Examining the Performance of Multidisciplinary Greenhouse Gas Assurance Engagement Teams

Author:
Ekasingh, Erboon

Publication Date:
2015

DOI:
https://doi.org/10.26190/unsworks/17257

Examining the Performance of Multidisciplinary Greenhouse Gas Assurance Engagement Teams

Erboon Ekasingh

A Thesis in fulfilment of the requirements of
the degree of Doctor of Philosophy

School of Accounting
UNSW Business School

January 2015

PLEASE TYPE

THE UNIVERSITY OF NEW SOUTH WALES
Thesis/Dissertation Sheet

Surname or Family name: **Ekasingh**

First name: **Erboon** Other name/s:

Abbreviation for degree as given in the University calendar: **PhD**

School: **School of Accounting** Faculty: **UNSW Business School**

Title: Examining the Performance of Multidisciplinary Greenhouse Gas Assurance Engagement Teams.

Abstract 350 words maximum: (PLEASE TYPE)

The growing pressure for public disclosure on Greenhouse Gas (GHG) emissions has led to a significant demand for assurance on GHG statements. Given the complexity of GHG assurance, the international GHG assurance standard requires multidisciplinary GHG assurance teams (MDGHGTs) comprising practitioners from accounting and non-accounting background to discuss and assesses the risks of material misstatements in the planning stage of the engagement. However, prior research highlights the difficulties, including impaired effectiveness, associated with integrating practitioners from a range of disciplines, and as such suggests the importance of identifying mechanisms to ensure the effectiveness of MDGHGTs is not impaired. This dissertation addresses the effectiveness of MDGHGTs through two studies.

Study One utilises a retrospective recall approach to explore the factors that GHG assurers perceive could affect the effectiveness of MDGHGTs. This study finds that team processes are crucial to the success of MDGHGTs. In particular, having sufficient elaboration on different perspectives significantly increases the perceived effectiveness of MDGHGTs. This study also finds that the perceived sufficiency of elaboration increases when MDGHGTs become more educationally diverse, while the perceived sufficiency of discussion time decreases when the MDGHGTs become larger.

Study Two focuses on MDGHGT processes through a controlled experiment examining the effect of three team formats (nominal, interacting and review teams), on MDGHGT risk assessment performance. This study finds that accountant and non-accountant practitioners differ in the types of risks they generate, supporting the need for additional expertise over and above that provided by an accountant assurer. However, this study also finds that while review and nominal teams achieve a similar performance level in the GHG risk assessment task, MDGHGTs do not benefit from performing this task in an interactive manner. These findings suggest that MDGHGTs may not be able to capitalise on the benefits of the diverse knowledge and perspectives brought to the team by individual team members due to process losses occurring when these perspectives require discussion and reconciliation by the team. These findings thereby have implications for the team processes employed by assurance firms undertaking GHG assurance.

COPYRIGHT STATEMENT

‘I hereby grant the University of New South Wales or its agents the right to archive and to make available my thesis or dissertation in whole or part in the University libraries in all forms of media, now or here after known, subject to the provisions of the Copyright Act 1968. I retain all proprietary rights, such as patent rights. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

I also authorise University Microfilms to use the 350 word abstract of my thesis in Dissertation Abstract International (this is applicable to doctoral theses only).

I have either used no substantial portions of copyright material in my thesis or I have obtained permission to use copyright material; where permission has not been granted I have applied/will apply for a partial restriction of the digital copy of my thesis or dissertation.'

Signed……………………………..……

Date ..……14 January 2015…………

AUTHENTICITY STATEMENT

‘I certify that the Library deposit digital copy is a direct equivalent of the final officially approved version of my thesis. No emendation of content has occurred and if there are any minor variations in formatting, they are the result of the conversion to digital format.’

Signed……………………………..……

Date ..……14 January 2015…………

ORIGINALITY STATEMENT

'I hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, or substantial proportions of material which have been accepted for the award of any other degree or diploma at UNSW or any other educational institution, except where due acknowledgement is made in the thesis. Any contribution made to the research by others, with whom I have worked at UNSW or elsewhere, is explicitly acknowledged in the thesis. I also declare that the intellectual content of this thesis is the product of my own work, except to the extent that assistance from others in the project's design and conception or in style, presentation and linguistic expression is acknowledged.'

Signed…………………………………..……

Date ..……14 January 2015…………

ACKNOWLEDGEMENTS

I would like to use this opportunity to express my gratitude to a number of people who have contributed in so many ways to the completion of this dissertation.

