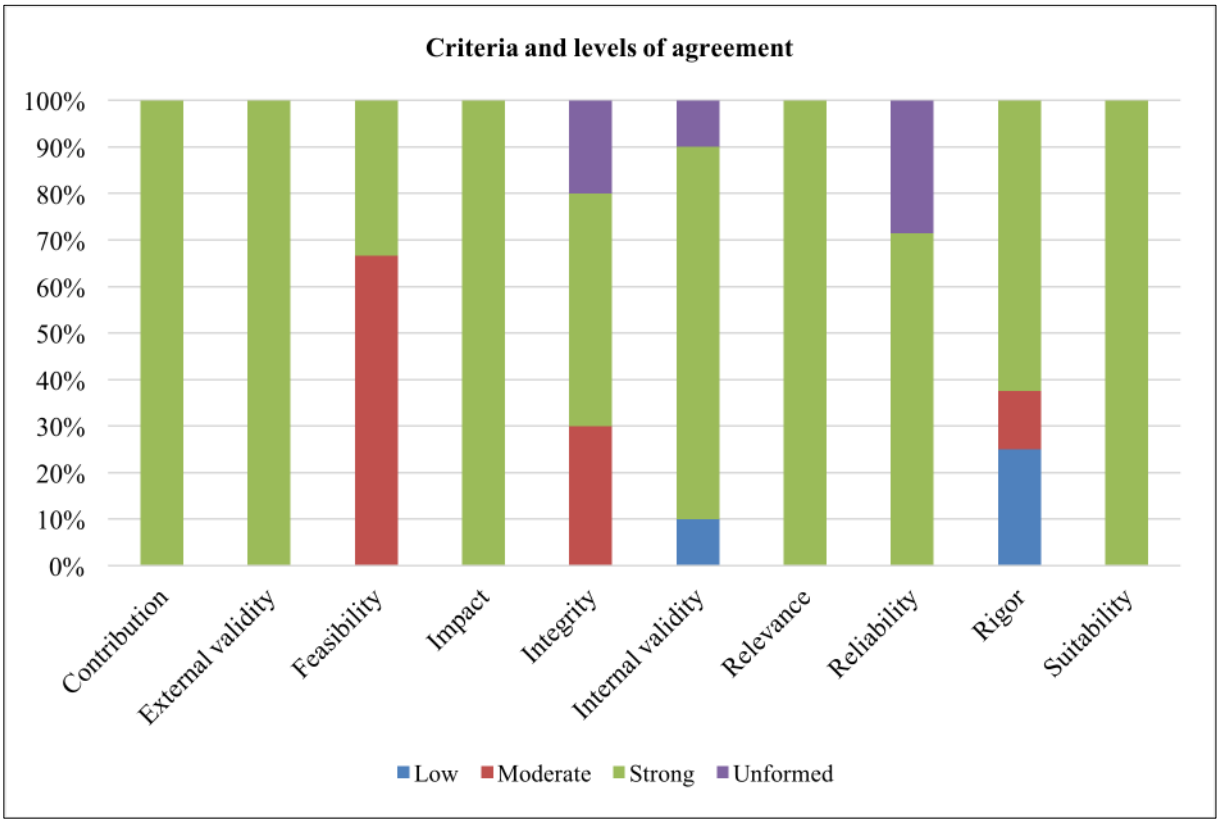


Figure 20: Criteria and levels of agreement.

The graphic representation reveals that statements regarding the contribution/quality of the theoretical perspective, external validity or generalization/fittingness, relevance, and suitability show a strong adherence in respondents' opinion. Approximately 65% of the items related to the feasibility criterion are at a moderate level. Although this level does not necessarily indicate the occurrence of serious failures, these results should be noted as studies are developed because they are associated with elements such as research design dimensioning when taking time constraints into account and considering the target audience when selecting one's research strategy. Another criterion that had moderate-level propositions was integrity, at 35%, which requires attention to issues such as how the researcher addressed errors and biases and the impact of the research team's participation. Low-level assertions concentrated on the criteria of internal validity/credibility or defensibility and rigor. Notably, approximately 1/4 of the items concerning rigor obtained a low level of agreement, indicating a point of concern to be considered by researchers in suggesting that there are weaknesses in the conduct of their studies.

The items that did not achieve consensus are associated with the integrity, internal validity/credibility or defensibility, and reliability/auditability criteria. The findings suggest that these are issues on which there are discrepancies or lower average acceptance among the

respondents. This situation is clear when looking at the individual data, which reveal that between 26.92% and 43.24% of the respondents attributed grades 5-7 to such statements. Moreover, some of these statements had the highest possible grade range (10, 6, 5, 5, and 10, respectively), which indicates a dispersion of the data that led to the non-formation of consensus, even after having obtained high relative scores, as shown in Table 36.

Finally, as displayed in Table 31, it is notable that the number of valid responses to each statement has varied significantly, which explains why the same number of answers indicating a particular grade does not represent the same relative share compared with other propositions analyzed. In addition, it should be emphasized that there is an imbalance in the number of statements associated with each criterion, as demonstrated in Figure 12. As a result, the analysis of the relative share of one attribute/relationship in the level-based distribution of answers representing the criteria in figures on this topic is proportionally different. For example, whereas only one item is associated with impact, ten items are related to the criteria of integrity and internal validity/credibility or defensibility. Figure 21 shows the distribution by key features and levels of agreement.

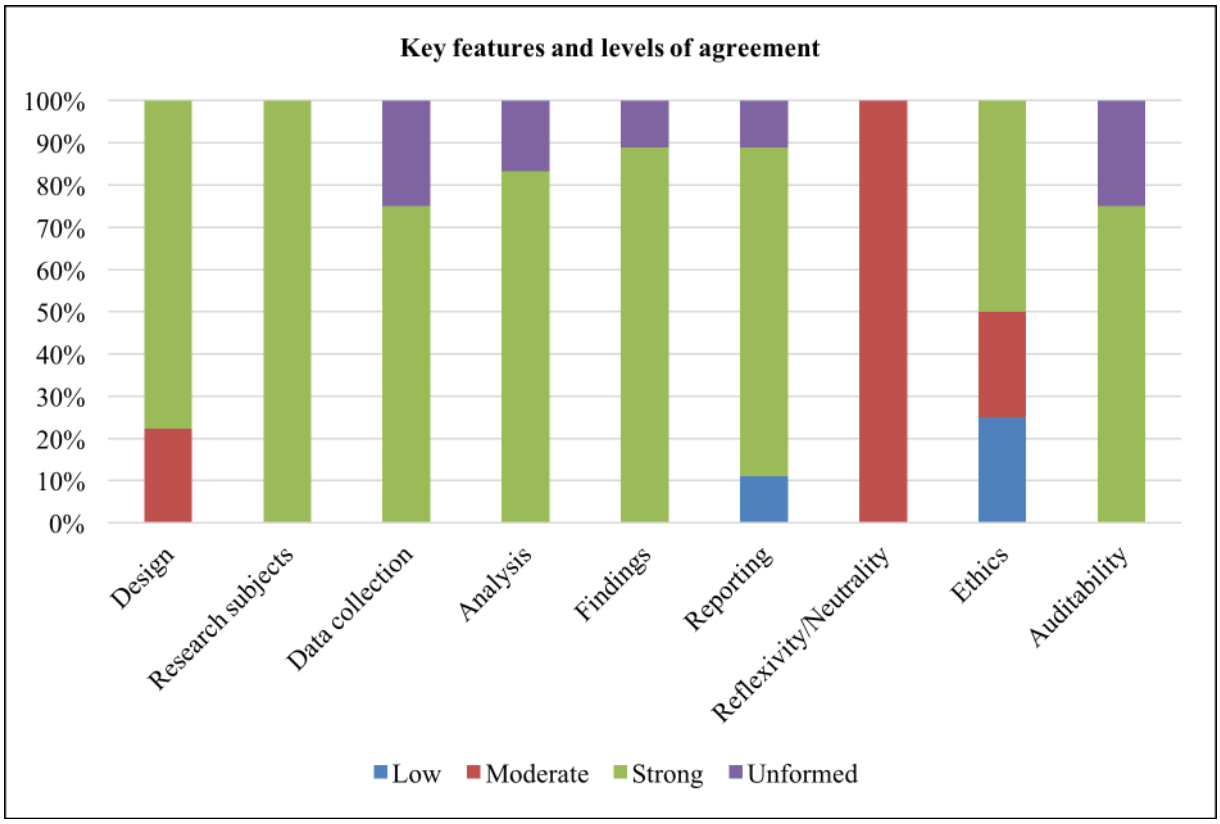


Figure 21: Key features and levels of agreement.

By focusing on the essential features and levels of agreement, Figure 21 showed that 'research subjects' was the only characteristic with a strong level of acceptance. Conversely, reflexivity/neutrality showed the moderate level of all propositions. In addition, 'ethics' and 'reporting' were the features in which low-level statements were concentrated. Approximately 25% of the ethics-related propositions presented a low level of acceptance. This finding suggests that researchers must care more about such procedures, which in addition to the associated legal aspects can indicate not only a failure to fulfill the essential characteristics of research integrity but also weaknesses in methodological rigor in conducting the study. Ethical issues were the only feature that had statements at the three levels of agreement: low, moderate, and strong. Additionally, data collection, analysis, findings, reporting, and auditability have concentrated propositions whose consensus was not formed. Finally, for the data collection and auditability stages, this result represented 25% of the items.

Intending to structure the proposal for an approach to assess quality in the research process in accounting, the structure of the logic model was employed. This model is widely applied to evaluate programs and projects for both processes and outcomes (or variance). A logic model for process evaluation is structured in inputs, processes/activities, and process indicators/outputs. The approach based on the logic model considers as inputs the stages of the research process (Figure 22). Initially, this approach admitted the three phases proposed by Brinberg and McGrath (1985), which this work calls the first (planning), second (execution) and third (results monitoring – outcomes) phases. Additionally, the nine key features (Mays & Pope, 2006; Spencer et al., 2003) were grouped into these three stages in the following distribution: (1) design and research subjects; (2) data collection, analysis, findings, and reporting; and (3) reflexivity/neutrality, ethics, and auditability.

The attributes and relationships used in the Delphi data collection instruments were considered to characterize the processes and activities proposed in the logic model. This association was made on the grounds that those attributes and relationships adequately represented the processes, considering that the building of the device itself was based on the respondents' research practices, i.e., in activities performed at different stages of their studies. The statements were grouped according to the key feature and the quality criterion that are associated with one another.

With respect to the process indicators/outputs, the criteria used to group the statements were allocated, i.e., they are associated with the attributes/relationships that are defined as processes/activities. Additionally, some potential implications of either violating or ignoring

these criteria were submitted and therefore, used the domain structure of the validity network schema (VNS).

In conclusion, it must be emphasized that this approach intends to contribute to the discussion of quality criteria for building and performing research (the process). It is less focused on final results (product evaluation). This approach merely attempts to help accounting researchers self-evaluate their studies' quality. It does not intend to replace other frameworks for research evaluation.

LOGIC MODEL TO EVALUATE THE RESEARCH PROCESS

INPUT	PROCESSES/ACTIVITIES		PROCESS INDICATORS/OUTPUTS	
Stage	**Attribute/relationship**		**Criterion**	**Some possible implications**
1st – Planning				
Design	• Useful strategy for purpose • Clear overview • Overview x strategy	• Goal/problem definition • Overview x theory • Arguments to technique choices	Suitability	SUBSTANTIVE DOMAIN • Unclear gap to justify the research • Low contribution to advancing knowledge • Use of inadequate research strategy • Waste of resources by inappropriate use of time • Failure to show research impacts
	• Access to the data • Time constraints	• Strategy x target audience	Feasibility	
Research subjects	• Criteria to design/select subjects	• Representativeness of subjects	Internal validity	
2nd – Execution				
Data collection	• Notes for each research step	• Notes for divergent events	Rigor	
	• Voluntary participation of subjects	• Formal agreement of subjects	Integrity	
Analysis	• Description of nature and form of data		Integrity	CONCEPTUAL DOMAIN • Lack of clarity of main concepts related to the study • Absence of a theory to support data analysis • Difficulty justifying hypotheses based on literature
	• Description of tools and procedures	• Implicit/explicit links – findings x aims	Rigor	
	• Context x impact on data analysis	• Other views to know context	Reliability	
	• Significance of data to reach the aim		Relevance	
Findings	• Path to achieve conclusions • Link findings x evidence	• Check links findings x purpose • Impact from the nature of divergences	Internal validity	
	• Compare results with other studies		External validity	
	• New areas based on findings	• Insights for thinking about the field	Contribution	
	• Previous findings x hypothesis		Relevance	
	• Context to allow replication		Reliability	
Reporting	• Discussion of impact on knowledge		Impact	METHODOLOGICAL DOMAIN • Failures in ethical procedures • Difficulty in allowing replication • Limitations to generate comparison • Research context inadequately described • Insufficient records of steps taken to conduct the research • Conclusions not supported by findings • Risk of derailing the publication of results
	• Disclosure of limitations	• Boundaries of the study	Contribution	
	• Reasons for limitations		Integrity	
	• Literature review x main concepts • Theory to support propositions	• Conclusions x aim • Appreciation of drafts by colleagues	Internal validity	
	• Explicit possibility of generalization		External validity	
3rd – Outcomes				
Reflexivity/Neutrality	• How to address errors and biases	• Impact of team participation	Integrity	
Ethics	• Ethics committee to register the studies • Rules for registering with the CEP	• Formal submission to the CEP • Compulsory submission to CEP	Rigor	
	• Use of ethics code • Respect for human subjects formally	• Strategy to mitigate harm • Confidentiality of participants' data	Integrity	
Auditability	• Records of the design changes • Records of reasons for changes	• Safeguard databases for checks • Guard docs to reduce risks	Reliability	

Figure 22: Logic model to evaluate the research process.

CHAPTER 5 –
CONCLUSIONS

This chapter presents the conclusions of the study and is divided into three topics. The first topic presents final remarks with a summary of the findings and a description of how the goals that were initially proposed were met (or not met). The second topic presents the impacts and recommendations of this study. The third topic makes suggestions for future studies.

5.1 Final Remarks

My personal motivation for discussing research quality arose just over five years ago and spurred my interest in developing this study during my graduate program. This motivation has encouraged me to participate in research forums, conferences, and other discussions and to deepen my reading of the aspects related to what may or may not direct the judgment of research quality in accounting in Brazil. Every research activity requires choices; here, the strategy selected was study perceptions and procedures using the Modified Delphi Technique. That technique was chosen because it has been shown to be feasible and allows experts to build a consensus regarding research practices. This technique has been widely used in other areas of knowledge, particularly in studies designed to assess quality attributes and to create clinical guidelines and protocols.

It must be remembered that there is always a risk associated with proposing a discussion of research quality that focuses on more general criteria. As evidenced in the literature review, different paradigms make different assumptions in analyzing a phenomenon, and the choice of particular criteria can always be a point of criticism, depending on who is making the judgment. Additionally, the literature shows that each stakeholder group will have its own yardstick for judging quality; moreover, each group judges quality beginning from different interests, a fact that in itself demonstrates the difficulty of selecting a set of general criteria to guide judgment regarding the process of producing scientific knowledge.

To identify the quality attributes of good research based on the literature, which was undertaken in Stage 1 as described in the research design, a search was performed that led to the identification of the VNS (Brinberg & McGrath, 1985), and the framework for assessing research evidence (Mays & Pope, 2006; Spencer et al., 2003), along with other sources that

help elucidate the evaluation of the research process. The collected sources supported the choice of a set of 10 quality criteria and 53 attributes/relationships related to the research process to guide the construction of the data collection instrument for the first round of the modified Delphi. Additionally, criteria for evaluating the scientific production of graduate programs' faculty that were established by regulatory quality were presented. These standards are critical because they are considered not only in the definition of the triennial grade assigned to each graduate program but also by development institutions as a measure of the impact and relevance of a researcher's production during evaluation processes for research grants and funding.

Questions about respondents' characterization revealed that most of the experts surveyed have a doctoral degree in accounting that was obtained from another national institution, most did not participate in an exchange program during their doctorate work and did not undertake post-doctoral work, and more than 70% have less than seven years of experience in a graduate program. The data revealed that most of the respondents are members of an editorial board, all the respondents act as referees for scientific periodicals, most of the respondents have or have previously obtained research funding, approximately one-third of the respondents have productivity grants, and almost half of the respondents serve or have served as a journal editor. These findings demonstrate that the expert panel consisted of a qualified audience to evaluate the research process.