My sincere gratitude goes to my supervisors, Associate Professor Wendy Green, and Scientia Professor Ken Trotman for their invaluable expertise, advice, comments, suggestions, dedication, and support throughout my PhD. My special thanks also go to my dissertation panel member, Scientia Professor Roger Simnett for his continued help, support, and generosity from the very beginning until the end of my PhD, specifically for his supervision on the first study of this dissertation. Without my three supervisors, the completion of this dissertation would not have been possible.

I would also like to thank the School of Accounting staff at UNSW who provided constructive feedback, comments, and suggestions on my proposal and working papers, and for continually tracking my progress. In particular, I wish to thank Professors Gary Monroe and Mandy Cheng, Associate Professors Peter Roebuck and Noel Harding, Dr Yee Shih Phua, Dr Sarah Kim, and Dr Trish Strong. I also wish to thank Colin Withers, Lisa Lifman, and Kathy Song for their administrative support over the past four years.

This dissertation has also benefited from comments from the participants of the AFAANZ Doctoral Symposium 2011, the anonymous reviewers and conference participants at the 2014 Annual Meeting of the American Accounting Association, and the European Accounting Association Annual Congress 2014, as well as the workshop participants at UNSW. I would also like to thank Professors Steve Salterio, Karla Johnstone, and Robyn Moroney for their helpful comments and suggestions.

The official thanks are due for the financial support throughout the entire period of this dissertation by the joint scholarship between Chiang Mai University (CMU) and the UNSW Business School, and the scholarship from CMU Business School. I would also like to thank Scientia Professors Ken Trotman and Roger Simnett for providing additional funding from their ARC Discovery Grant, and to thank the ASB Experimental Lab for the IT support funding.

I would also like to thank Jenny Merkley for her assistance in the development of the research instrument for the second study, Jim Park for developing the online-instrument and his excellent computer programming support for all experimental sessions, and the participants completing the two studies for their time and cooperation. The completion of this dissertation would not have been possible without the research assistance of Anna Huggins and Ray Koroy, as well the assistance in supervising experimental sessions of Shan Zhou, Ashna Prasad, Janhanzeb Khan, Lin Liao, and Weihang Zhou.

I am very grateful to the friendship of all my PhD friends at UNSW, especially Shan, Ashna, Ray, Tanya, Tina, Jahan, and Hien for always being there for me as true friends during the brightest and darkest hours of my long journey. Thanks for being my listening ear, helping hand, and for keeping on telling me that I can do it. I cannot imagine how I could go through this tough process without you all around. I also want to thank my close friends in Thailand for always believing in me, keeping me motivated and cheering me up every time I was down.

A special thanks goes to my colleagues at the Department of Accounting, Chiang Mai University, for their kind understanding and support, and for taking the responsibility for my work load during my absence.

My deepest appreciation goes to my family, especially my dad who always supports me with his love and encouragement. His courage to go through a massive surgery in order to make sure that he has a chance to see me finishing my PhD has motivated me through to the last stage of the study. I would also like to thank my aunties and their family in Australia for their love, help, and support while I was living in Sydney.

Lastly, most credit goes to my beloved mom who never had anything else in her sight, than to give me the best in this world and the highest education possible. I know you are here by me in every step I took, and I would not have been as strong and determined as I am to complete my dissertation without you looking down on me from heaven. My PhD and this dissertation are dedicated to you, my dearest mom.

ABSTRACT

The growing pressure for public disclosure on Greenhouse Gas (GHG) emissions has led to a significant demand for assurance on GHG statements. Given the complexity of GHG assurance, the international GHG assurance standard requires multidisciplinary GHG assurance teams (MDGHGTs) comprising practitioners from accounting and non-accounting background to discuss and assesses the risks of material misstatements in the planning stage of the engagement. However, prior research highlights the difficulties, including impaired effectiveness, associated with integrating practitioners from a range of disciplines, and as such suggests the importance of identifying mechanisms to ensure the effectiveness of MDGHGTs is not impaired. This dissertation addresses the effectiveness of MDGHGTs through two studies.

Study One utilises a retrospective recall approach to explore the factors that GHG assurers perceive could affect the effectiveness of MDGHGTs. This study finds that team processes are crucial to the success of MDGHGTs. In particular, having sufficient elaboration on different perspectives significantly increases the perceived effectiveness of MDGHGTs. This study also finds that the perceived sufficiency of elaboration increases when MDGHGTs become more educationally diverse, while the perceived sufficiency of discussion time decreases when the MDGHGTs become larger.

Study Two focuses on MDGHGT processes through a controlled experiment examining the effect of three team formats (nominal, interacting and review teams), on MDGHGT risk assessment performance. This study finds that accountant and non-accountant practitioners differ in the types of risks they generate, supporting the need for additional expertise over and above that provided by an accountant assurer. However, this study also finds that while review and nominal teams achieve a similar performance level in the GHG risk assessment task, MDGHGTs do not benefit from performing this task in an interactive manner. These findings suggest that MDGHGTs may not be able to capitalise on the benefits of the diverse knowledge and perspectives brought to the team by individual team members due to process losses occurring when these perspectives require discussion and reconciliation by the team. These findings thereby have implications for the team processes employed by assurance firms undertaking GHG assurance.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Introduction

In response to increasing pressure across stakeholder groups to account for the environmental impact of their activities, companies have become more actively engaged in sustainability projects and reporting corporate responsibility (CR) information in the public domain (KPMG 2008, 2013; Simnett et al. 2009b; Cohen et al. 2012; GRI 2013). This reported CR information covers social and environmental performances, including greenhouse gas (GHG) emissions disclosures. The increased demand for CR reporting has been highlighted by a KPMG (2013) survey on CR reporting. This survey reports significant growth (from 20 percent in 2011 to 51 percent in 2013) in the number of companies worldwide that include CR information in their annual financial reports. It also shows that CR reporting is now a mainstream business practice worldwide, as 71 percent of the 4,100 companies surveyed (76 percent of companies in the Americas, 73 percent in Europe, and 71 percent in Asia Pacific) now report on CR. Companies have also responded to the escalating demand that GHG emissions information be disclosed. Using the Carbon Disclosure Project (CDP) database, which provides GHG emissions information from the largest organisations in the world, Zhou et al. (2013) find that 761 of the 1,483 (51.3 percent) responding companies in 2011 disclose their GHG emissions to the public. In 2013, the Global Reporting Initiative (GRI 2013) reports a major increase in the practice of sustainability reporting since 2001, with 95 percent of the world's largest corporations that currently publish sustainability reports including their GHG emissions information.

The growing pressure for public disclosure and transparency of GHG emissions has led to a significant demand for assurance on GHG emissions statements (Simnett et al. 2009b; Dhaliwal 2011; Huggins et al. 2011). In 2012, over 46 percent of the reports listed in GRI's Sustainability Disclosure database assure CR disclosures, including GHG emissions (GRI 2013), with 59 percent (an increase of 13 percent from 2011) of the world's 250 largest companies assuring their CR disclosures in 2013 (KPMG 2013).

Further, among the Global 500 companies, 270 companies verify and/or assure their GHG emissions in 2013 (CDP 2013).

The need for GHG emissions information to be reported and assured is supported by the fact that global climate finance reached US$364 billion in 2011 (World Bank 2013) and by the value of the global carbon market, which is currently valued around US$30 billion (World Bank 2014) after reaching US$176 billion in 2012 (World Bank 2012). Further, 40 countries and 20 sub-country jurisdictions, including the United States and China (i.e., the world's two largest emitters), are implementing various carbon pricing approaches such as carbon taxes, emissions trading schemes, and crediting mechanisms (World Bank 2014). In 2012, the International Auditing and Assurance Standards Board (IAASB) recognises the need to ensure the credibility of GHG emissions information and issues International Standard on Assurance Engagements (ISAE) 3410, "Assurance engagement on greenhouse gas statements" (IFAC 2012a), to provide comprehensive guidance for the assurance of GHG emissions reports. The assurance of GHG emissions is expected to result in increased public confidence on the quality of GHG data and facilitate the efficient operation of Emissions Reporting Schemes and Emissions Trading Schemes around the world (PwC 2007; KPMG 2008, 2013; Simnett et al. 2009a).

ISAE 3410 addresses the responsibilities and work effort of practitioners undertaking GHG assurance engagements. A key aspect of this standard is that given the high complexity and specific knowledge needed to undertake GHG assurance engagements, ISAE 3410 recognises the need for a multidisciplinary capacity to achieve GHG assurance quality. For example, ISAE 3410 states: "The engagement may be performed by a multidisciplinary team that includes one or more experts, particularly on relatively complex engagements when specialist competence in the quantification and reporting of emissions is likely to be required" (IFAC 2012a, para. A42).

Because the subject matter of GHG assurance, that is, emissions data, is subject to scientific estimation and uncertainties (Green and Li 2009; Simnett et al. 2009a), GHG assurance teams commonly consist of accountants, engineers, and scientists (Nugent 2008; Huggins et al. 2011). As such, the terms 'accountant' and non-accountant' are used to refer to practitioners' educational background. Although circumstances can arise when it is necessary to engage non-accountant practitioners (e.g., actuaries, IT specialist) for some aspects of financial statement audits, practitioners with scientific backgrounds are required for many GHG assurance engagements and are indispensable for complex GHG assurance engagements (IFAC 2012a, para. A19 and A42). As such, while the use of such non-accountant practitioners for financial audits is usually on an ad hoc basis and for consultation purposes (Selley 1999; Griffith 2014), non-accountant GHG assurance practitioners are included as an integral part of the GHG assurance team (IFAC 2012a, para. 16b).