The first goal is to observe the stages of the research process in which the attributes of good research are revealed in the practices of Brazilian accounting researchers. To meet this goal, the second stage of the study was conducted by data collecting using the Modified Delphi Technique for consensus building. The first round of Delphi explored 53 attributes/relationships identified in the literature, and the findings revealed that 25 of those attributes/relationships achieved consensus during this step. The data revealed that the attributes are not distributed evenly among the nine stages of the research process used to organize the data collection. They also showed that the extent to which certain attributes/relationships are present in the research practices of Brazilian accounting researchers varies significantly, which is evidenced by the high range of answers, considering the scale used. Finally, it was observed that there is a high level of acceptance of certain attributes/relationships, shown by the number of proposals that have reached the consensus in the Delphi first round.

To achieve the second aim, which was to compare the literature attributes with those identified in accounting research practices, the consolidated results of two rounds of Delphi

were examined. The purpose of this examination was to determine the extent to which the attributes/relationships adhered to research practices in Brazilian accounting; in other words, we sought to identify the level of acceptance of declarations related to those attributes/relationships through the respondents' perceptions. Approximately three-quarters of the statements indicated a strong level of agreement, and 10 of the statements achieved a percentage higher than 90%. Conversely, attributes that achieved only low or moderate levels of acceptance presented significant elements that might compromise the quality and integrity of accounting research. The attributes related to formal and non-formal ethical principles, the need to demonstrate how researchers addressed errors and biases, and the disclosure of the impact of the research team's participation on the results all fit this description. Additionally, some attributes/relationships did not achieve consensus by obtaining a greater number of values between 5-7; those attributes/relationships also deserve attention for mitigating the possible effects that reduce the quality of scientific production in the accounting field.

The study's final goal was to present an approach based on the attributes of good research to judge the quality of scientific production in accounting. The first matter analyzed pursuant to this approach was the relationship between criteria and levels of agreement. A visual representation of that matter has shown that rigor/thoroughness and internal validity/credibility or defensibility have concentrated the low-level statements. In other words, there were evident discrepancies between the attributes/relationships derived from the literature and research practices' adherence to those attributes/relationships in the field, based on the respondents' perceptions. Additionally, more than half of the items associated with feasibility were classified at the moderate level, and most of the elements that did not achieve consensus concern reliability, followed by integrity and internal validity/credibility or defensibility. The second aspect analyzed in the approach was the relationship between criteria and key features. The visual representation stressed that the items with a low level of agreement are linked to ethics and reporting. Furthermore, all the propositions concerning reflexivity/neutrality reached the moderate level, which also occurred with some items related to ethics and research design. Finally, the statements that did not achieve consensus are distributed by five of the nine key features.

The findings that emerged from the Delphi Technique application – in addition to two aspects initially discussed in the approach – validate Hypothesis 1, which established that some current practices used in building the scientific-production process in accounting in Brazil do not adequately meet the quality attributes of good research described in the literature. This position is supported both by the high ranges characterizing the answers to

several items, and by the low acceptance of elements related to ethics, for which Brazilian law does not provide flexibility. In turn, the occurrence of a low level of agreement on issues related to rigor and ethics, along with statements regarding integrity and feasibility that reached the moderate level, are consistent with the acceptance of Hypothesis 2. This second hypothesis stated that the Brazilian institutional environment contributes to reduce the quality of scientific production in accounting as the result of institutional failures that may have substantial effects on research integrity. One striking example of this phenomenon is the compulsory submission of research projects involving human subjects to ethics committees, which is required by Brazilian law but which has achieved only a low level of acceptance. Other ethical issues involving formal and non-formal elements also experienced moderate or low levels of acceptance.

To complete the approach regarding attributes of good research to judge the quality of production, a logic model was designed to help evaluate the research process. The model considered the stages of the research process as inputs, and the processes and activities were linked to attributes/relationships. Certain criteria and possible implications in every domain of scientific investigation were presented as part of the process indicators/outputs. It should be stressed that this approach is not intended to be a comprehensive guideline; instead, it only seeks to identify elements to enable researchers to self-evaluate their studies for compliance with the criteria for the planning, execution, and evaluation of research results.

Given these considerations, this study concludes that there is a clear answer to the research problem, based on the findings and the proposed approach, thus confirming the stated thesis that some Brazilian accounting researchers' practices, in combination with local institutional environments, contribute to reduce the quality of accounting research.

5.2 Impacts and Recommendations

Although the approach encompasses the three domains of investigation proposed in the VNS, the methodological dimension is considered to be better represented in statements because of the nature of such statements. The methodological dimension enables researchers to perceive/address the fact that choices made at each step can result in careful and useful research while helping researchers avoid mistakes discussed in the literature that compromise the quality and integrity of studies in the field.

With respect to this study's expected impacts, it is believed that the issues that it presents, along with the logic model used to evaluate the research process, can contribute to

encouraging the target audience to reconsider their research strategies and re-orient certain practices that have developed that might not be consistent with good practices. Furthermore, it is expected that upon learning about the findings, graduate programs will seek to improve aspects of new researchers' training so that such weaknesses in knowledge production related to accounting can be avoided.

With respect to this study's recommendations, it is important to stress the need to intensify training related to ethical issues in graduate accounting programs to mitigate failures in the conduct of studies that might compromise those studies' integrity and even make it impossible to publish their results. Furthermore, graduate accounting programs are encouraged to specify that the submission of research projects involving human subjects is an internal requirement to sit for the qualification examination. Finally, it is recommended to require approval by the Ethics Committee on Research with Human Beings as a necessary condition for delivering a thesis or dissertation for final defense.

5.3 Suggestions for Future Studies

Studies involving analysis of research processes logically pique interest in evaluating not only each stage of the process but also the role of stakeholders in the production of knowledge. Therefore, future studies should be undertaken to understand graduate students' perceptions of the attributes/relationships analyzed to identify possible mistakes and failures related to training new accounting researchers. Additionally, another relevant inquiry might be to discuss the roles of editors and referees as responsible parties in validating quality, in contrast to their perception of research quality.

Other general topics also affect the quality of the scientific production process and can be the subject of future studies. For example, subjects of future studies might include the use of traditional and less innovative research designs, the relation between choices and methods to establish the research's actual contribution, consciousness of replication to avoid rehashing theory, and features regarding the endogeneity of referees who comprise the body of evaluators in the field's journals. Finally, institutional issues of research practice involving scientific production, such as ethical issues, productivism, and salami science to achieve scores established by the graduate regulator should also be considered as subjects of future research.

REFERENCES

Abbagnano, N. (2007). *Dicionário de Filosofia*. São Paulo: Martins Fontes.

Adler, R., & Liyanarachchi, G. (2011). An empirical examination of the editorial review processes of accounting journals. *Accounting and Finance*. http://doi.org/10.1111/j.1467-629X.2010.00378.x

AEA, A. E. A. American Evaluation Association Guiding Principles For Evaluators (2004). Retrieved from http://www.eval.org/p/cm/ld/fid=51

Amara, N., & Landry, R. (2012). Counting citations in the field of business and management: Why use Google Scholar rather than the Web of Science. *Scientometrics*, *93*, 553–581. http://doi.org/10.1007/s11192-012-0729-2

Amara, N., Landry, R., & Halilem, N. (2015). What can university administrators do to increase the publication and citation scores of their faculty members? *Scientometrics*, *103*(2), 489–530. http://doi.org/10.1007/s11192-015-1537-2

Amin, M., & Mabe, M. (2007). Impact factors: use and abuse. *Perspectives in Publishing*, *1*(October), 1–6. Retrieved from http://cdn.elsevier.com/assets/pdf_file/0014/111425/Perspectives1.pdf

Andrade, J. X. (2011). *Má conduta na pesquisa em ciências contábeis (Doctoral dissertation)*. Universidade de São Paulo. Retrieved from http://www.teses.usp.br/teses/disponiveis/12/12136/tde-10062011-172859/

André, M. (2001). Pesquisa em educação: buscando rigor e qualidade. *Cadernos de Pesquisa*, (113), 51–64. http://doi.org/10.1590/S0100-15742001000200003

Antunes, M. T. P., Mendonça Neto, O. R. de, Oyadomari, J. C. T., & Okimura, R. T. (2011). Conduta ética dos pesquisadores em contabilidade: diferenças entre a crença e a práxis. *Revista Contabilidade & Finanças*, *22*(57), 319–337. http://doi.org/10.1590/S1519-70772011000300006

Aragão, I. R. B. N., Oliveira, J. R. S., & Lima, G. A. S. F. de. (2014). Resonance of articles and impact factor of Brazilian Accounting Journals. *Revista de Educação E Pesquisa Em Contabilidade (Journal of Education and Research in Accounting)*, *8*(1), 5–21.

Babbie, E. R. (2010). *The practice of social research* (12th ed.). Belmont: Wadsworth/CENGAGE.

Ballas, A., & Theoharakis, V. (2003). Exploring diversity in accounting through Faculty journal perceptions. *Contemporary Accounting Research*, *20*(4), 619–44. http://doi.org/10.1506/MLWH-KBTM-ET47-LYKH

Baxter, J., & Chua, W. W. F. (2003). Alternative management accounting research—whence and whither. *Accounting, Organizations and Society*, *28*(2-3), 97–126. http://doi.org/10.1016/S0361-3682(02)00022-3

Beck, C. T. (1993). Qualitative Research: The Evaluation of Its Credibility, Fittingness, and

Auditability. *Western Journal of Nursing Research*, *15*(2), 263–266. http://doi.org/10.1177/019394599301500212

Brasil. Resolução 466, de 12 de dezembro de 2012 (2012). Conselho Nacional de Saúde. Retrieved from http://conselho.saude.gov.br/resolucoes/2012/Reso466.pdf

Brasil. Norma Operacional CNS Nº 001 / 2013 (2013). Conselho Nacional de Saúde. Retrieved from http://conselho.saude.gov.br/web_comissoes/conep/aquivos/CNS Norma Operacional 001 - conep finalizada 30-09.pdf

Brinberg, D. (1982). Validity Concepts in Research: an Integrative Approach. *Advances in Consumer Research*, *9*(1), 40–44.

Brinberg, D., & McGrath, J. E. (1985). *Validity and the research process*. Newbury Park: Sage Publications.

Brinn, T., Jones, M., & Pendlebury, M. (2001). Why Do Uk Accounting and Finance Academics Not Publish in Top Us Journals? *The British Accounting Review*, *33*(2), 223–232. http://doi.org/10.1006/bare.2001.0160

Brown, L. D., & Gardner, J. C. (1985). Using Citation Analysis to Assess the Impact of Journals and Articles on Contemporary Accounting Research (CAR). *Journal of Accounting Research*, *23*(1), 84–109. http://doi.org/10.2307/2490908

Brown, L. D., & Huefner, R. J. (1994). The Familiarity with and Perceived Quality of Accounting Journals: Views of Senior Accounting Faculty in Leading U.S. MBA Programs. *Contemporary Accounting Research*, *11*(1), 223–250. http://doi.org/10.1111/j.1911-3846.1994.tb00442.x

Campbell, S. M., Braspenning, J., Hutchinson, A., & Marshall, M. (2002). Research methods used in developing and applying quality indicators in primary care. *Quality and Safety in Health Care*, *11*(4), 358–364. http://doi.org/10.1136/qhc.11.4.358

CAPES, (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior). (2012). *Comunicado No. 002/2012 - Área de Administração, Ciências Contábeis e Turismo Atualização do Webqualis da Área*. Brasília. Retrieved from https://www.capes.gov.br/images/stories/download/avaliacao/Qualis_Administracao.pdf

CAPES, (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior). (2013). *Documento de Área: Administração, Ciências Contábeis e Turismo, Avaliação Trienal 2013*. Brasília, Brasil. Retrieved from https://www.capes.gov.br/images/stories/download/avaliacaotrienal/Docs_de_area/Administração_doc_area_e_comissão_16out.pdf

CAPES, (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior). (2015a). Ofício Circular nº 022/2015 - DAV/CAPES. *Ofício Circular. Assunto: Divulgacão Da Atualização Do Qualis Periódicos 2013 E 2014.*

CAPES, (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior). (2015b). *Relatório do processo de classificação de periódicos. Área de Administração, Ciências Contábeis e Turismo. Quadriênio 2013-2016 (Maio de 2015)*. Retrieved from https://www.capes.gov.br/images/documentos/RelatorioQualisAdministracao-

2015final.pdf

Carmona, S. (2006). Performance Reviews, the Impact of Accounting Research, and the Role of Publication Forms. *Advances in Accounting*, *22*(06), 241–267. http://doi.org/10.1016/S0882-6110(06)22011-X

Chan, K. C., Chang, C. H., Tong, J. Y., & Zhang, F. (2012). An analysis of the accounting and finance research productivity in Australia and New Zealand in 1991-2010. *Accounting and Finance*, *52*(July 2011), 249–265. http://doi.org/10.1111/j.1467-629X.2011.00440.x

Clarkson, P. M. (2012). Publishing: art or science? Reflections from an editorial perspective. *Accounting & Finance*, *52*(2), 359–376. http://doi.org/10.1111/j.1467-629X.2012.00472.x

Cooper, D. S., & Schindler, P. S. (2003). *Métodos de pesquisa em administração* (7th ed.). Porto Alegre: Bookman.

Creswell, J. W., Klassen, L. a. C., Clark, V. L. P., & Smith, L. K. C. (2011). Best practices for mixed methods research in the health sciences. *Office of Behavioral and Social Sciences Research National Institutes of Health. Retrieved September*, *29*, 2011. Retrieved from https://tigger.uic.edu/jaddams/college/business_office/Research/Best_Practices_for_Mixed_Methods_Research.pdf

Crotty, M. (1998). Introduction: the research process. In *The foundations of social research: Meaning and perspective in the research process* (p. 256). Sage Publications.

Cunha, J. V. A. da. (2007). *Doutores em Ciências Contábeis da FEA/USP: análise sob a óptica da teoria do capital humano (Doctoral dissertation)*. Universidade de São Paulo. Retrieved from http://www.teses.usp.br/teses/disponiveis/12/12136/tde-17102007-173046/pt-br.php

Davidson, E. J. (2005). *Evaluation methodology basics: the nuts and bolts of sound evaluation*. (Sage Publications, Ed.) (1st ed.). Thousand Oaks: Sage Publications.

Davies, D., & Dodd, J. (2002). Qualitative Research and the Question of Rigor. *Qualitative Health Research*, *12*(2), 279–289. http://doi.org/10.1177/104973230201200211

Deckers, K., van Boxtel, M. P., Schiepers, O. J., de Vugt, M., Muñoz Sánchez, J. L., Anstey, K. J., … Köhler, S. (2015). Target risk factors for dementia prevention: a systematic review and Delphi consensus study on the evidence from observational studies. *Int J Geriatr Psychiatry*, *30*, 234–246. http://doi.org/10.1002/gps.4245

Denscombe, M. (2010). *The Good Research Guide*. Open University Press.

Doyle, J., & Arthurs, A. (1995). Judging the quality of research in business schools: the UK as a case study. *Omega*, *23*(3), 257–270. Retrieved from http://www.sciencedirect.com/science/article/pii/030504839500013E

Dunbar, A. E., & Weber, D. P. (2014). What influences accounting research? A citations-based analysis. *Issues in Accounting Education*, *29*(1), 1–60. http://doi.org/10.2308/iace-50603

Elwyn, G., O'Connor, A., Stacey, D., Volk, R., Edwards, A., Coulter, A., … Whelan, T. (2006). Developing a quality criteria framework for patient decision aids: online international Delphi consensus process. *BMJ (Clinical Research Ed.)*, *333*(7565), 417. http://doi.org/10.1136/bmj.38926.629329.AE

European Science Foundation. (2011). *The European Code of Conduct for Research Integrity*. Strasbourg. Retrieved from http://www.esf.org/fileadmin/Public_documents/Publications/Code_Conduct_ResearchIntegrity.pdf

Evans, J. H., Feng, M., Hoffman, V. B., Moser, D. V., & Van der Stede, W. A. (2015). Points to Consider When Self-Assessing Your Empirical Accounting Research. *Contemporary Accounting Research*, *32*(3), 1162–1192. http://doi.org/10.1111/1911-3846.12133

FAPESP. (2011). Boas práticas científicas: carta de encaminhamento da Diretoria Científica ao Conselho Superior da FAPESP. São Paulo: Fundação de Amparo à Pesquisa do Estado de São Paulo.

FAPESP. (2012a). *Code of good scientific practice*. Sao Paulo: Fundação de Amparo à Pesquisa do Estado de São Paulo.