Given that GHG statements cover a wide range of company-specific circumstances, the unique and complementary skills of accountant and non-accountant practitioners in multidisciplinary GHG assurance teams (hereafter, MDGHGTs) are expected to improve assurance quality (Huggins et al. 2011). Consequently, ISAE 3410 requires multidisciplinary teams (hereafter, MDTs) to be involved in planning assurance engagement, including discussions to assess the entity's potential material misstatements due to fraud or error: "The engagement partner and other key members of the engagement team, and any key practitioner's external experts, shall discuss the susceptibility of the entity's GHG statement to material misstatement whether due to fraud or error, and the application of the applicable criteria to the entity's facts and circumstances" (IFAC 2012a, para. 29).

However, the difficulties that may be associated with GHG assurance teams comprising a range of disciplines and backgrounds were flagged during the IAASB Consultation Paper[1] stage of the ISAE 3410 standard's due process:

"Given that engagements, in particular complex engagements, are ordinarily undertaken by a multidisciplinary team, does the working draft adequately reflect how multidisciplinary teams should operate? For example, does the working draft adequately address the collective competence and capabilities of the team? What further improvements could be made?" (IAASB 2009, question 3, p. 8).

[1] A draft assurance standard, ISAE 3410 "Assurance on a Greenhouse Gas Statement", was presented for discussion to the IAASB meeting in September 2009. This process resulted in a Consultation Paper on the draft assurance standard that was open for consultation between October 2009 and February 2010. The purpose of this Consultation Paper was to seek views from practitioners and other stakeholders in relation to the IAASB's project to develop ISAE 3410.

Therefore, while the specific requirements and guidance provided in ISAE 3410 aim to enhance GHG assurance quality, a GHG assurance team's ability to deliver effective assurance depends upon the functionality of the MDTs.

MDTs are defined as teams comprising individuals from different educational backgrounds who are diverse in their knowledge and skill domains (Van der Vegt and Bunderson 2005). Although some reasons support the expectation that interactions among MDT members could result in better decision making (Guzzo and Dickson 1996; Williams and O'Reilly 1998), psychology research in team diversity has found mixed evidence regarding the benefits and limitations of MDTs on team performance (van Knippenberg and Schippers 2007). Diverse members bring to the task greater knowledge and skill-sets that enhance team creativity and decision making (Guzzo and Dickson 1996; Williams and O'Reilly 1998). However, individuals with diverse educational backgrounds may have different frames of reference, professional language, and problem-solving styles that impede the optimum sharing and recognition of diverse ideas and information (van Someren et al. 1998; van Asselt 2000). O'Dwyer (2011) notes the tension that arises from different mindsets in the sustainability assurance setting, in which accountant and non-accountant assurers work together on engagement. However, no research has yet explored ways to enhance the effectiveness of MDTs in a complex setting, such as conducting GHG assurance.

The overarching aim of this dissertation is to examine opportunities to improve the effectiveness of MDGHGTs. The specific research questions examined in this thesis are achieved through two related studies. The first study examines MDGHGTs currently involved in GHG assurance in practice to determine factors perceived as associated with the effectiveness of those teams. A retrospective recall methodology (Gibbins et al. 2001; Gibbins and Trotman 2002; Fargher et al. 2005; Gibbins et al. 2007) is adopted requiring assurance professionals to report on the effectiveness of GHG assurance teams they have previously worked on and data specific to the actual engagements recalled. The second study uses a controlled experimental setting to examine the effect of team processes, specifically team format, on the performance of MDGHGTs in a GHG assurance task. Three different team formats (nominal, interacting, and review teams) suggested by the previous literature to enhance the performance of audit teams are compared (e.g., Trotman 1985; Ismail and Trotman 1995; Carpenter 2007; Chen et al. 2014). A two-person nominal team (Diehl and Stroebe 1987, 1991) is included to set a baseline or control for evaluating team performance; this team was formed by combining outputs from one member with an accounting degree and/or financial audit experience (hereafter, an accountant practitioner) and one member who did not have such a degree and/or experience (hereafter, a non-accountant practitioner). The interacting team comprised both an accounting and a non-accounting practitioner, and team members were required to interact with each other through team discussion. The review team was operationalised by an accountant practitioner reviewing the work of a non-accountant practitioner.

1.2 Research Aims

The aims of this thesis are twofold. The first aim, to provide empirical evidence on factors with a significant impact on MDGHGTs' perceived effectiveness, is addressed by Study One. A research framework is proposed using various factors informed by team effectiveness frameworks in social psychology (e.g., Cohen and Bailey 1997; Ilgen et al. 2005; Mathieu et al. 2008) and studies on audit quality (e.g., O'Keefe et al. 1994; Hackenbrack and Knechel 1997; Reynolds and Francis 2001; Gibbins and Trotman 2002; Asare et al. 2005; Carey and Simnett 2006; Knechel et al. 2009; Li 2009). The framework includes environmental factors, task characteristics, team composition, and team process variables. Using the retrospective recall approach by GHG assurance team members in a field setting allows this study to explore GHG assurance engagement characteristics, team composition, and team process features not currently explored in the literature. The data are reported by GHG assurance professionals who were in a position to report on their own experiences as part of a GHG assurance team for two separate engagements: one example in which they felt the team worked more effectively together and one in which they felt the team worked less effectively together.

Because Study One focuses on factors that are under the control of the assurance firms, the impact of variables related to team composition (i.e., team size and the level of educational diversity) and team processes (i.e., sufficiency of team discussion and elaboration of different information and perspectives) on the effectiveness of MDGHGTs are tested. Further, the study examines the relationship between the team composition variables and the team process variables. Other factors inherent to the GHG assurance setting, including environmental factors (i.e., number of facilities, complexity of the GHG emissions profile, type of company, quality of client's internal control, familiarity with client, and client importance) and task characteristics (i.e., task interdependence and type of task) are treated as control variables in this study. The use of this approach also demonstrates how MDGHGTs are operationalised in current practice by gathering a large amount of descriptive data about GHG assurance engagement characteristics, team composition, and team processes, which provides insights that are new to the literature. This evidence is then used to inform the experimental design of Study Two.

The second aim is to explore ways to improve MDGHGTs' performance in the planning stage risk assessment for a GHG assurance engagement by comparing the performance of three different team formats: nominal, interacting, and review teams. This aim is addressed by Study Two. This study focuses exclusively on MDGHGTs' team processes, which the first study found to be significant determinants of team effectiveness. A controlled experiment using GHG assurance practitioners is employed to address this aim. A GHG assurance case scenario developed in conjunction with experts from a Big Four audit firm is used to examine the risk generation and risk selection performance of teams comprising an accountant practitioner and a non-accountant practitioner in terms of the quantity, breadth, and depth of risks generated. It also examines the difference between the types of risks accountant and non-accountant practitioners generate and select.

Given that accountant and non-accountant practitioners are expected to work together on GHG assurance engagements, three additional research questions are specifically addressed in this study: (1) is there cognitive diversity in MDGHGTs working together on risk generation and selection tasks; (2) how do different team formats affect the generation of risks; and (3) how do different team formats affect the utilisation of diverse information and perspectives. In addition, the study explores two further processes: information elaboration (van Knippenberg et al. 2004) and cross-understanding (Huber and Lewis 2010). The social psychology literature suggests that these processes are important mechanisms underlying the positive effect of cognitive diversity on team performance.

1.3 Contributions

The key theoretical and practical contributions made by this dissertation are outlined below.

1.3.1 Theoretical Contributions

First, Study One contributes to the audit literature by examining the factors contributing to the effectiveness of assurance teams comprising multidisciplinary practitioners. While the effectiveness of hierarchical financial audit teams comprising members with accounting backgrounds has been addressed in the audit literature (see Rich et al. 1997b and Nelson and Tan 2005 for reviews), very limited research has been conducted on assurance teams composed of practitioners from different disciplines. To date, O'Dwyer (2011) is the only study examining how multidisciplinary assurance practitioners carry out sustainability assurance engagements in practice. Study One adds to the extant literature by deepening the understanding of factors underlying the success of MDTs by employing the GHG assurance context, in which practitioners from accounting backgrounds work with practitioners from non-accounting backgrounds.

Second, Study One contributes to the auditing and social psychology literature by applying team effectiveness frameworks (e.g., Cohen and Bailey 1997; Ilgen et al. 2005; Mathieu et al. 2008) in the context of the MDTs used for GHG assurance. Since team effectiveness depends heavily on context, rapid changes in business environments, emerging markets, and regulations have highlighted the importance of understanding team effectiveness in different disciplines. Study One, therefore adds to the existing framework by introducing additional variables unique to the assurance setting, including variables related to the client's inherent risks and the client-assurer relationship.