FAPESP. (2012b). *Código de boas práticas científicas*. São Paulo: Fundação de Amparo à Pesquisa do Estado de São Paulo.

FAPESP. Portaria PR Nº 05 / 2013 - Dispõe sobre a divulgação das práticas de más condutas científicas apuradas pela FAPESP (2013). Brasil: Fundação de Amparo à Pesquisa do Estado de São Paulo.

Flick, U. (2009). *Qualidade na pesquisa qualitativa* (Coleção Pe). Porto Alegre: Artmed.

Garfield, E. (1972). Citation Analysis as a Tool in Journal Evaluation. *Science (New York, N.Y.)*, *178*(60), 471–479.

Garfield, E. (1999). Journal impact factor: a brief review. *CMAJ: Canadian Medical Association Journal = Journal de l'Association Medicale Canadienne*, *161*(8), 979–980.

Gephart, R. (1999). Paradigms and research methods. *Academy of Management Research Methods Forum*, *4*(Summer 1999).

Given, L. M. (2008). *The Sage Encyclopedia of Qualitative Research Methods*. *The MIT Press*. http://doi.org/10.4135/9781412963909

Godoy, A. S. (2005). Refletindo sobre critérios de qualidade da pesquisa qualitativa. *Revista Eletrônica de Gestão Organizacional*, *3*(2, mai/ago.), 81–89. Retrieved from http://www.revista.ufpe.br/gestaoorg/index.php/gestao/article/viewFile/136/118

Graham, B., Regehr, G., & Wright, J. G. (2003). Delphi as a method to establish consensus for diagnostic criteria. *Journal of Clinical Epidemiology*, *56*(12), 1150–1156. http://doi.org/10.1016/S0895-4356(03)00211-7

Greene, J. C. (2007). *Mixed Methods in Social Inquiry* (1st ed.). San Francisco: Wiley.

Greene, J. G. (1988). Stakeholder Participation and Utilization in Program Evaluation. *Evaluation Review*, *12*(2), 91–116. http://doi.org/10.1177/0193841X8801200201

Guba, E. G., & Lincoln, Y. S. (1994). Competing paradigms in qualitiative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 105–117). Sage Publications.

Healy, M., & Perry, C. (2000). Comprehensive criteria to judge validity and reliability of qualitative research within the realism paradigm. *Qualitative Market Research: An International Journal*, *3*(3), 118–126. http://doi.org/10.1108/13522750010333861

Hostetler, K. (2005). What Is "Good" Education Research? *Educational Researcher*, *34*(6), 16–21. http://doi.org/10.3102/0013189X034006016

Hug, S. E., Ochsner, M., & Daniel, H. D. (2013). Criteria for assessing research quality in the humanities: A delphi study among scholars of english literature, german literature and art history. *Research Evaluation*, *22*(August), 369–383. http://doi.org/10.1093/reseval/rvt008

Jones, M. J., Brinn, T., & Pendlebury, M. (1996). Judging the quality of research in business schools: A comment from accounting. *Omega*, *24*(5), 597–602. http://doi.org/10.1016/0305-0483(96)00036-9

Jones, M. J., & Xiao, J. Z. (2004). Financial reporting on the Internet by 2010: A consensus view. *Accounting Forum*, *28*(July 2001), 237–263. http://doi.org/10.1016/j.accfor.2004.07.002

Justo, C. (2005). A Técnica Delphi de Formação de Consensos. Retrieved from http://www.opss.pt/sites/opss.pt/files/A-Tecnica-Delphi-de-Formacao-de-Consensos.pdf

Kalichman, M., & Plemmons, D. (2007). Reported goals for responsible conduct of research courses. *Academic Medicine*, *82*(9), 846–852. Retrieved from http://journals.lww.com/academicmedicine/Abstract/2007/09000/Reported_Goals_for_Responsible_Conduct_of_Research.6.aspx

Kelbaugh, B. M. (2003). *Exploration of teamwork in extension: identifying indicators of success using a modified delphi study (Ph.D. dissertation)*. The Ohio State University. Retrieved from https://etd.ohiolink.edu/ap:10:0:::10:P10_ACCESSION_NUM:osu1054604099

Kim, K. K., O'Bryan, C. a., Crandall, P. G., Ricke, S. C., & Neal, J. a. (2013). Identifying baseline food safety training practices for retail delis using the Delphi expert consensus method. *Food Control*, *32*(1), 55–62. http://doi.org/10.1016/j.foodcont.2012.10.009

Le Pair, C. (1995). Formal evaluation methods: their utility and limitations. *International Forum on Information and Documentation*, *20*, 16–24 ST – Formal evaluation methods: Their utili. Retrieved from <Go to ISI>://WOS:A1995TV97400002

LeCompte, M. D., & Goetz, J. P. (1982). Problems of Reliability and Validity in Ethnographic Research. *Review of Educational Research*, *52*(1), 31–60. http://doi.org/10.3102/00346543052001031

Libby, R., Bloomfield, R., & Nelson, M. W. (2002). Experimental research in financial accounting. *Accounting, Organizations and Society*, *27*(8), 775–810. http://doi.org/10.1016/S0361-3682(01)00011-3

Lowe, A., & Locke, J. (2005). Perceptions of journal quality and research paradigm: results of a web-based survey of British accounting academics. *Accounting, Organizations and Society*, *30*(1), 81–98. http://doi.org/10.1016/j.aos.2004.05.002

Lowensohn, S., & Samelson, D. P. (2006). An Examination of Faculty Perceptions of Academic Journal Quality within Five Specialized Areas of Accounting Research. *Issues in Accounting Education*, *21*(3), 219–239.

Martins, O. S., & Lucena, W. G. L. (2014). Produtivismo acadêmico: as práticas dos docentes dos programas de pós-graduação em contabilidade. *Advances in Scientific and Applied Accounting*, *7*(1), 66–96. http://doi.org/10.14392/asaa.2014070103

Mays, N., & Pope, C. (1995). Qualitative Research: Rigour and qualitative research. *BMJ*, *311*(6997), 109–112. http://doi.org/10.1136/bmj.311.6997.109

Mays, N., & Pope, C. (2006). Quality in qualitative health research. In *Qualitative research in health care* (3rd ed., pp. 82–101). London: Blackwell Pub./BMJ Books. Retrieved from http://eprints.soton.ac.uk/64477/

McGrath, J. E., & Brinberg, D. (1983). External Validity and the Research Process : A Comment on the Calder / Lynch Dialogue. *Journal of Consumer Research*, *10*(1), 115–124.

Mendonça Neto, O. R. de O., Riccio, E. L., & Sakata, M. C. G. (2009). Dez anos de pesquisa contábil no Brasil: análise dos trabalhos apresentados nos Enanpads de 1996 a 2005. *Revista de Administração de Empresas*, *49*(1), 62–73.

Meyrick, J. (2006). What is good qualitative research? A first step towards a comprehensive approach to judging rigour/quality. *Journal of Health Psychology*, *11*(5), 799–808. http://doi.org/10.1177/1359105306066643

Michaelis. (2012). *Michaelis Moderno Dicionário da Língua Portuguesa (versão online)*. Editora Melhoramentos. Retrieved from http://michaelis.uol.com.br/moderno/portugues/index.php

Miranda, G. J., Azevedo, R. F. L., & Martins, G. de A. (2011). Teses das teses em contabilidade na USP. *REPEC Revista de Educação E Pesquisa Em Contabilidade*, *5*(2, mai/ago), 21–42.

Moizer, P. (2009). Publishing in accounting journals: A fair game? *Accounting, Organizations and Society*, *34*(2), 285–304. http://doi.org/10.1016/j.aos.2008.08.003

Morell, J. A. (2010). Logic Models - Beyond the Traditional View: Metrics, Methods, Expected and Unexpected Change. In *American Evaluation Association Annual Meeting – November 2-5*. Anaheim, California: American Evaluation Association.

National Institutes of Health. (1979). The Belmont Report. *The Belmont Report Ethical Principles and Guidelines for the Protection of Human Subjects of Research*, 4–6.

Retrieved from http://www.salesianos-cadiz.com/Descargas/Escolar/ESO-Bachillerato/Bachillerato/2do Bachillerato/2do A/InformeBelmont.pdf

Northcott, D., & Linacre, S. (2010). Producing spaces for academic discourse: The impact of research assessment exercises and journal quality rankings. *Australian Accounting Review*, *20*, 38–54. http://doi.org/10.1111/j.1835-2561.2010.00079.x

OADS. (2012). *CDC Guidance on Scientific Integrity*. Retrieved from http://www.cdc.gov/od/science/docs/CDCSIGuide_021612.pdf

OHRP, & DHHS. (2009). Code of Federal Regulations - Title 45: Public Welfare and Title 46: Protection of Human Subjects. *45 CFR 46*, 1–14. Retrieved from papers2://publication/uuid/9C582C3B-CB51-4C48-A6D0-4AC885DDB8FA

Patton, M. Q. (2002). *Qualitative research & evaluation methods* (3rd ed.). Thousand Oaks: Sage Publications.

Pinto, L. A. (2008). Cientometria: é possível avaliar qualidade da pesquisa científica? *Scientia*, *18*(2, abr./jun.), 64–65.

Pozzebon, M. (2004). Conducting and evaluating critical interpretative research: examining criteria as a key component in building a research tradition. *Information Systems Research: Relevant Theory and …*, (July 2002), 1–20. Retrieved from http://link.springer.com/content/pdf/10.1007/1-4020-8095-6_16.pdf

Reiter, S. A., & Williams, P. F. (2002). The structure and progressivity of accounting research: the crisis in the academy revisited. *Accounting, Organizations and Society*, *27*(6), 575–607. http://doi.org/10.1016/S0361-3682(01)00050-2

Rodríguez-Mañas, L., Féart, C., Mann, G., Viña, J., Chatterji, S., Chodzko-Zajko, W., … Vega, E. (2013). Searching for an operational definition of frailty: A delphi method based consensus statement. the frailty operative definition-consensus conference project. *Journals of Gerontology - Series A Biological Sciences and Medical Sciences*, *68*(1), 62–67. http://doi.org/10.1093/gerona/gls119

Rossi, P. H., Lipsey, M. W., & Freeman, H. E. (2004). *Evaluation: a systematic approach* (7th ed.). Thousand Oaks: Sage Publications.

Rynes, S., & Gephart, R. P. (2004). From the editors: Qualitative Research and the Academy of Management Journal. *Academy of Management Journal*, *47*(4), 454–462. http://doi.org/10.5465/AMJ.2004.14438580

Sakamoto, A. (2011). *Inovação em rede de PMEs: fatores determinantes e consequências sociais*. Fundação Getúlio Vargas. Retrieved from http://bibliotecadigital.fgv.br/dspace/handle/10438/8188

Schwartzman, S. (1988). Qualidade e relevância da pesquisa universitária: 10 teses para discussão. Retrieved from http://www.schwartzman.org.br/simon/10teses.htm

Shamoo, A., & Resnik, D. (2003). *Responsible conduct of research*. New York: Oxford University Press. Retrieved from http://books.google.com/books?hl=en&lr=&id=RLZin-f9eooC&oi=fnd&pg=PA3&dq=responsible+conduct+of+research&ots=RGF5kYyHAc&

sig=qubYPUXknzdQJEw7JEdDeIabm9s

Skulmoski, G. J., Graham, F. T., & Krahn, J. (2007). The Delphi method for graduate research. *Journal of Information Technology Education*, *6*, 1–21. Retrieved from http://bern.library.nenu.edu.cn/upload/soft/0-article/+03/JITEv6p001-021Skulmoski212.pdf

Smith, J. K. (1984). Judging Interpretive Inquiry. *Educational Evaluation and Policy Analysis*, *6*(4), 379–391.

Spencer, L., Ritchie, J., Lewis, J., & Dillon, L. (2003). Quality in Qualitative Evaluation : A framework for assessing research evidence A Quality Framework. *Social Research*, *108*(21/08/05), 170.

Steneck, N. H. (2007). *ORI Introduction to the Responsible Conduct of Research*. Department of Health & Human Services-USA.

Stufflebeam, D. L. (2000). Lessons in Contracting for Evaluations. *American Journal of Evaluation*, *21*(3), 293–314. http://doi.org/10.1177/109821400002100302

Taylor-Powell, E., & Henert, E. (2008). *Developing a logic model : Teaching and training guide*.

Taylor, J. (2011). The assessment of research quality in UK universities: Peer review or metrics? *British Journal of Management*, *22*(2), 202–217. http://doi.org/10.1111/j.1467-8551.2010.00722.x

Tetzlaff, J. M., Moher, D., & Chan, A.-W. (2012). Developing a guideline for clinical trial protocol content: Delphi consensus survey. *Trials*, *13*, 176. http://doi.org/10.1186/1745-6215-13-176

Theóphilo, C., & Iudícibus, S. (2005). Uma análise crítico-epistemológica da produção científica em contabilidade no Brasil. *UnB Contábil*, *8*(2, jul/dez), 147–175. Retrieved from http://www.cgg-amg.unb.br/index.php/contabil/article/viewArticle/164

Uncles, M. D., & Kwok, S. (2013). Designing research with in-built differentiated replication. *Journal of Business Research*, *66*(9), 1398–1405. http://doi.org/10.1016/j.jbusres.2012.05.005

Valentine, J. C. (2009). Judging the quality of primary research. In H. Cooper, L. V Hedges, & J. C. Valentine (Eds.), *The Handbook of Research Synthesis and Meta-Analysis* (pp. 129–146). Russell Sage Foundation.

Van de Ven, A. H., & Poole, M. S. (2005). Alternative Approaches for Studying Organizational Change. *Organization Studies*, *26*(9), 1377–1404. http://doi.org/10.1177/0170840605056907

Van der Stede, W. a., Young, S. M., & Chen, C. X. (2005). Assessing the quality of evidence in empirical management accounting research: The case of survey studies. *Accounting, Organizations and Society*, *30*(7-8), 655–684. http://doi.org/10.1016/j.aos.2005.01.003

Webster, J., & Watson, R. T. (2002). Analyzing the Past to Prepare for the Future: Writing a

Literature Review. *MIS Quarterly*, *26*, xiii – xxiii. http://doi.org/10.1.1.104.6570

Weiss, C. H. (1997). *Evaluation* (2nd ed.). New Jersey: Prentice Hall.

Whittemore, R., Chase, S. K., & Mandle, C. L. (2001). Validity in qualitative research. *Qualitative Health Research*, *11*(4), 522–537. http://doi.org/10.1177/104973201129119299

Williams, P. F. (2014). The Myth of Rigorous Accounting Research. *Accounting Horizons*, *28*(4), 869–887. http://doi.org/10.2308/acch-50880

Worrell, J. L., Di Gangi, P. M., & Bush, A. A. (2013). Exploring the use of the Delphi method in accounting information systems research. *International Journal of Accounting Information Systems*, *14*(3), 193–208. http://doi.org/10.1016/j.accinf.2012.03.003

Yarbrough, D. B., Shulha, L. M., Hopson, R. K., & Caruthers, F. A. (2011). *The program evaluation standards: a guide for evaluators and evaluation users* (3rd ed.). Thousand Oaks: Sage Publications.

Appendix A: Matrix of Orientation for Data Collection Instruments (Delphi Online)

N.	Stage	Dom.	Key Feature Spencer et al.	Criterion	Focus	Source(s)	Question (Spencer et al)	Attribute/ Relationship	Statement for Delphi
1	One	S	Design	Suitability	Usefulness	Denscombe	How defensible is the research design?	Useful strategy for purpose	My studies show reasons/justification for choosing the research strategy.
2	One	S	Design	Suitability	Appropriateness	Denscombe	How defensible is the research design?	Clear overview that guides the study	I am able to recognize the assumptions of the overview (paradigm or theoretical lens) that guides the building of my studies.
3	One	S	Design	Suitability	Appropriateness	Denscombe	How defensible is the research design?	Overview x strategy	I consider a particular overview and its assumptions when I choose the strategy to build my research design.
4	One	S	Design	Suitability	Appropriateness	Denscombe	How defensible is the research design?	Goal/problem shown precisely	My studies show precisely the goal/problem that they want to reach or solve.
5	One	C	Design	Suitability	Appropriateness	Denscombe	How defensible is the research design?	Overview x theory	I consider a particular paradigm and its assumptions when I choose the theory to support my research (theoretical platform).
6	One	S	Design	Suitability	Usefulness	Denscombe	How defensible is the research design?	Arguments for technique choices	I present arguments for choosing techniques and procedures to gather and analyze the data for my research.
7	One	M	Design	Feasibility	Access to data sources	Denscombe	How defensible is the research design?	Access to the data	The research strategy is chosen considering the access to the data sources that are required to perform the research.
8	One	M	Design	Feasibility	Time constraints	Denscombe	How defensible is the research design?	Time constraints	I choose the research strategy based on the time required to collect the data and present the results in accordance with the purpose of the research.
9	One	S	Design	Feasibility	Audience	Denscombe	How defensible is the research design?	Research strategy x target audience	I consider the profile of the likely audience of my study when I choose the research strategy.
10	One	M	Research subjects	Internal validity / credibility or defensibility	Representativeness	André, Brinberg & Mcgrath, Spencer et al	How well defended is the sample design/target collection of cases/documents?	Criteria to design/select subjects	I present the criteria used to design/select the research subjects explicitly based on the population of interest.
11	One	M	Research subjects	Internal validity / credibility or defensibility	Logical coherence	Spencer et al, Brinberg & Mcgrath	Sample composition/case inclusion – how well is the eventual coverage described?	Representativeness of subjects	My studies formally describe methods and inclusion/exclusion cases that might have an impact on the representativeness of the research subjects.