Third, Study One contributes to the literature on group decision making in psychology. This literature finds that expanding discussion time in the start-up phase of a task increases the likelihood that unique information will be shared and considered by team participants who are homogeneous in their educational backgrounds yet informationally diverse (Larson et al. 1994). Study One demonstrates that this finding also holds for educationally diverse teams. Study One also examines van Knippenberg et al.'s (2004) proposition that information elaboration is a key process underlying the positive effects of diversity on team effectiveness by examining how the perceived sufficiency of information elaboration affects MDGHGTs' effectiveness and how the educational diversity level affects the perceived sufficiency of information elaboration.

Fourth, Study Two contributes to the limited empirical evidence on multidisciplinary assurance teams by demonstrating how the cognitive diversity between accountant and non-accountant practitioners affects the performance of assurance teams in the GHG assurance setting. In the sustainability assurance setting, O'Dwyer (2011) finds that cognitive diversity between accountant and non-accountant practitioners inhibits a team's ability to work well together. Different mindsets on how to approach the data and different concepts of materiality lead to tension between accountant and non-accountant practitioners in this setting. Study Two adds to the existing evidence by demonstrating how cognitive diversity can benefit multidisciplinary teams in the GHG assurance context.

Fifth, Study Two contributes to the audit brainstorming literature by examining the effects of different team formats suggested by previous fraud brainstorming studies (Carpenter 2007; Lynch et al. 2009; Chen et al. 2014) on the risk generation performance of assurance teams comprising practitioners from different educational backgrounds. While the effects of different team formats on brainstorming audit teams' performance have been examined in the fraud audit setting (Carpenter 2007; Hoffman and Zimbelman 2009; Trotman et al. 2009; Chen et al. 2014), teams in these studies are different in their hierarchical nature but not different in their educational backgrounds (i.e., they typically have accounting/financial audit backgrounds). This dissertation extends the previous literature by testing the findings from previous audit brainstorming studies in a GHG assurance setting, in which practitioners from different disciplines (i.e., accounting, environmental science, and engineering) are required to work together to assess the risk of material misstatement. Study Two also adds the use of the review process to the brainstorming phase. As such, it adds to the audit review literature by examining the review process in a case in which reviewers and reviewees have different educational backgrounds, i.e., in which accountant practitioners review the work of non-accountant practitioners.

Sixth, Study Two examines how MDGHGTs utilise their diverse knowledge and perspectives when generating and selecting risks in the planning stage of GHG assurance engagements. To achieve this aim, two important aspects of quality—the breadth and depth of risks generated and selected—are examined. The breadth and depth of ideas has been used to measure the quality of ideas generated by cognitively diverse teams in psychology literature (Stroebe and Diehl 1994; Nijstad et al. 2002; Dahlin et al. 2005; Kohn and Smith 2011) and the quality of audit procedures identified in the audit literature (Asare et al. 2000; Green and Trotman 2003). This study contributes to the audit literature by examining the breadth and depth of risks in both idea generation and selection tasks and by testing previous studies' findings in the context of GHG assurance, in which teams have highly diverse educational backgrounds.

Both studies build on previous information elaboration studies (van Ginkel et al. 2009; van Ginkel and van Knippenberg 2009) by examining the effect of information elaboration on decision-making tasks. Study One contributes to the literature by exploring the relationship between information elaboration, team diversity, and team effectiveness. Study Two builds on Study One to extend both the auditing and social psychology literatures by testing the effect of information elaboration on idea generation tasks. Study Two also explores the possibility that a high level of cross-understanding—"extensive and accurate understanding among members about the factual knowledge, beliefs, sensitivities, and preferences of other members" (Huber and Lewis 2010, p.12)—could cancel the negative effects of cognitive diversity found in MDTs (such as difficulties communicating and understanding different knowledge and perspectives) and thus improve their performance. Testing these social psychology findings in the context of GHG assurance provides an understanding of the underlying mechanisms of the differential outcomes between nominal, interacting, and review teams.

1.3.2 Practical Contributions

Given the relative novelty of contemporary GHG assurance and the unique multidisciplinary nature of GHG audit teams, assurance firms remain in the early stages of developing and implementing these teams. GHG assurance engagements are becoming a major business stream for leading assurance firms (KPMG 2011, 2013; O'Dwyer 2011). The results of the two studies in this dissertation provide specific implications.

Study One provides an understanding of team composition and team processes that may improve MDGHGT effectiveness. This dissertation addresses the dearth of information relating to the factors that help MDTs undertake effective GHG assurance engagements. These factors include the manner in which such teams are operationalised. In its application of frameworks and findings from the social psychology literature, this dissertation identifies and addresses challenges to GHG assurance practice by engaging the specific factors that are perceived to affect successful GHG assurance outcomes. The findings contribute to the standards of assurance practice and inform practitioner and scholarly understanding of GHG assurance by exploring factors that affect multidisciplinary audit team effectiveness. This understanding is important in advancing the effective use of such teams in the newly emerging area of GHG assurance. Study One contributes to an understanding of the GHG assurance practice by providing rich data on client characteristics, GHG engagement characteristics, team composition, and the processes involved in conducting GHG assurance engagements.

Study Two provides evidence that the team format used for GHG assurance engagements can affect MDGHGTs' performance. ISAE 3410 suggests that MDTs should work together to undertake GHG assurance engagements and that the discussion between MDGHGT members in the planning stage of GHG assurance engagement is beneficial to assessing the risks of material misstatements due to fraud and errors (IFAC 2012a, para. A42 and 29). Study Two provides empirical evidence on how different team formats can affect risk generation performance and information use by MDGHGT members. This study also examines how information elaboration and cross-understanding enhance or inhibit the risk generation and selection performance of multidisciplinary teams. All the firms are in the early stages of undertaking these engagements, and evidence on the performance of different team formats informs how these teams can work together more effectively in practice.

Another key contribution of this dissertation is informing the growing number of international regulations and standard setting processes that provide guidance to GHG assurance providers. Specifically, the findings of this dissertation will provide insights relating to the question raised by the IAASB Consultation Paper on how to operationalise GHG assurance multidisciplinary teams.

1.4 Structure of the Dissertation

To achieve the research aims of this dissertation, two studies were developed to investigate MDGHGTs' effectiveness. Chapter 2 provides background information relating to GHG assurance and reviews the literature related to factors influencing team effectiveness. This chapter engages literature on team effectiveness and on group decision making, particularly in auditing. The review of this literature highlights a research gap in the study of team formats that can be employed to improve MDGHGTs' performance. The chapter also explains the roles of information elaboration and cross-understanding on MDT performance. Chapter 3 describes Study One, which provides empirical evidence on those factors GHG assurance practitioners perceive as affecting the successful execution of multidisciplinary GHG assurance teams. A retrospective recall study is conducted with participants from Big Four audit firms who are members of MDGHGTs. Chapter 4 describes Study Two, which builds on the results of Study One to demonstrate how the team format affects MDGHGTs' performance by comparing between nominal, interacting, and review teams. A controlled computerised experiment is conducted with participants from Big Four audit firms. This chapter also reports the exploratory analyses carried out to test whether information elaboration and cross-understanding are associated with MDGHGTs' performance. Finally, Chapter 5 summarises the findings from Study One and Study Two and discusses the important contributions made by this dissertation. Implications for the assurance profession and suggestions for future research complete this chapter.

CHAPTER 2

BACKGROUND AND LITERATURE REVIEW

2.1 Introduction

This chapter provides background information on the nature of greenhouse gas (GHG) assurance engagements and the multidisciplinary GHG assurance teams (MDGHGTs) undertaking such engagements. The chapter also reviews literature from various disciplines including psychology, management, and auditing to provide a theoretical and empirical basis for understanding how to improve the effectiveness of MDGHGTs.

Section 2.2 of this chapter provides background information on GHG assurance engagements and MDGHGTs, including the unique nature of GHG assurance engagements, the need for MDGHGTs, and how these teams are operationalised. The literature review is provided in subsequent sections. Section 2.3 reviews studies examining the benefits of adopting a MDT approach and the problems faced by MDTs. Section 2.4 reviews the team effectiveness frameworks developed by studies in psychology and management, which provide a theoretical framework for examining the factors associated with MDGHGTs' effectiveness. Based on the inputs-processes-outcomes (IPO) approach adopted by these team effectiveness frameworks (Cohen and Bailey 1997; Kozlowski and Ilgen 2006; Mathieu et al. 2008), the following sections discuss three aspects: team outcomes, team inputs, and team processes. Section 2.5 explores various team outcomes used in the group decision-making literature to measure team effectiveness. Section 2.6 reviews studies investigating the effects of team input variables on team performance with a focus on three categories of team inputs: environmental factors (i.e., client characteristics and risks and the client–assurer relationship), task characteristics (i.e., task interdependence and task type), and team composition (i.e., team size and diversity). Section 2.7 focuses on literature studying the effects of team processes on team effectiveness, with a particular focus on the effects of three different team formats—nominal, interacting, and review teams—on team performance. Finally, Section 2.8 summarises this chapter.