N.	Stage	Dom.	Key Feature Spencer et al.	Criterion	Focus	Source(s)	Question (Spencer et al)	Attribute/ Relationship	Statement for Delphi
12	Two	M	Data collection	Rigor / thoroughness	Methodological rigor	Spencer et al, Brinberg & Mcgrath	How well was the data collection carried out?	Notes for each research step	Typically, I take notes on the steps and actions performed during data collection.
13	Two	M	Data collection	Rigor / thoroughness	Methodological rigor	Spencer et al, Brinberg & Mcgrath	How well was the data collection carried out?	Notes for divergent events	I typically write down details about events that diverge from the research design that arise during data collection.
14	Two	M	Data collection	Integrity	Ethics	Denscombe	What evidence is there of attention to ethical issues?	Voluntary participation of subjects	When my study involves human beings, they are invited, and they participate only voluntarily.
15	Two	M	Data collection	Integrity	Ethics	Denscombe	What evidence is there of attention to ethical issues?	Formal agreement from humans	All the participants of my studies that involve human beings agree to and sign (or expressly declare) informed consent.
16	Two	M	Analysis	Integrity	Logic of inquiry	Spencer et al, Brinberg & Mcgrath	How well has the approach to and formulation of the analysis been conveyed?	Description of nature and form of data	My studies show a detailed description of the nature and form of the collected original data.
17	Two	M	Analysis	Rigor / thoroughness	Methodological rigor	Spencer et al, Brinberg & Mcgrath	How well has the approach to and formulation of the analysis been conveyed?	Description of tools and procedures	Tools and procedures for data analysis are carefully presented.
18	Two	S	Analysis	Reliability / auditability	Consistency	Spencer et al, Brinberg & Mcgrath	Contexts of data sources – how well are they retained and portrayed?	Context x impact on data analysis	I show details of the context in which the data were collected that might have/have had an impact on the data analysis.
19	Two	S	Analysis	Reliability / auditability	Consistency	Spencer et al, Brinberg & Mcgrath	How well has diversity of perspective and content been explored?	Other views provided to explain context	I present other views/perspectives so that the reader may better understand the context of the analysis.
20	Two	S	Analysis	Relevance	Significance	André, Brinberg & Mcgrath, Spencer et al	How well has detailed, depth and complexity (i.e. richness) of the data been conveyed?	Significance of data to reaching goals	I discuss the significance of the data to achieving the study's goals.
21	Two	S	Analysis	Rigor / thoroughness	Logic of inquiry	André, Brinberg & Mcgrath, Spencer et al	How well has detailed, depth and complexity (i.e. richness) of the data been conveyed?	Implicit/explicit links - findings x aims	I discuss implicit and explicit links between data/findings and the purpose of my studies.
22	Two	C	Findings	Internal validity / credibility or defensibility	Logical coherence	Spencer et al, Mays & Pope	How credible are the findings?	Path to achieve conclusions	My studies clearly show the target audience how I arrived at my conclusions from the findings reported.
23	Two	C	Findings	Internal validity / credibility or defensibility	Support for data and evidence	Spencer et al, Mays & Pope	How credible are the findings?	Link between findings x evidence	The conclusions of my studies explicitly present the correspondence between findings and the data/evidence.
24	Two	C	Findings	External validity or generalization / fittingness	Applicability	Spencer et al, Mays & Pope	How credible are the findings?	Comparison of results with those of other studies	I use the empirical findings of previous studies to compare with my results (when applicable).

N.	Stage	Dom.	Key Feature Spencer et al.	Criterion	Focus	Source(s)	Question (Spencer et al)	Attribute/ Relationship	Statement for Delphi
25	Two	S	Findings	Contribution / quality of the theoretical perspective	New areas for future studies	Spencer et al, Mays & Pope	How has knowledge/understanding been extended by the research?	New areas based on the findings	My research highlights new areas for investigation based on findings.
26	Two	S	Findings	Contribution / quality of the theoretical perspective	Advancement of knowledge	Spencer et al, Mays & Pope	How has knowledge/understanding been extended by the research?	Insights for thinking about the field	Findings are presented in a way that creates perspectives (insights) of thinking about the field.
27	Two	S	Findings	Relevance	Gap	Spencer et al, Brinberg & Mcgrath	How well does the evaluation address its original aims and purpose?	Previous findings x hypothesis	I use empirical findings from previous studies to justify my propositions or hypotheses (when applicable).
28	Two	M	Findings	Internal validity / credibility or defensibility	Logical coherence	Spencer et al, Brinberg & Mcgrath	How well does the evaluation address its original aims and purpose?	Check links of findings x purpose	I check whether my findings are clearly linked to the purpose of my studies.
29	Two	S	Findings	Reliability / auditability	Robustness	Spencer et al, Brinberg & Mcgrath	Scope for drawing wider inference – how well is this explained?	Does the context allow replication	To allow future replications, I describe in detail the context in which my studies were conducted.
30	Two	M	Findings	Internal validity / credibility or defensibility	Logical coherence	Spencer et al, Mays & Pope	How clear is the basis of evaluative appraisal?	Impact from the nature of any divergences	My studies precisely discuss the nature and sources of possible divergences in the data that might impact the conclusions.
31	Two	S	Reporting	Impact	Advancement of knowledge	Spencer et al, Mays & Pope	How has knowledge/understanding been extended by the research?	Discussion of impact on knowledge	My reports discuss the impacts of findings to clearly expand knowledge.
32	Two	M	Reporting	Contribution / quality of the theoretical perspective	Limitations	Spencer et al, Brinberg & Mcgrath	How has knowledge/understanding been extended by the research?	Disclosure of limitations	My reports explicitly include information about the limitations of my findings.
33	Two	M	Reporting	Integrity	Limitations	Spencer et al, Brinberg & Mcgrath	How well does the evaluation address its original aims and purpose?	Reasons for limitations	I present a discussion about the reasons for my study's limitations.
34	Two	M	Reporting	Contribution / quality of the theoretical perspective	Boundaries	Spencer et al, Brinberg & Mcgrath	How has knowledge/understanding been extended by the research?	Boundaries of the study	My research reports formally describe the boundaries (scope) of my studies.
35	Two	C	Reporting	Internal validity / credibility or defensibility	Logical coherence	Spencer et al, Brinberg & Mcgrath	How clear and coherent is the reporting?	Literature review x main concepts	My literature reviews describe the central concepts inherent in the studied topics.
36	Two	C	Reporting	Internal validity / credibility or defensibility	Logical coherence	Brinberg & McGrath	How clear and coherent is the reporting?	Theory to support propositions	I provide theory-based explanations to justify the propositions in my studies.

N.	Stage	Dom.	Key Feature Spencer et al.	Criterion	Focus	Source(s)	Question (Spencer et al)	Attribute/ Relationship	Statement for Delphi
37	Two	M	Reporting	Internal validity / credibility or defensibility	Logical coherence	Spencer et al, Brinberg & Mcgrath	How well does the evaluation address its original aims and purpose?	Conclusions x aim	My conclusions show explicit linkage to my study's goals.
38	Two	M	Reporting	External validity or generalization / fittingness	Scope	Spencer et al, Mays & Pope	Scope for drawing wider inference – how well is this explained?	Explicit possibility of generalization	The possibility of generalization (when applicable) is explicitly evinced based on the scope of my study.
39	Two	M	Reporting	Internal validity / credibility or defensibility	Logical coherence	Brinberg & McGrath	How clear are the links between data, interpretation and conclusions – i.e. how well can the route to any conclusions be seen?	Appreciation of drafts by colleagues	I submit my drafts to colleagues before making a final submission to a journal/conference.
40	Three	M	Reflexivity and neutrality	Integrity	Robustness	Spencer et al, Mays & Pope	How clear are the assumptions/ theoretical perspectives/values that have shaped the form and output of the evaluation?	How to address errors and biases	I explicitly discuss in my studies any errors or biases may have arisen during the search process and how I managed them.
41	Three	M	Reflexivity and neutrality	Integrity	Robustness	Spencer et al, Brinberg & Mcgrath	How clear are the assumptions/ theoretical perspectives/values that have shaped the form and output of the evaluation?	Impact of team participation on the research	I explicitly discuss the possible impact of my participation (and that of my team) in the various phases of the research process (where applicable).
42	All	M	Ethics	Rigor / thoroughness	Ethics	Denscombe, CNS, Brasil (2012, 2013)	What evidence is there of attention to ethical issues?	Registration of the study with the Ethics Committee (CEP)	My institution/academic unit has a Research Ethics Committee to register research designs.
43	All	M	Ethics	Rigor / thoroughness	Ethics	Denscombe, CNS, Brasil (2012, 2013)	What evidence is there of attention to ethical issues?	Rules for registering studies with the CEP	My institution/academic unit has clear rules about submitting research designs to the Research Ethics Committee.
44	All	M	Ethics	Rigor / thoroughness	Ethics	Denscombe, CNS, Brasil (2012, 2013)	What evidence is there of attention to ethical issues?	Formal submission to the CEP	My studies that involve human beings are formally submitted to the Research Ethics Committee.
45	All	M	Ethics	Rigor / thoroughness	Ethics	Denscombe, CNS, Brasil (2012, 2013), André	What evidence is there of attention to ethical issues?	Compulsory submission to the CEP	The submission of the research projects to Research Ethics Committee is compulsory in my university/academic unit.
46	All	M	Ethics	Integrity	Code of Ethics	Denscombe	What evidence is there of attention to ethical issues?	Use of the principles of an ethics code	I use principles of ethics in research, as formally described in an Ethics Code (of a development institution, for example), to guide my studies.
47	All	M	Ethics	Integrity	Respect for human subjects	Denscombe, National Institute of Health, Conama, FAPESP.	What evidence is there of attention to ethical issues?	Formal respect for human subjects	I provide informed consent to human subjects involved in my studies, with details about the research purpose, risks, and other aspects foreseen by the legislation.

N.	Stage	Dom.	Key Feature Spencer et al.	Criterion	Focus	Source(s)	Question (Spencer et al)	Attribute/ Relationship	Statement for Delphi
48	All	M	Ethics	Integrity	Respect for human subjects	Denscombe, National Institute of Health, CNS, FAPESP.	What evidence is there of attention to ethical issues?	Strategy to mitigate possible harm	I take care to choose the research strategy that brings the least harm to the subjects involved in my studies.
49	All	M	Ethics	Integrity	Confidentiality	Denscombe, National Institute of Health, CNS, FAPESP.	What evidence is there of attention to ethical issues?	Confidentiality to participants' data	My studies provide specific procedures to treat participants' data in a confidential manner.
50	All	M	Auditability	Reliability / auditability	Methodological rigor	Denscombe, Spencer et al., Mays & Pope	How adequately has the research process been documented?	Records of design changes	I safely store complete records of changes to the design of investigations that have altered the scope/extent of the research process.
51	All	M	Auditability	Reliability / auditability	Consistency	Denscombe, Spencer et al., Mays & Pope	How adequately has the research process been documented?	Records of the reasons for changes	I safely store complete records of the reasons for the changes made to the studies' design.
52	All	M	Auditability	Reliability / auditability	Consistency	Denscombe, Spencer et al., Mays & Pope	How adequately has the research process been documented?	Safeguard databases for checks	For future checks, I save databases of my studies with an adequate description of their composition.
53	All	M	Auditability	Reliability / auditability	Methodological rigor	Denscombe, Spencer et al., Mays & Pope	How adequately has the research process been documented?	Guard documents to reduce risks	The informed consents signed (or formally stated) by each individual participant in my research are safely stored to preserve the subjects' consent to possible risks.

Note: Only a few sources were included in the matrix. Other sources that have provided support are explored in the literature review.

Appendix B: Matrix of Orientation for Data Collection Instruments (Delphi Online) (in Portuguese)

MATRIZ DE ORIENTAÇÃO PARA A CONSTRUÇÃO DO INSTRUMENTO DE COLETA DE DADOS (DELPHI ONLINE)

N.	Fase	Dom.	Caract. chave Spencer et al.	Critério	Foco	Fonte (s)	Questão (Spencer et al)	Atributo/Relação	Proposição para a Delphi
1	Um	S	Desenho	Adequabilidade	Utilidade	Denscombe	Quão defensável é o desenho da pesquisa?	Estratégia útil ao propósito	Meus estudos apresentam razões/justificativas para a escolha da estratégia de pesquisa.
2	Um	S	Desenho	Adequabilidade	Adequação	Denscombe	Quão defensável é o desenho da pesquisa?	Clara visão geral que orienta o estudo	Sou capaz de reconhecer os pressupostos da visão geral (paradigma ou lente teórica) que guiam a construção dos meus estudos.
3	Um	S	Desenho	Adequabilidade	Adequação	Denscombe	Quão defensável é o desenho da pesquisa?	Visão geral x estratégia	Levo em conta um paradigma e seus pressupostos quando escolho a estratégia para construir o desenho das minhas pesquisas.
4	Um	S	Desenho	Adequabilidade	Adequação	Denscombe	Quão defensável é o desenho da pesquisa?	Objetivo/problema evidenciado precisamente	Meus estudos apresentam, precisamente, o objetivo/problema que se pretende alcançar ou resolver.
5	Um	C	Desenho	Adequabilidade	Adequação	Denscombe	Quão defensável é o desenho da pesquisa?	Visão geral x teoria	Levo em conta um paradigma e seus pressupostos quando escolho a teoria para suportar meus estudos (plataforma teórica).
6	Um	S	Desenho	Adequabilidade	Utilidade	Denscombe	Quão defensável é o desenho da pesquisa?	Argumentos para a escolha da técnica	Apresento argumentos para a escolha das técnicas e procedimentos para coleta e análise dos dados das minhas pesquisas.
7	Um	M	Desenho	Viabilidade	Acesso às fontes de dados	Denscombe	Quão defensável é o desenho da pesquisa?	Acesso aos dados	A estratégia de pesquisa é escolhida considerando o acesso às fontes de dados para realizar os meus estudos.
8	Um	M	Desenho	Viabilidade	Restrições de tempo	Denscombe	Quão defensável é o desenho da pesquisa?	Restrições de tempo	Escolho uma estratégia de pesquisa baseado no tempo necessário para coletar os dados e apresentar os resultados.
9	Um	S	Desenho	Viabilidade	Público-alvo	Denscombe	Quão defensável é o desenho da pesquisa?	Estratégia de pesquisa x público-alvo	Levo em conta o provável público-alvo dos meus estudos quando escolho a estratégia da pesquisa.
10	Um	M	Sujeitos da Pesquisa	Validade interna/ credibilidade ou defensibilidade	Representatividade	André, Brinberg & Mcgrath, Spencer et al	Quão bem defendido é o desenho da amostra/coleção de casos-alvo/ documentos?	Critérios para desenho/seleção de sujeitos	Apresento, explicitamente, os critérios usados para desenhar/selecionar os sujeitos da pesquisa baseada na população de interesse.
11	Um	M	Sujeitos da Pesquisa	Validade interna/ credibilidade ou defensibilidade	Coerência lógica	Spencer et al, Brinberg & Mcgrath	Composição da amostra / inclusão de casos - o quão bem descrita está a eventual cobertura?	Representatividade dos sujeitos	Meus estudos descrevem, formalmente, os métodos e os casos de inclusão/exclusão que têm impacto sobre a representatividade dos sujeitos da pesquisa.

N.	Fase	Dom.	Caract. chave Spencer et al.	Critério	Foco	Fonte (s)	Questão (Spencer et al)	Atributo/Relação	Proposição para a Delphi
12	Dois	M	Coleta de dados	Rigor/ meticulosidade	Rigor metodológico	Spencer et al, Brinberg & Mcgrath	Quão bem realizada foi a coleta de dados?	Anotações de cada passo da pesquisa	Normalmente, tomo nota dos passos e ações realizadas durante a coleta de dados.
13	Dois	M	Coleta de dados	Rigor/ meticulosidade	Rigor metodológico	Spencer et al, Brinberg & Mcgrath	Quão bem realizada foi a coleta de dados?	Anotações dos eventos divergentes	Normalmente, tomo nota de detalhes das ocorrências divergentes do desenho da pesquisa que aparecem durante a coleta de dados.
14	Dois	M	Coleta de dados	Integridade	Ética	Denscombe	Que evidências existem de atenção às questões éticas?	Participação voluntária dos sujeitos	Quando meus estudos envolvem seres humanos, estes são convidados e participarem voluntariamente.
15	Dois	M	Coleta de dados	Integridade	Ética	Denscombe	Que evidências existem de atenção às questões éticas?	Concordância formal dos sujeitos humanos	Todos os participantes dos meus estudos que envolvam seres humanos assinam (ou declaram expressamente) um Termo de Consentimento Livre e Esclarecido.
16	Dois	M	Análise	Integridade	Lógica de investigação	Spencer et al, Brinberg & Mcgrath	Quão bem transmitida foi a abordagem para e a formulação da análise?	Descrição da natureza e forma dos dados	Meus estudos apresentam uma descrição detalhada da natureza e forma dos dados originais coletados.
17	Dois	M	Análise	Rigor/ meticulosidade	Rigor metodológico	Spencer et al, Brinberg & Mcgrath	Quão bem transmitida foi a abordagem para e a formulação da análise?	Descrição das ferramentas e procedimentos	Ferramentas e procedimentos para análise de dados são, cuidadosamente, apresentados.
18	Dois	S	Análise	Confiabilidade/ auditabilidade	Consistência	Spencer et al, Brinberg & Mcgrath	Contextos de fontes de dados - o quão bem mantidos e retratados eles estão?	Contexto x impacto sobre a análise dos dados	Explicito os detalhes do contexto em que os dados foram coletados e o que pode ter/tem impacto na análise de dados.
19	Dois	S	Análise	Confiabilidade/ auditabilidade	Consistência	Spencer et al, Brinberg & Mcgrath	Quão bem exploradas foram a diversidade de perspectivas e o conteúdo?	Visões alternativas para entender o contexto	Apresento outras visões/perspectivas possíveis para que o leitor possa entender melhor o contexto da análise dos meus dados.
20	Dois	S	Análise	Relevância	Significância	André, Brinberg & Mcgrath, Spencer et al	Quão bem detalhada e transmitida foi a profundidade e complexidade (i.e. riqueza) dos dados?	Significância dos dados para alcançar o objetivo	Discuto a significância dos dados para alcançar os objetivos em meus estudos.
21	Dois	S	Análise	Rigor/ meticulosidade	Lógica de investigação	André, Brinberg & Mcgrath, Spencer et al	Quão bem detalhada e transmitida foi a profundidade e complexidade (i.e. riqueza) dos dados?	Links implícitos/ explícitos - achados x objetivos	Discuto os links, implícitos e explícitos, entre os dados/achados e o propósito nos meus estudos.
22	Dois	C	Achados	Validade interna/ credibilidade ou defensibilidade	Coerência lógica	Spencer et al, Mays & Pope	Quão críveis são os achados?	Caminho para chegar às conclusões	Meus estudos apresentam ao público-alvo, de maneira clara, como cheguei às conclusões a partir dos achados reportados.
23	Dois	C	Achados	Validade interna/ credibilidade ou defensibilidade	Apoio em dados e evidências	Spencer et al, Mays & Pope	Quão críveis são os achados?	Links entre achados x evidências	As conclusões dos meus estudos apresentam, de maneira explícita, a correspondência entre achados e evidências/dados.
24	Dois	C	Achados	Validade externa ou generalização/ conformidade	Aplicabilidade	Spencer et al, Mays & Pope	Quão críveis são os achados?	Comparação de resultados com outros estudos	Uso achados empíricos de estudos anteriores para comparar com os meus achados (quando aplicável).

N.	Fase	Dom.	Caract. chave Spencer et al.	Critério	Foco	Fonte (s)	Questão (Spencer et al)	Atributo/Relação	Proposição para a Delphi
25	Dois	S	Achados	Contribuição/ qualidade da persp. teórica.	Novas áreas para futuros estudos	Spencer et al, Mays & Pope	Como o conhecimento/entendimento foi ampliado pela pesquisa?	Novas áreas baseadas nos achados	Baseados nos achados, meus estudos sugerem novas áreas para investigação.
26	Dois	S	Achados	Contribuição/ qualidade da persp. teórica.	Avanço do conhecimento	Spencer et al, Mays & Pope	Como o conhecimento/entendimento foi ampliado pela pesquisa?	Insights para pensar o campo de conhecimento	Os achados são apresentados de uma maneira que cria perspectivas (insights) de pensar o conhecimento na área.
27	Dois	S	Achados	Relevância	Lacuna/Gap	Spencer et al, Brinberg & Mcgrath	Quão bem a avaliação leva em conta seus objetivos e propósito originais?	Achados anteriores x hipóteses	Uso achados empíricos de estudos anteriores para suportar minhas proposições e/ou hipóteses (quando aplicável).
28	Dois	M	Achados	Validade interna/ credibilidade ou defensibilidade	Coerência lógica	Spencer et al, Brinberg & Mcgrath	Quão bem a avaliação leva em conta seus objetivos e propósito originais?	Checagem da ligação dos achados x propósito	Checo se meus achados estão, claramente, ligados ao propósito dos meus estudos.
29	Dois	S	Achados	Confiabilidade/ auditabilidade	Robustez	Spencer et al, Brinberg & Mcgrath	Quão bom é o escopo para obter maior inferência explicada?	Contexto para permitir replicação	Descrevo em detalhes o contexto no qual os estudos foram conduzidos, a fim de permitir replicações.
30	Dois	M	Achados	Validade interna/ credibilidade ou defensibilidade	Coerência lógica	Spencer et al, Mays & Pope	Quão clara é a base da apreciação valorativa?	Impacto da natureza das fontes	Meus estudos discutem, precisamente, a natureza e a origem de possíveis divergências nos dados que poderiam impactar as conclusões.
31	Dois	S	Relatório	Impacto	Avanço do conhecimento	Spencer et al, Mays & Pope	Como o conhecimento/entendimento foi ampliado pela pesquisa?	Discussão sobre o impacto para o conhecimento	Meus relatórios apresentam, claramente, a discussão sobre o impacto dos achados para a expansão do conhecimento na área.
32	Dois	M	Relatório	Contribuição/ qualidade da persp. teórica.	Limitações	Spencer et al, Brinberg & Mcgrath	Quão claro e coerente é o relatório?	Evidenciação de limitações	Meus relatórios contêm, explicitamente, as limitações dos achados das minhas pesquisas.
33	Dois	M	Relatório	Integridade	Limitações	Spencer et al, Brinberg & Mcgrath	Quão bem a avaliação leva em conta seus objetivos e propósito originais?	Razões das limitações	Apresento discussão acerca das razões que geraram as limitações dos meus estudos.
34	Dois	M	Relatório	Contribuição/ qualidade da persp. teórica.	Fronteiras	Spencer et al, Brinberg & Mcgrath	Como o conhecimento/entendimento foi ampliado pela pesquisa?	Fronteiras do estudo	Meus relatórios de pesquisa evidenciam, formalmente, as fronteiras (alcance) dos achados dos meus estudos.
35	Dois	C	Relatório	Validade interna/ credibilidade ou defensibilidade	Coerência lógica	Spencer et al, Brinberg & Mcgrath	Quão claro e coerente é o relatório?	Revisão de literatura x conceitos principais	As revisões de literatura que faço descrevem os conceitos-chave inerentes aos assuntos estudados.
36	Dois	C	Relatório	Validade interna/ credibilidade ou defensibilidade	Coerência lógica	Brinberg & McGrath	Quão claro e coerente é o relatório?	Teoria para apoiar proposições	Uso explicação baseada na teoria para justificar as proposições dos meus estudos.
37	Dois	M	Relatório	Validade interna/ credibilidade ou defensibilidade	Coerência lógica	Spencer et al, Brinberg & Mcgrath	Quão bem a avaliação leva em conta seus objetivos e propósito originais?	Conclusões x objetivo	Minhas conclusões mostram, explicitamente, a ligação com os objetivos dos estudos.

N.	Fase	Dom.	Caract. chave Spencer et al.	Critério	Foco	Fonte (s)	Questão (Spencer et al)	Atributo/Relação	Proposição para a Delphi
38	Dois	M	Relatório	Validade externa ou generalização/ conformidade	Escopo	Spencer et al, Mays & Pope	Quão bom é o escopo para obter maior inferência explicada?	Possibilidade explícita de generalização	A possibilidade de generalização (quando aplicável) é explicitamente evidenciada, com base no escopo dos meus estudos.
39	Dois	M	Relatório	Validade interna/ credibilidade ou defensibilidade	Coerência lógica	Brinberg & McGrath	Quão claras são as ligações entre dados, interpretação e conclusões – i.e. quão bem pode ser visto o caminho até as conclusões?	Apreciação de rascunhos por colegas	Submeto minhas versões preliminares (drafts) para a apreciação de colegas antes da submissão final a um periódico/evento.
40	Três	M	Reflexividade e neutralidade	Integridade	Robustez	Spencer et al, Mays & Pope	Quão claras são as suposições/perspectivas teóricas/valores que moldaram a forma e saídas da avaliação?	Como lidou com erros e vieses	Discuto em meus estudos, de forma explícita, como erros ou viéses podem ter surgido durante o processo de pesquisa e como lidei com eles.
41	Três	M	Reflexividade e neutralidade	Integridade	Robustez	Spencer et al, Brinberg & Mcgrath	Quão claras são as suposições/perspectivas teóricas/valores que moldaram a forma e saídas da avaliação?	Impactos da participação da equipe sobre a pesquisa	Discuto, de forma explícita, possíveis impactos decorrentes da minha participação (e da minha equipe) nas fases do processo de pesquisa (quando aplicável).
42	Todas	M	Ética	Rigor/ meticulosidade	Ética	Denscombe, CNS, Brasil (2012, 2013)	Que evidências existem de atenção às questões éticas?	Comitê de Ética (CEP) para registrar os estudos	Minha Instituição/Unidade Acadêmica tem um Comitê de Ética em Pesquisa para submissão e aprovação dos projetos de pesquisa.
43	Todas	M	Ética	Rigor/ meticulosidade	Ética	Denscombe, CNS, Brasil (2012, 2013)	Que evidências existem de atenção às questões éticas?	Regras para registrar os estudos no CEP	Minha Instituição/Unidade Acadêmica tem regras claras para submissão dos projetos de pesquisa ao Comitê de Ética em Pesquisa.
44	Todas	M	Ética	Rigor/ meticulosidade	Ética	Denscombe, CNS, Brasil (2012, 2013)	Que evidências existem de atenção às questões éticas?	Submissão formal ao CEP	Meus estudos que envolvem seres humanos são, formalmente, submetidos ao Comitê de Ética em Pesquisa.
45	Todas	M	Ética	Rigor/ meticulosidade	Ética	Denscombe, CNS, Brasil (2012, 2013), André	Que evidências existem de atenção às questões éticas?	Submissão compulsória ao CEP	A submissão dos projetos de pesquisa que envolvem seres humanos ao Comitê de Ética é compulsória em minha Instituição/Unidade Acadêmica.
46	Todas	M	Ética	Integridade	Código de Ética	Denscombe	Que evidências existem de atenção às questões éticas?	Uso de princípios de um Código de Ética	Uso princípios de ética na pesquisa descritos formalmente em um Código de Ética (de uma instituição de fomento, por exemplo), para orientar meus estudos.
47	Todas	M	Ética	Integridade	Respeito aos sujeitos humanos	Denscombe, National Institute of Health, Conama, Fapesp.	Que evidências existem de atenção às questões éticas?	Respeito aos subjeitos humanos formalmente	Entrego o Termo de Consentimento aos sujeitos humanos envolvidos em meus estudos, com detalhes sobre propósito do estudo, riscos e outros aspectos previstos na legislação.

N.	Fase	Dom.	Caract. chave Spencer et al.	Critério	Foco	Fonte (s)	Questão (Spencer et al)	Atributo/Relação	Proposição para a Delphi
48	Todas	M	Ética	Integridade	Respeito aos sujeitos humanos	Denscombe, National Institute of Health, CNS, Fapesp.	Que evidências existem de atenção às questões éticas?	Estratégia para mitigar possíveis danos	Me preocupo em escolher a estratégia de pesquisa que traga os menores danos para os sujeitos envolvidos nos meus estudos.
49	Todas	M	Ética	Integridade	Confidencialidade	Denscombe, National Institute of Health, CNS, Fapesp.	Que evidências existem de atenção às questões éticas?	Confidencialidade para os dados dos participantes	Minhas pesquisas fornecem procedimentos específicos para tratar os dados dos participantes de maneira confidencial.
50	Todas	M	Auditabilidade	Confiabilidade/ auditabilidade	Rigor metodológico	Denscombe, Spencer et al., Mays & Pope	Quão adequadamente documentado é o processo de pesquisa?	Registros das mudanças no desenho	Mantenho registros completos das alterações feitas no desenho das pesquisas que modificaram o alcance/amplitude destas.
51	Todas	M	Auditabilidade	Confiabilidade/ auditabilidade	Consistência	Denscombe, Spencer et al., Mays & Pope	Quão adequadamente documentado é o processo de pesquisa?	Registros das razões para as mudanças	Mantenho registros completos das razões que motivaram as alterações feitas no desenho das pesquisas.
52	Todas	M	Auditabilidade	Confiabilidade/ auditabilidade	Consistência	Denscombe, Spencer et al., Mays & Pope	Quão adequadamente documentado é o processo de pesquisa?	Salvaguarda das bases de dados para verificações	Mantenho as bases de dados dos meus estudos, com adequada descrição da sua composição, para futuras verificações.
53	Todas	M	Auditabilidade	Confiabilidade/ auditabilidade	Rigor metodológico	Denscombe, Spencer et al., Mays & Pope	Quão adequadamente documentado é o processo de pesquisa?	Guarda de documentos para reduzir riscos	Os termos de consentimento, assinados (ou com declaração explícita) individualmente por cada participante, são mantidos em segurança para preservar os sujeitos das pesquisas de possíveis riscos.

Nota: Apenas algumas fontes foram incluídas na matriz. Outras fontes que deram suporte estão exploradas na revisão de literatura.

Appendix C: Registration with the Ethical Committee for Human Beings (in Portuguese)

PLATAFORMA BRASIL
http://aplicacao.saude.gov.br/plataformabrasil/login.jsf

DADOS DO PROJETO DE PESQUISA

Título Público: Qualidade no processo de produção científica em Ciências Contábeis no Brasil
Pesquisador Responsável: JOSE RENATO SENA OLIVEIRA
Contato Público: JOSE RENATO SENA OLIVEIRA
Condições de saúde ou problemas estudados:
Descritores CID - Gerais:
Descritores CID - Específicos:
Descritores CID - da Intervenção:
Data de Aprovação Ética do CEP/CONEP: 06/12/2015

DADOS DA INSTITUIÇÃO PROPONENTE

Nome da Instituição: Universidade Estadual de Feira de Santana
Cidade: FEIRA DE SANTANA

DADOS DO COMITÊ DE ÉTICA EM PESQUISA

Comitê de Ética Responsável: 53 - Universidade Estadual de Feira de Santana - UEFS
Endereço: Avenida Transnordestina, s/n - Novo Horizonte, UEFS
Telefone: (75)3161-8067
E-mail: cep@uefs.br

CENTRO(S) PARTICIPANTE(S) DO PROJETO DE PESQUISA

CENTRO(S) COPARTICIPANTE(S) DO PROJETO DE PESQUISA

Appendix D: List of Selected Graduate Programs (in Portuguese)

LISTA DE PROGRAMAS DE PÓS-GRADUAÇÃO SELECIONADOS

GRANDE ÁREA: CIÊNCIAS SOCIAIS APLICADAS
ÁREA: ADMINISTRAÇÃO

N.	PROGRAMA	IES	UF	M	D	F	WEBSITE	E-MAIL
1	Administração e Controladoria	UFC	CE	4	4	-	http://www.ppac.ufc.br/index.php/pt-br/	maac@ufc.br
2	Ciências Contábeis	UNB	DF	4	4	-	http://ppgcont.unb.br	ppgcont@unb.br
3	Ciências Contábeis	UFES	ES	3	-	-	http://www.cienciascontabeis.ufes.br/	pos.cienciascontabeis@ufes.br
4	Ciências Contábeis	FUCAPE	ES	4	4	-	http://fucape.br/cursos/mestradoacademicocontabeis/	arilton@fucape.br
5	Ciências Contábeis	UFMG	MG	4	-	-	http://cepcon.face.ufmg.br/	cepcon@face.ufmg.br
6	Ciências Contábeis	UFU	MG	3	-	-	http://www.ppgcc.facic.ufu.br/	lailamelo@facic.ufu.br
7	Ciências Contábeis	UFPB/J.P.	PB	4	4	-	http://www.ccsa.ufpb.br/ppgcc/	ppgcc@ccsa.ufpb.br
8	Ciências Contábeis	UFPE	PE	4	-	-	https://www.ufpe.br/ppgcontabeis/	mestrado.contabeis@ufpe.br
9	Ciências Contábeis	UEM	PR	3	-	-	http://www.pco.uem.br/	sec-pco@uem.br
10	Ciências Contábeis	UFRJ	RJ	5	4	-	http://ppgcc.ufrj.br/	ppgcc@facc.ufrj.br
11	Ciências Contábeis	UERJ	RJ	3	-	-	http://ppgcc.faf.uerj.br/	ppgcc@uerj.br
12	Ciências Contábeis	UFRN	RN	3	-	-	https://sigaa.ufrn.br/sigaa/public/programa/equipe.jsf?lc=pt_BR&id=9066	ppgcc@ccsa.ufrn.br
13	Ciências Contábeis	UNISINOS	RS	5	4	-	http://www.unisinos.br/mestrado-e-doutorado/ciencias-contabeis	ppgeconomicas@unisinos.br
14	Ciências Contábeis	FURB	SC	4	4	-	http://www.furb.br/web/1854/cursos/programa-pos-graduacao/ciencias-contabeis/apresentacao	ppgcc-mestrado@furb.br
15	Ciências Contábeis	UNIFECAP	SP	4	-	-	http://www.fecap.br/portalinstitucional/mestrado/	secretaria.mestrado@fecap.br
16	Ciências Contábeis e Atuariais	PUC/SP	SP	3	-	-	http://www.pucsp.br/pos-graduacao/mestrado-e-doutorado/ciencias-contabeis-e-atuariais	poscont@pucsp.br
17	Contabilidade	UFBA	BA	3	-	-	http://www.contabeis.ufba.br/	ppgcont@ufba.br
18	Contabilidade	UFPR	PR	4	4	-	http://www.ppgcontabilidade.ufpr.br/	ppgccont@ufpr.br
19	Contabilidade	UNIOESTE	PR	3	-	-	http://www.unioeste.br/pos/cienciascontabeis/	cascavel.ppgc@unioeste.br
20	Contabilidade	UFSC	SC	4	4	-	http://www.ppgc.ufsc.br/	ppgc@contato.ufsc.br
21	Controladoria	UFRPE	PE	3	-	-	http://www.prppg.ufrpe.br/	cppg@prppg.ufrpe.br
22	Controladoria E Contabilidade	USP	SP	6	6	-	http://prpg.usp.br/ppgcc	ppgcc@usp.br
23	Controladoria E Contabilidade	USP/RP	SP	4	4	-	http://www.fearp.usp.br/pt-br/ppgcc.html	rcc@fearp.usp.br

Fonte: SNPG.
Data Atualização: **20/03/2015**
Data da Coleta: 04/09/2015.

Appendix E: Invitation Letter to the Researchers (in Portuguese)

São Paulo, ___ de outubro de 2015.

Prezado (a) Pesquisador (a),

A qualidade da produção científica deve ser objeto constante de reflexões da comunidade acadêmica, na busca de ampliar as fronteiras do conhecimento e oferecer soluções aos problemas sociais, com rigor e relevância. É um elemento importante para os PPG, dado o seu impacto sobre o que é produzido e sua influência nas avaliações periódicas dos mesmos.

Neste sentido, pedimos a sua participação em uma pesquisa de doutorado "Qualidade no processo de produção científica em Ciências Contábeis no Brasil", para uma tese a ser defendida no PPGCC/FEA/USP. O estudo utilizará a técnica Delphi e a coleta será feita por questionários eletrônicos com questões fechadas, com uso de login e senha. Serão apenas 3 rodadas e a previsão é que sejam necessários menos de 20 minutos para concluir cada rodada. Sua função será a de compor o painel de especialistas da Delphi para formação de consensos sobre aspectos da qualidade, a partir de suas práticas nos diferentes estágios do processo de pesquisa.

Ao final do estudo, teremos prazer de encaminhar um sumário com os principais achados e conclusões. Caso concorde, pedimos que responda a este e-mail até o dia 17 de outubro de 2015. Em seguida, encaminharemos uma mensagem com o endereço eletrônico e instruções para acesso e participação na primeira rodada.

Quaisquer dúvidas, sinta-se à vontade para nos contatar. Desde já agradecemos.

Cordialmente,

José Renato Sena Oliveira
Doutorando em Controladoria e Contabilidade – PPGCC/FEA/USP
jrsena@usp.br / jrsenna@uefs.br

Prof. Dr. Gilberto de Andrade Martins
Professor Orientador – PPGCC/FEA/USP
Professor Titular – EAC/FEA/USP
martins@usp.br – (11) 3091-5820, ramal 124.

Appendix F: Invitation Letter to the Graduate Coordinators (in Portuguese)

São Paulo, ___ de outubro de 2015

Prezado (a) Coordenador (a),

A qualidade da produção científica deve ser objeto constante de reflexões da comunidade acadêmica, na busca de ampliar as fronteiras do conhecimento e oferecer soluções aos problemas sociais, com rigor e relevância. É um elemento importante para os PPG, dado o seu impacto sobre o que é produzido e sua influência nas avaliações periódicas dos mesmos.

Neste sentido, pedimos a sua colaboração no sentido de incentivar os (as) docentes do seu Programa a participarem, voluntariamente, da pesquisa "Qualidade no processo de produção científica em Ciências Contábeis no Brasil", para a tese do doutorando José Renato Sena Oliveira, a ser defendida no PPGCC/FEA/USP, sob a orientação do Prof. Dr. Gilberto de Andrade Martins. Os (as) docentes serão convidados (as) a comporem o painel de especialistas da Delphi sobre aspectos da prática cotidiana, relativos à qualidade, em diferentes estágios do processo de pesquisa.

Quaisquer dúvidas, sinta-se à vontade para nos contatar ou para entrar em contato com o doutorando pelos e-mails jrsena@usp.br ou jrsenna@uefs.br.

Agradecemos sua colaboração.

Cordialmente,

Prof. Dr. Andson Braga de Aguiar
Coordenador do PPGCC/FEA/USP
ppgcc@usp.br – (11) 3091-5920.

Prof. Dr. Gilberto de Andrade Martins
Professor Orientador – PPGCC/FEA/USP
martins@usp.br – (11) 3091-5820, ramal 124.

Appendix G: Data Collection Instrument – First Round (in Portuguese)

INSTRUMENTO DE COLETA DE DADOS (DELPHI: 1ª RODADA)

INSTRUÇÃO:
Para as proposições a seguir, atribua uma nota em uma escala de 1 a 10, de acordo com o nível de concordância quanto à aderência da proposição à sua prática de pesquisa.

PROPOSIÇÃO	NOTA									
	1	2	3	4	5	6	7	8	9	10
Desenho da Pesquisa										
Meus estudos apresentam razões/justificativas para a escolha da estratégia de pesquisa.										
Sou capaz de reconhecer os pressupostos da visão geral (paradigma ou lente teórica) que guiam a construção dos meus estudos.										
Levo em conta um paradigma e seus pressupostos quando escolho a estratégia para construir o desenho das minhas pesquisas.										
Meus estudos apresentam, precisamente, o objetivo/problema que se pretende alcançar ou resolver.										
Levo em conta um paradigma e seus pressupostos quando escolho a teoria para suportar meus estudos (plataforma teórica).										
Apresento argumentos para a escolha das técnicas e procedimentos para coleta e análise dos dados das minhas pesquisas.										
A estratégia de pesquisa é escolhida considerando o acesso às fontes de dados para realizar os meus estudos.										
Escolho uma estratégia de pesquisa baseado no tempo necessário para coletar os dados e apresentar os resultados.										
Levo em conta o provável público-alvo dos meus estudos quando escolho a estratégia da pesquisa.										
Sujeitos da Pesquisa										
Apresento, explicitamente, os critérios usados para desenhar/selecionar os sujeitos da pesquisa baseada na população de interesse.										
Meus estudos descrevem, formalmente, os métodos e os casos de inclusão/exclusão que têm impacto sobre a representatividade dos sujeitos da pesquisa.										
Coleta de Dados										
Normalmente, tomo nota dos passos e ações realizadas durante a coleta de dados.										
Normalmente, tomo nota de detalhes das ocorrências divergentes do desenho da pesquisa que aparecem durante a coleta de dados.										
Quando meus estudos envolvem seres humanos, estes são convidados a participarem voluntariamente.										
Todos os participantes dos meus estudos que envolvam seres humanos assinam (ou declaram expressamente) um Termo de Consentimento Livre e Esclarecido.										
Análise de Dados										
Meus estudos apresentam uma descrição detalhada da natureza e forma dos dados originais coletados.										
Ferramentas e procedimentos para análise de dados são, cuidadosamente, apresentados.										
Explicito os detalhes do contexto em que os dados foram coletados e o que pode ter/tem impacto na análise de dados.										

PROPOSIÇÃO	NOTA									
	1	2	3	4	5	6	7	8	9	10
Apresento outras visões/perspectivas possíveis para que o leitor possa entender melhor o contexto da análise dos meus dados.										
Discuto a significância dos dados para alcançar os objetivos em meus estudos.										
Discuto os links, implícitos e explícitos, entre os dados/achados e o propósito nos meus estudos.										
Achados										
Meus estudos apresentam ao público-alvo, de maneira clara, como cheguei às conclusões a partir dos achados reportados.										
As conclusões dos meus estudos apresentam, de maneira explícita, a correspondência entre achados e evidências/dados.										
Uso achados empíricos de estudos anteriores para comparar com os meus achados (quando aplicável).										
Baseados nos achados, meus estudos sugerem novas áreas para investigação.										
Os achados são apresentados de uma maneira que cria perspectivas (insights) de pensar o conhecimento na área.										
Uso achados empíricos de estudos anteriores para suportar minhas proposições e/ou hipóteses (quando aplicável).										
Checo se meus achados estão, claramente, ligados ao propósito dos meus estudos.										
Descrevo em detalhes o contexto no qual os estudos foram conduzidos, a fim de permitir replicações.										
Meus estudos discutem, precisamente, a natureza e a origem de possíveis divergências nos dados que poderiam impactar as conclusões.										
Relatório										
Meus relatórios apresentam, claramente, a discussão sobre o impacto dos achados para a expansão do conhecimento na área.										
Meus relatórios contêm, explicitamente, as limitações dos achados das minhas pesquisas.										
Apresento discussão acerca das razões que geraram as limitações dos meus estudos.										
Meus relatórios de pesquisa evidenciam, formalmente, as fronteiras (alcance) dos achados dos meus estudos.										
As revisões de literatura que faço descrevem os conceitos-chave inerentes aos assuntos estudados.										
Uso explicação baseada na teoria para justificar as proposições dos meus estudos.										
Minhas conclusões mostram, explicitamente, a ligação com os objetivos dos estudos.										
A possibilidade de generalização (quando aplicável) é explicitamente evidenciada, com base no escopo dos meus estudos.										
Submeto minhas versões preliminares (drafts) para a apreciação de colegas antes da submissão final a um periódico/evento.										
Reflexividade e Neutralidade										
Discuto em meus estudos, de forma explícita, como erros ou vieses podem ter surgido durante o processo de pesquisa e como lidei com eles.										
Discuto, de forma explícita, possíveis impactos decorrentes da minha participação (e da minha equipe) nas fases do processo de pesquisa (quando aplicável).										

PROPOSIÇÃO	NOTA									
	1	2	3	4	5	6	7	8	9	10
Aspectos Éticos										
Minha Instituição/Unidade Acadêmica tem um Comitê de Ética em Pesquisa para submissão e aprovação dos projetos de pesquisa.										
Minha Instituição/Unidade Acadêmica tem regras claras para submissão dos projetos de pesquisa ao Comitê de Ética em Pesquisa.										
Meus estudos que envolvem seres humanos são, formalmente, submetidos ao Comitê de Ética em Pesquisa.										
A submissão dos projetos de pesquisa que envolvem seres humanos ao Comitê de Ética é compulsória em minha Instituição/Unidade Acadêmica.										
Uso princípios de ética na pesquisa descritos formalmente em um Código de Ética (de uma instituição de fomento, por exemplo), para orientar meus estudos.										
Entrego o Termo de Consentimento aos sujeitos humanos envolvidos em meus estudos, com detalhes sobre propósito do estudo, riscos e outros aspectos previstos na legislação.										
Me preocupo em escolher a estratégia de pesquisa que traga os menores danos para os sujeitos envolvidos nos meus estudos.										
Minhas pesquisas fornecem procedimentos específicos para tratar os dados dos participantes de maneira confidencial.										
Auditabilidade										
Mantenho registros completos das alterações feitas no desenho das pesquisas que modificaram o alcance/amplitude destas.										
Mantenho registros completos das razões que motivaram as alterações feitas no desenho das pesquisas.										
Mantenho as bases de dados dos meus estudos, com adequada descrição da sua composição, para futuras verificações.										
Os termos de consentimento, assinados (ou com declaração explícita) individualmente por cada participante, são mantidos em segurança para preservar os sujeitos das pesquisas de possíveis riscos.										

Aqui termina a parte sobre a qualidade no processo de pesquisa.

Por fim, deixe-nos saber um pouco sobre você e o PPG onde atua.

12. Área de conhecimento do seu Doutorado:

- ☐ Administração/Finanças
- ☐ Ciências Sociais
- ☐ Contabilidade/Controladoria
- ☐ Economia
- ☐ Educação
- ☐ Engenharia de Produção
- ☐ Matemática/Estatística
- ☐ Outra (qual?)_____

13. Sobre o seu Doutorado, escolha a opção que melhor o caracteriza:

- ☐ Fiz o Doutorado no mesmo PPG onde atuo.
- ☐ Fiz o Doutorado na mesma Instituição, mas em um PPG diferente do que atuo.
- ☐ Fiz o Doutorado em outra Instituição nacional.
- ☐ Fiz o Doutorado em uma Instituição estrangeira.

14. Fez Estágio Doutoral (sanduíche) durante o Doutorado?

- [ ] Não fiz.
- [ ] Sim, fiz Estágio Doutoral em outra Instituição nacional.
- [ ] Sim, fiz Estágio Doutoral em uma Instituição estrangeira.

15. Já fez Pós-doutorado?

- [ ] Não fiz.
- [ ] Sim, fiz Pós-doutorado em uma Instituição nacional.
- [ ] Sim, fiz Pós-doutorado em uma Instituição estrangeira.

16. Tempo de atuação como membro do corpo docente do PPG (em anos):

[ ]

17. Seu credenciamento para orientação de discentes do PPG é:

- [ ] Somente para Mestrado
- [ ] Somente para Doutorado
- [ ] Para Mestrado e Doutorado

18. Possui projeto de pesquisa financiado por Instituição de Fomento?

- [ ] Sim
- [ ] Não

19. Possui bolsa de produtividade em pesquisa fornecida por Instituição de Fomento?

- [ ] Sim
- [ ] Não

20. Atua ou já atuou como Editor (a) de periódico classificado nos estratos do Qualis/CAPES (exceto Qualis C)?

- [ ] Sim
- [ ] Não

21. É ou já foi membro de Conselho Editorial de periódico classificado nos estratos do Qualis/CAPES (exceto Qualis C)?

- [ ] Sim
- [ ] Não

22. É ou já foi Avaliador (a) de periódico classificado nos estratos do Qualis/CAPES (exceto Qualis C)?

- [ ] Sim
- [ ] Não

23. Considerando a última Avaliação Trienal, quantos artigos você publicou em periódico classificado nos estratos superiores (A1, A2 ou B1) do Qualis/CAPES?

- [ ] 0
- [ ] 1
- [ ] 2
- [ ] 3
- [ ] 4
- [ ] 5
- [ ] Mais de 5
- [ ]

24. Característica do PPG onde atua:

- [ ] Oferece apenas Mestrado
- [ ] Oferece Mestrado e Doutorado

25. Nota atual do PPG onde atua na avaliação da CAPES:

☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7

26. Use este espaço para deixar seu comentário, crítica ou sugestão (opcional)

Esta primeira rodada da Delphi termina aqui.
Em breve você receberá um novo e-mail com um resumo das respostas e o link para acesso ao questionário da próxima rodada.
Muito obrigado!

Appendix H: Informed Consent (in Portuguese)

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO (TCLE)

Prezado (a) Pesquisador (a),

Você está sendo convidado (a) como voluntário (a) a participar da pesquisa intitulada **"Qualidade no processo de produção científica em Ciências Contábeis no Brasil"**, como parte do projeto de tese cujo pesquisador responsável é JOSÉ RENATO SENA OLIVEIRA, discente do Doutorado Interinstitucional (parceria entre a Universidade de São Paulo e a Universidade Estadual de Feira de Santana e associadas) do Programa de Pós-Graduação em Controladoria e Contabilidade da Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo (PPGCC/FEA/USP). O pesquisador é orientado pelo Prof. Dr. Gilberto de Andrade Martins.

O objetivo do estudo é levantar e discorrer sobre atributos de qualidade de uma boa pesquisa no processo de construção da produção científica em Ciências Contábeis, em que serão levantados tais atributos nas práticas de investigação dos (das) respondentes e confrontados com aqueles identificados na literatura. O estudo é importante para compreender as práticas de produção científica na área, a fim de avaliar a qualidade da pesquisa a partir do seu processo de construção. Poderá contribuir para redirecionar ações, redefinir práticas e estratégias metodológicas, especialmente nos Programas de Pós-Graduação, bem como para identificar elementos que colaborem para maior aceitabilidade pela produção na comunidade científica.

A coleta de dados será feita eletronicamente com o uso da técnica Delphi modificada (com 3 rodadas para aplicação dos instrumentos de coleta), a fim de formar consenso entre os respondentes. Em cada rodada você responderá a um questionário objetivo, cujo acesso para a primeira rodada será pelo site http://pt.surveymonkey.com/r/QualiPPGCCFEAUSP, com senha e código de respondente personalizado que você receberá em seu e-mail. O código de respondente é individual e tem a finalidade de permitir a conexão entre as respostas de cada respondente nas diferentes rodadas. A fim de preservar a sua privacidade, o código não guarda qualquer associação que permita identificação dos (das) respondentes na base de dados.

Você deverá atribuir uma nota (1 a 10) sobre o seu nível de concordância a cada proposição relativa a uma determinada característica das diferentes fases do processo de pesquisa. Como a finalidade é a formação do consenso sobre as práticas, o questionário será modificado nas fases seguintes em função do resultado alcançado e de sugestões dos (das) experts na fase anterior. O questionário da primeira rodada da Delphi apresentará, também, algumas questões para caracterização do perfil dos (das) participantes. A previsão é que sejam necessários menos de 20 minutos para cada rodada. Você receberá por e-mail um resumo personalizado com a sua nota e com a mediana do grupo para cada proposição, após cada rodada.

Considerando que a coleta de dados ocorrerá por meio eletrônico, há um risco associado a uma possível identificação de senha e/ou código de respondente do (da) participante. Para minimizá-lo, recomenda-se a manutenção do e-mail sem acesso por terceiros, bem como evitar a impressão da mensagem eletrônica. Outro aspecto é um possível desconforto com alguma questão específica e, a fim de mitigá-lo, os dados não serão analisados individualmente, mas no agregado da formação do consenso dos (das) respondentes para que não haja exposição individual. Quando aos benefícios esperados à comunidade científica, estão explicitados nas contribuições esperadas apresentadas no parágrafo relativo ao objetivo e às justificativas.

Você será esclarecido (a) sobre os aspectos relativos a esta pesquisa sempre que desejar. Você é livre para recusar-se a participar, retirar seu consentimento ou interromper a participação a qualquer momento. A sua participação é voluntária e a recusa em participar não irá acarretar qualquer penalidade. Os pesquisadores tratarão a sua identidade com padrões profissionais de sigilo e os dados serão analisados de forma agregada, sem identificação individual de participantes. Os dados serão divulgados na tese e em futuras publicações de natureza científica, seguindo as diretrizes éticas da pesquisa e assegurando a privacidade dos (das) respondentes. Caberá ao pesquisador responsável a

preservação do sigilo e a guarda dos dados, os quais serão mantidos em meio eletrônico indefinidamente para uso em futuras publicações de natureza científica. Entretanto, os dados que possam representar risco à privacidade dos (das) respondentes serão destruídos após a defesa da tese.

A participação no estudo não acarretará custos para você e não será disponibilizada nenhuma compensação financeira. Qualquer dano que, porventura, o (a) participante vier a sofrer como resultado da participação dele (a) nesta pesquisa é de responsabilidade dos pesquisadores e das Instituições envolvidas, conforme previsto na Res. CNS n. 466/12 e será apurado nos termos da legislação em vigor. Ao final do estudo, teremos prazer de encaminhar a você, por e-mail, um sumário com os principais achados e conclusões.

Em caso de dúvidas você pode indagar ao pesquisador José Renato Sena Oliveira pelos meios de contato ou no Departamento de Ciências Sociais Aplicadas da UEFS, situado à Av. Transnordestina, S/N, Módulo 3, MA 34 – Novo Horizonte – Feira de Santana/BA. Alternativamente, indagar ao pesquisador e/ou seu orientador pelos meios de contato ou no Departamento de Contabilidade e Atuária da FEA/USP, situado à Av. Prof. Luciano Gualberto, 908 – Butantã – São Paulo/SP. Para dúvidas de natureza ética, você poderá contatar o Comitê de Ética em Pesquisa da UEFS (CEP/UEFS), pelo e-mail cep@uefs.br ou no endereço Av. Transnordestina, S/N, Módulo 1, MA 17, Bairro Novo Horizonte – Feira de Santana/BA.

Cordialmente**,**

José Renato Sena Oliveira
Doutorando em Controladoria e Contabilidade – PPGCC/FEA/USP
jrsena@usp.br / jrsenna@uefs.br

Prof. Dr. Gilberto de Andrade Martins
Professor Orientador – PPGCC/FEA/USP
Professor Titular – EAC/FEA/USP
martins@usp.br – (11) 3091-5820, ramal 124

Por favor, insira no campo abaixo o código de respondente que você recebeu por e-mail:

☐ Concordo com as condições para participação conforme descritas acima.

ATENÇÃO: guarde uma cópia deste TCLE após o preenchimento destes campos.

Appendix I: Invitation Letter for Delphi First Round (in Portuguese)

Prezado (a) Pesquisador (a),

Inicialmente, gostaríamos de agradecer por ter aceito o convite para participar voluntariamente do painel de especialistas da Delphi no estudo "Qualidade no processo de produção científica em Ciências Contábeis no Brasil", como parte do projeto de tese de José Renato Sena Oliveira no PPGCC/FEA/USP, sob a orientação do Prof. Dr. Gilberto de Andrade Martins.

A partir da presente data estamos iniciando a coleta de dados da primeira rodada da Delphi. A fim de preservar a privacidade dos (das) respondentes, na base de dados não haverá qualquer identificação pessoal. Cada usuário (a) deverá utilizar senha e Código de Respondente individual, que foi criado de modo a não apresentar qualquer elemento de identificação e, ao mesmo tempo, permitir o vínculo das respostas nas diferentes rodadas. Pedimos que preserve estes dados, pois são importantes para que possamos fazer tal vínculo.

Seus dados personalizados são:

SENHA: xxxxxxxxxx
CÓDIGO DE RESPONDENTE: delphixxx

Para acessar o instrumento visite o endereço http://pt.surveymonkey.com/r/QualiPPGCCFEAUSP. A tela de abertura já requererá esta senha.

Na tela seguinte você encontrará o Termo de Consentimento Livre e Esclarecido (TCLE), com informações adicionais sobre o estudo. Antes de passar para a página seguinte, pedimos que insira este Código de Respondente no campo ao final do TCLE e sugerimos que gere e guarde uma cópia do referido Termo.

O instrumento ficará disponível até o dia ___ de _____. Após o término desta etapa encaminharemos um novo e-mail com as instruções para a próxima rodada.
Qualquer dúvida, sinta-se à vontade para nos contatar.

Agradecemos por sua disponibilidade.

Cordialmente,

José Renato Sena Oliveira
Doutorando em Controladoria e Contabilidade – PPGCC/FEA/USP
jrsena@usp.br / jrsenna@uefs.br

Prof. Dr. Gilberto de Andrade Martins
Professor Orientador – PPGCC/FEA/USP
Professor Titular – EAC/FEA/USP
martins@usp.br – (11) 3091-5820, ramal 124

Appendix J: Customized Report for Experts (in Portuguese)

USUÁRIO: delphiXXX **RODADA:** 1

ETAPA	PROPOSIÇÃO	Sua Nota	Mediana
Desenho	Meus estudos apresentam razões/justificativas para a escolha da estratégia…	5	9.00
	Sou capaz de reconhecer os pressupostos da visão geral (paradigma…	5	9.00
	Levo em conta um paradigma e seus pressupostos quando escolho a estratégia…	5	9.00
	Meus estudos apresentam, precisamente, o objetivo/problema…	5	10.00
	Levo em conta um paradigma e seus pressupostos quando escolho a teoria…	5	9.00
	Apresento argumentos para a escolha das técnicas e procedimentos para coleta…	5	9.00
	A estratégia de pesquisa é escolhida considerando o acesso às fontes de dados…	5	9.00
	Escolho uma estratégia de pesquisa baseado no tempo necessário para coletar…	5	8.00
	Levo em conta o provável público-alvo dos meus estudos quando escolho…	5	7.00
Sujeitos da Pesquisa	Apresento, explicitamente, os critérios usados para desenhar/selecionar…	5	9.00
	Meus estudos descrevem, formalmente, os métodos e os casos…	5	9.00
Coleta de Dados	Normalmente, tomo nota dos passos e ações realizadas durante a coleta…	5	9.00
	Normalmente, tomo nota de detalhes das ocorrências divergentes do desenho…	5	9.00
	Quando meus estudos envolvem seres humanos, estes são convidados…	5	10.00
	Todos os participantes dos meus estudos que envolvam seres humanos assinam	5	8.00
Análise dos Dados	Meus estudos apresentam uma descrição detalhada da natureza e forma…	5	9.00
	Ferramentas e procedimentos para análise de dados são, cuidadosamente,…	5	9.00
	Explicito os detalhes do contexto em que os dados foram coletados e o que pode…	5	8.00
	Apresento outras visões/perspectivas possíveis para que o leitor…	5	7.00
	Discuto a significância dos dados para alcançar os objetivos em meus estudos.	5	9.00
	Discuto os links, implícitos e explícitos, entre os dados/achados e o propósito…	5	9.00
Achados	Meus estudos apresentam ao público-alvo, de maneira clara, como cheguei…	5	9.00
	As conclusões dos meus estudos apresentam, de maneira explícita, a…	5	9.00
	Uso achados empíricos de estudos anteriores para comparar com os meus…	5	10.00
	Baseados nos achados, meus estudos sugerem novas áreas para investigação.	5	9.00
	Os achados são apresentados de uma maneira que cria perspectivas (insights) …	5	8.00
	Uso achados empíricos de estudos anteriores para suportar minhas proposições…	5	9.00
	Checo se meus achados estão, claramente, ligados ao propósito…	5	9.00
	Descrevo em detalhes o contexto no qual os estudos foram conduzidos, …	5	9.00
	Meus estudos discutem, precisamente, a natureza e a origem de possíveis...	5	8.00
Relatório	Meus relatórios apresentam, claramente, a discussão sobre o impacto…	5	8.00
	Meus relatórios contêm, explicitamente, as limitações dos achados…	5	9.00
	Apresento discussão acerca das razões que geraram as limitações…	5	8.00
	Meus relatórios de pesquisa evidenciam, formalmente, as fronteiras (alcance)…	5	8.00
	As revisões de literatura que faço descrevem os conceitos-chave …	5	9.00
	Eu uso explicação baseada na teoria para justificar as proposições…	5	9.00
	Minhas conclusões mostram, explicitamente, a ligação com os objetivos …	5	10.00
	A possibilidade de generalização (quando aplicável) é explicitamente…	5	9.00
	Submeto minhas versões preliminares (drafts) para a apreciação de colegas …	5	7.00
Reflexivid. Neutralid.	Discuto em meus estudos, de forma explícita, como erros ou viéses podem…	5	7.00
	Discuto, de forma explícita, possíveis impactos decorrentes da minha particip…	5	7.00
Aspectos Éticos	Minha Instituição/Unidade Acadêmica tem um Comitê de Ética para submissão...	5	10.00
	Minha Instituição/Unidade Acadêmica tem regras claras para submissão…	5	9.00
	Meus estudos que envolvem seres humanos são, formalmente, submetidos…	5	6.00
	A submissão dos projetos de pesquisa que envolvem seres humanos ao Comitê…	5	6.00
	Uso princípios de ética na pesquisa descritos formalmente em um Código de…	5	9.00
	Entrego o Termo de Consentimento aos sujeitos humanos envolvidos…	5	8.00
	Me preocupo em escolher a estratégia de pesquisa que traga os menores danos…	5	9.00
	Minhas pesquisas fornecem procedimentos específicos para tratar os dados…	5	10.00
Auditabil.	Mantenho registros completos das alterações feitas no desenho…	5	8.00
	Mantenho registros completos das razões que motivaram as alterações…	5	8.00
	Mantenho as bases de dados dos meus estudos, com adequada descrição…	5	9.00
	Os termos de consentimento, assinados individualmente por cada participante…	5	8.00

Appendix K: Invitation Letter – Second Round (in Portuguese)

Prezado (a) Pesquisador (a),

Mais uma vez gostaríamos de agradecer por sua valiosa e voluntária contribuição como especialista no painel da Delphi no estudo "Qualidade no processo de produção científica em Ciências Contábeis no Brasil", como parte do projeto de tese de José Renato Sena Oliveira no PPGCC/FEA/USP, sob a orientação do Prof. Dr. Gilberto de Andrade Martins.

A partir de hoje iniciamos a coleta de dados da segunda rodada da Delphi. Fizemos alguns ajustes no novo instrumento:
- Por sugestão de alguns (mas) participantes, incluímos a alternativa "Nunca fiz/não faço pesquisas envolvendo seres humanos" aos grupos Coleta de Dados, Aspectos Éticos e Auditabilidade, a fim de evitar possíveis vieses na análise de algumas proposições. A sugestão também será levada em conta na análise dos dados já coletados na primeira rodada;
- Foi incluído o conceito de "pesquisa envolvendo seres humanos" previsto na Resolução no 466/12 do Conselho Nacional de Saúde;
- As proposições que já atenderam aos critérios e parâmetros de consenso formado que estamos utilizando no estudo não fazem mais parte desta segunda rodada. Como efeito, o novo instrumento de coleta é significativamente menor.

Seus dados personalizados são:
SENHA: xxxxxxxxxx
CÓDIGO DE RESPONDENTE: delphixxx

Para acessar o novo instrumento visite o endereço: http://pt.surveymonkey.com/r/DelphiQuali2. As orientações são idênticas às do questionário anterior.

A fim de facilitar suas respostas nesta etapa, segue anexo um sumário personalizado com suas notas e as medianas das notas atribuídas pelo grupo de especialistas na primeira rodada, mas apenas com as proposições que farão parte desta segunda rodada. Para cada item você poderá optar por manter a mesma nota ou escolher outra alternativa disponível como resposta.

O novo instrumento ficará disponível até o dia __ de _____. Após o término desta etapa encaminharemos um novo e-mail com as instruções para a rodada final, caso seja necessária.
Qualquer dúvida, sinta-se à vontade para nos contatar.

Agradecemos por sua disponibilidade.

Cordialmente,

José Renato Sena Oliveira
Doutorando em Controladoria e Contabilidade – PPGCC/FEA/USP
jrsena@usp.br / jrsenna@uefs.br

Prof. Dr. Gilberto de Andrade Martins
Professor Orientador – PPGCC/FEA/USP
Professor Titular – EAC/FEA/USP
martins@usp.br – (11) 3091-5820, ramal 124

Appendix L: Data Collection Instrument – Second Round (in Portuguese)

INSTRUMENTO DE COLETA DE DADOS (DELPHI: 2ª RODADA)

INSTRUÇÃO:
Para as proposições a seguir, atribua uma nota em uma escala de 1 a 10, de acordo com o nível de concordância quanto à aderência da proposição à sua prática de pesquisa.

PROPOSIÇÃO	NOTA									
	1	2	3	4	5	6	7	8	9	10
Desenho da Pesquisa										
Levo em conta um paradigma e seus pressupostos quando escolho a estratégia para construir o desenho das minhas pesquisas.										
Escolho uma estratégia de pesquisa baseado no tempo necessário para coletar os dados e apresentar os resultados.										
Levo em conta o provável público-alvo dos meus estudos quando escolho a estratégia da pesquisa.										
Sujeitos da Pesquisa										
Meus estudos descrevem, formalmente, os métodos e os casos de inclusão/exclusão que têm impacto sobre a representatividade dos sujeitos da pesquisa.										

PROPOSIÇÃO	NOTA										*
	1	2	3	4	5	6	7	8	9	10	
Coleta de Dados											
Normalmente, tomo nota de detalhes das ocorrências divergentes do desenho da pesquisa que aparecem durante a coleta de dados.											
Quando meus estudos envolvem seres humanos, estes são convidados e participarem voluntariamente**.											
Todos os participantes dos meus estudos que envolvam seres humanos assinam (ou declaram expressamente) um Termo de Consentimento Livre e Esclarecido.											

* Nunca fiz/não faço pesquisas envolvendo seres humanos.
** De acordo com a Resolução 466/2012 do Conselho Nacional de Saúde, pesquisa envolvendo seres humanos é a "pesquisa que, individual ou coletivamente, tenha como participante o ser humano, em sua totalidade ou partes dele, e o envolva de forma direta ou indireta, incluindo o manejo de seus dados, informações ou materiais biológicos" (Res. CNS n. 466/2012, item II.14).

PROPOSIÇÃO	NOTA									
	1	2	3	4	5	6	7	8	9	10
Análise de Dados										
Apresento outras visões/perspectivas possíveis para que o leitor possa entender melhor o contexto da análise dos meus dados.										
Achados										
Os achados são apresentados de uma maneira que cria perspectivas (insights) de pensar o conhecimento na área.										
Meus estudos discutem, precisamente, a natureza e a origem de possíveis divergências nos dados que poderiam impactar as conclusões.										

PROPOSIÇÃO	NOTA									
	1	2	3	4	5	6	7	8	9	10
Relatório										
Meus relatórios apresentam, claramente, a discussão sobre o impacto dos achados para a expansão do conhecimento na área.										
Meus relatórios contém, explicitamente, as limitações dos achados das minhas pesquisas.										
Apresento discussão acerca das razões que geraram as limitações dos meus estudos.										
Meus relatórios de pesquisa evidenciam, formalmente, as fronteiras (alcance) dos achados dos meus estudos.										
Submeto minhas versões preliminares (drafts) para a apreciação de colegas antes da submissão final a um periódico/evento.										
Reflexividade e Neutralidade										
Discuto em meus estudos, de forma explícita, como erros ou vieses podem ter surgido durante o processo de pesquisa e como lidei com eles.										
Discuto, de forma explícita, possíveis impactos decorrentes da minha participação (e da minha equipe) nas fases do processo de pesquisa (quando aplicável).										

PROPOSIÇÃO	NOTA										*
	1	2	3	4	5	6	7	8	9	10	
Aspectos Éticos											
Minha Instituição/Unidade Acadêmica tem um Comitê de Ética em Pesquisa para submissão e aprovação dos projetos de pesquisa.											
Minha Instituição/Unidade Acadêmica tem regras claras para submissão dos projetos de pesquisa ao Comitê de Ética em Pesquisa.											
Meus estudos que envolvem seres humanos são, formalmente, submetidos ao Comitê de Ética em Pesquisa.											
A submissão dos projetos de pesquisa que envolvem seres humanos ao Comitê de Ética é compulsória em minha Instituição/Unidade Acadêmica.											
Uso princípios de ética na pesquisa descritos formalmente em um Código de Ética (de uma instituição de fomento, por exemplo), para orientar meus estudos.											
Entrego o Termo de Consentimento aos sujeitos humanos envolvidos em meus estudos, com detalhes sobre propósito do estudo, riscos e outros aspectos previstos na legislação.											
Me preocupo em escolher a estratégia de pesquisa que traga os menores danos para os sujeitos envolvidos nos meus estudos.											
Minhas pesquisas fornecem procedimentos específicos para tratar os dados dos participantes de maneira confidencial.											
Auditabilidade											
Mantenho registros completos das alterações feitas no desenho das pesquisas que modificaram o alcance/amplitude destas.											
Mantenho registros completos das razões que motivaram as alterações feitas no desenho das pesquisas.											
Os termos de consentimento, assinados (ou com declaração explícita) individualmente por cada participante, são mantidos em segurança para preservar os sujeitos das pesquisas de possíveis riscos.											

Use este espaço para deixar seu comentário, crítica ou sugestão (opcional)

Esta segunda rodada da Delphi termina aqui.
Em breve você receberá um novo e-mail com um resumo das respostas e o link para acesso ao questionário da rodada final, se necessário.
Muito obrigado!

Appendix M: Thank You Letter for Participating in the Delphi Panel (in Portuguese)

Prezado (a) Pesquisador (a),

Chegamos ao final da segunda e última rodada da Delphi no estudo "Qualidade no processo de produção científica em Ciências Contábeis no Brasil", como parte do projeto de tese de José Renato Sena Oliveira no PPGCC/FEA/USP, sob a orientação do Prof. Dr. Gilberto de Andrade Martins. Agradecemos por sua valiosa e voluntária contribuição como especialista no painel.

Para a composição do painel foram convidados (as) os (as) docentes dos 23 Programas de Pós-Graduação acadêmicos da área contábil reconhecidos/recomendados pela CAPES, com base nos dados do Portal da CAPES e da Plataforma Sucupira (set/2015). Ao final, contamos com a participação de docentes de 19 dos 23 PPG (10 mestrados e 9 mestrados/doutorados).

Após a segunda rodada os resultados indicaram uma saturação do desvio-padrão, razão pela qual não realizaremos a terceira rodada.

Em respeito ao nosso compromisso firmado no Termo de Consentimento, após a defesa da tese encaminharemos um sumário dos principais achados a todos (as) os (as) participantes do painel.

Também agradecemos pelos comentários e sugestões, os quais ajudaram no ajuste do instrumento para a segunda rodada, bem como ajudarão na análise dos dados.

Colocamo-nos à disposição.

Cordialmente,

José Renato Sena Oliveira
Doutorando em Controladoria e Contabilidade – PPGCC/FEA/USP
jrsena@usp.br / jrsenna@uefs.br

Prof. Dr. Gilberto de Andrade Martins
Professor Orientador – PPGCC/FEA/USP
Professor Titular – EAC/FEA/USP
martins@usp.br – (11) 3091-5820, ramal 124

Appendix N: Summary of Final Results

Design – final classification (ordered by proposition)

Propos.	Attribute/relationship	N	Round	Score	%	Mean	Median	Ranking	Level	Criterion
R1Des01	Useful strategy for purpose	41	1	369	90.00	9.00	9.00	2º	Strong	Suitability
R1Des02	Clear overview that guides the study	41	1	358	87.32	8.73	9.00	4º	Strong	Suitability
R2Des03	Overview x strategy	37	2	316	85.41	8.54	9.00	6º	Strong	Suitability
R1Des04	Goal/problem shown precisely	41	1	395	96.34	9.63	10.00	1º	Strong	Suitability
R1Des05	Overview x theory	41	1	349	85.12	8.51	9.00	7º	Strong	Suitability
R1Des06	Arguments for technique choices	41	1	366	89.27	8.93	9.00	3º	Strong	Suitability
R1Des07	Access to the data	41	1	351	85.61	8.56	9.00	5º	Strong	Feasibility
R2Des08	Time constraints	37	2	261	70.54	7.05	7.00	9º	Moderate	Feasibility
R2Des09	Research strategy x target audience	37	2	264	71.35	7.14	8.00	8º	Moderate	Feasibility

Research Subjects – final classification (ordered by proposition)

Propos.	Attribute/relationship	N	Round	Score	%	Mean	Median	Ranking	Level	Criterion
R1Sub01	Criteria to design/select subjects	41	1	352	85.85	8.59	9.00	1º	Strong	Internal validity
R2Sub02	Representativeness of subjects	37	2	287	77.57	7.76	8.00	2º	Strong	Internal validity

Data collection – final classification (ordered by proposition)

Propos.	Attribute/relationship	N	Round	Score	%	Mean	Median	Ranking	Level	Criterion
R1Col01	Notes for each research step	41	1	361	88.05	8.80	9.00	2º	Strong	Rigor
R2Col02	Notes for divergent events	28	2	237	84.64	8.46	9.65	3º	Strong	Rigor
R2Col03	Voluntary participation of subjects	23	2	222	96.52	9.65	8.15	1º	Strong	Integrity
R2Col04	Formal agreement from humans	26	2	212	81.54	8.15	7.51	4º	Strong	Integrity

Analysis – final classification (ordered by proposition)

Propos.	Attribute/relationship	N	Round	Score	%	Mean	Median	Ranking	Level	Criterion
R1Ana01	Description of nature & form of data	41	1	351	85.61	8.56	9.00	2º	Strong	Integrity
R1Ana02	Description of tools and procedures	41	1	358	87.32	8.73	9.00	1º	Strong	Rigor
R1Ana03	Context x impact on data analysis	41	1	341	83.17	8.32	8.00	5º	Strong	Reliability
R2Ana04	Other views to explain context	37	2	278	75.14	7.51	8.00	6º	Strong	Reliability
R1Ana05	Significance of data to reaching aims	41	1	351	85.61	8.56	9.00	2º	Strong	Relevance
R1Ana06	Implicit/explicit links - find. x aims	41	1	349	85.12	8.51	9.00	4º	Strong	Rigor

Findings – final classification (ordered by proposition)

Propos.	Attribute/relationship	N	Round	Score	%	Mean	Median	Ranking	Level	Criterion
R1Fin01	Path to achieve conclusions	41	1	359	87.56	8.76	9.00	6º	Strong	Internal validity
R1Fin02	Link between findings x evidence	41	1	363	88.54	8.85	9.00	4º	Strong	Internal validity
R1Fin03	Comparison results x other studies	41	1	373	90.98	9.10	10.00	1º	Strong	External validity
R1Fin04	New areas based on the findings	41	1	361	88.05	8.80	9.00	5º	Strong	Contribution
R2Fin05	Insights for thinking about the field	37	2	302	81.62	8.16	8.00	8º	Strong	Contribution
R1Fin06	Previous findings x hypothesis	41	1	369	90.00	9.00	9.00	3º	Strong	Relevance
R1Fin07	Check links of findings x purpose	41	1	373	90.98	9.10	9.00	1º	Strong	Internal validity
R1Fin08	Does the context allow replication	41	1	351	85.61	8.56	9.00	7º	Strong	Reliability
R2Fin09	Impact from the nature of divergences	37	2	289	78.11	7.81	8.00	9º	Strong	Internal validity

Reporting – final classification (ordered by proposition)

Propos.	Attribute/relationship	N	Round	Score	%	Mean	Median	Ranking	Level	Criterion
R2Rep01	Discussion of impact on knowledge	37	2	301	81.35	8.14	8.00	6º	Strong	Impact
R2Rep02	Disclosure of limitations	37	2	315	85.14	8.51	9.00	4º	Strong	Contribution
R2Rep03	Reasons for limitations	37	2	293	79.19	7.92	8.00	8º	Strong	Integrity
R2Rep04	Boundaries of the study	37	2	294	79.46	7.95	8.00	7º	Strong	Contribution
R1Rep05	Literature review x main concepts	41	1	373	90.98	9.10	9.00	2º	Strong	Internal validity
R1Rep06	Theory to support propositions	41	1	371	90.49	9.05	9.00	3º	Strong	Internal validity
R1Rep07	Conclusions x aim	41	1	377	91.95	9.20	10.00	1º	Strong	Internal validity
R1Rep08	Explicit possibility of generalization	41	1	345	84.15	8.41	9.00	5º	Strong	External validity
R2Rep09	Appreciation of drafts by colleagues	37	2	229	61.89	6.19	7.00	9º	Low	Internal validity

Reflexivity and Neutrality – final classification (ordered by proposition)

Propos.	Attribute/relationship	N	Round	Score	%	Mean	Median	Ranking	Level	Criterion
R1Neu01	How to address errors and biases	37	2	266	71.89	7.19	8.00	1º	Moderate	Integrity
R1Neu02	Impact of team participation	37	2	249	67.30	6.73	7.00	2º	Moderate	Integrity

Ethics – final classification (ordered by proposition)

Propos.	Attribute/relationship	N	Round	Score	%	Mean	Median	Ranking	Level	Criterion
R2Eth01	Registration of the study/CEP	26	2	207	79.62	7.96	10.00	3º	Strong	Rigor
R2Eth02	Rules for registering studies/CEP	25	2	176	70.40	7.04	9.00	6º	Moderate	Rigor
R2Eth03	Formal submission to the CEP	25	2	144	57.60	5.76	6.00	7º	Low	Rigor
R2Eth04	Compulsory submission to the CEP	25	2	142	56.80	5.68	7.00	8º	Low	Rigor
R2Eth05	Use of the principles of a code	26	2	187	71.92	7.19	9.00	5º	Moderate	Integrity
R2Eth06	Formal respect for human subjects	25	2	195	78.00	7.80	9.00	4º	Strong	Integrity
R2Eth07	Strategy to mitigate possible harm	26	2	212	81.54	8.15	9.00	2º	Strong	Integrity
R2Eth08	Confidentiality to participants' data	26	2	248	95.38	9.54	10.00	1º	Strong	Integrity

Auditability – final classification (ordered by proposition)

Propos.	Attribute/relationship	N	Round	Score	%	Mean	Median	Ranking	Level	Criterion
R2Aud01	Records of design changes	29	2	242	83.45	8.34	9.00	2º	Strong	Reliability
R2Aud02	Records of the reasons for changes	29	2	237	81.72	8.17	8.00	3º	Strong	Reliability
R1Aud03	Safeguard databases for checks	41	1	367	89.51	8.95	9.00	1º	Strong	Reliability
R2Aud04	Guard documents to reduce risks	25	2	194	77.60	7.76	9.00	4º	Strong	Reliability

Appendix O: Summary of Statements by Criterion, Key Feature, and Level of Agreement

Row Labels	Analysis	Auditability	Data	Design	Ethics	Findings	Neutrality	Reporting	Subjects	Total
Contribution	0	0	0	0	0	2	0	2	0	4
Strong	0	0	0	0	0	2	0	2	0	4
External validity	0	0	0	0	0	1	0	1	0	2
Strong	0	0	0	0	0	1	0	1	0	2
Feasibility	0	0	0	3	0	0	0	0	0	3
Moderate	0	0	0	2	0	0	0	0	0	2
Strong	0	0	0	1	0	0	0	0	0	1
Impact	0	0	0	0	0	0	0	1	0	1
Strong	0	0	0	0	0	0	0	1	0	1
Integrity	1	0	2	0	4	0	2	1	0	10
Moderate	0	0	0	0	1	0	2	0	0	3
Strong	1	0	1	0	3	0	0	0	0	5
Unformed	0	0	1	0	0	0	0	1	0	2
Internal validity	0	0	0	0	0	4	0	4	2	10
Low	0	0	0	0	0	0	0	1	0	1
Strong	0	0	0	0	0	3	0	3	2	8
Unformed	0	0	0	0	0	1	0	0	0	1
Relevance	1	0	0	0	0	1	0	0	0	2
Strong	1	0	0	0	0	1	0	0	0	2
Reliability	2	4	0	0	0	1	0	0	0	7
Strong	1	3	0	0	0	1	0	0	0	5
Unformed	1	1	0	0	0	0	0	0	0	2
Rigor	2	0	2	0	4	0	0	0	0	8
Low	0	0	0	0	2	0	0	0	0	2
Moderate	0	0	0	0	1	0	0	0	0	1
Strong	2	0	2	0	1	0	0	0	0	5
Suitability	0	0	0	6	0	0	0	0	0	6
Strong	0	0	0	6	0	0	0	0	0	6
Grand Total	**6**	**4**	**4**	**9**	**8**	**9**	**2**	**9**	**2**	**53**