The majority of the literature reviewed in this dissertation is derived from studies on group decision making, work team diversity and team effectiveness. While it is important to note that some differences between ‘‘groups”[2] and “teams” exist, these terminologies are often used interchangeably in the organisational psychology literature (Guzzo and Dickson 1996). Because this dissertation examines assurance teams currently working in the field and conducts group experiments by adopting literature on group decision making and team effectiveness, which is consistent with the organisational psychology literature, the words “group” and “team” will be used interchangeably.

2.2 GHG Assurance and Multidisciplinary GHG Assurance Teams

2.2.1 GHG Assurance Engagements

Assurance engagements on GHG statements are an emerging area of assurance services that aim to enhance the reliability of emissions information reported by an entity (Simnett at al. 2009a). ISAE 3410 defines a GHG statement as “a statement setting out constituent elements and quantifying an entity’s GHG emissions for a period (sometimes known as an emissions inventory) and, where applicable, comparative information and explanatory notes including a summary of significant quantification and reporting policies. The GHG statement is the subject matter information of the engagement” (IFAC 2012a, para. 14m). Given the inherent uncertainty in quantifying precise activity data and emissions factors (Simnett et al. 2009a) and the fact that the process for generating GHG information is inherently less robust than the process for financial statements (Nugent 2008), the audit risk model adopted in the financial statement audit practice can also be applied to GHG statement assurance (Huggins et al. 2011). As a result, ISAE 3410 adopts a risk-based approach for both reasonable and limited assurance engagements; it requires GHG assurance practitioners to understand

[2] Groups has been defined as “two or more individuals who have some interdependence or relationship and who have an influence on each other through their interactions” (Paulus 2000, p. 238). A team has been defined as a group embedded in an organisation who work together interactively on highly interdependent tasks, who have distinctive roles and responsibilities, and who share common goals and values (Kozlowski and Ilgen 2006). Teams can also be classified as a type of group that develops a higher degree of interdependence and integration than other types of groups (Cohen and Bailey 1997).

the entity and its environment and to identify and assess risks of material misstatements due to fraud or error (IFAC 2012a, paras. 6 and 24).

Practitioners are required to assess the risks of material misstatement at the GHG statement level for a limited assurance engagement and at both the GHG statement and assertion levels for a reasonable assurance engagement (IFAC 2012a, paras. 25L and 25R). The assertions used in the financial audit setting translate well into the GHG assurance setting because, similar to an income statement, a GHG statement is prepared periodically (e.g., annually). In particular, this allows practitioners to use assertions regarding the quantification (e.g., occurrence, completeness, accuracy, cut-off, classification) and presentation and disclosure (e.g., occurrence, completeness, classification, accuracy, consistency) of emissions for the period assured to consider the different types of potential material misstatements that may occur (IFAC 2012a, para. A82).

Given that the risk-based and assertion-based approaches commonly used for financial audit engagements have been adopted by GHG assurance engagements, accountant practitioners are well placed to undertake these assurance engagements (Huggins et al. 2011). Further, accounting seems to be the only discipline providing assurance training to their members (Gray 2000), thus accountant practitioners are well recognised for their competencies in financial accounting and audit methodology (Huggins et al. 2011). However, unlike a financial audit, the subject matter assured in a GHG assurance engagement involves non-financial data, specifically emissions data. The quantification of GHG emissions relies heavily on scientific estimations and uncertainties (Green and Li 2009; Simnett et al. 2009a). Given that the scientific knowledge and skill-sets required to undertake GHG assurance engagements are traditionally outside the accounting discipline, GHG emissions assurances are currently performed by teams of accountant practitioners and non-accountant practitioners (e.g., engineers and environmental scientists) (Nugent 2008). The necessity of both accounting/assurance and scientific competencies is recognised by ISAE 3410, with the standard indicating that engagements under ISAE 3410 are usually expected to be undertaken by a "multidisciplinary team", and that such teams are indispensable in complex GHG assurance engagements (IFAC 2012a, paras. A19 and A42).

2.2.2 Multidisciplinary GHG Assurance Teams (MDGHGTs)

The use of MDTs is not new to the auditing profession because in some circumstances, it is necessary to engage non-accountant practitioners (IFAC 2010), such as when IT specialists are needed to address significant risks due to a client's complex IT systems or actuaries are needed to determine appropriate loan loss provisions. However, the use of such non-accounting practitioners for financial audits is usually on an ad hoc basis rather than as integral members of the assurance team, as is the practice for MDGHGTs (Selley 1999; Griffith 2014). In the traditional financial audit setting, non-accountant practitioners are usually treated as consultants, and their suggestions are not fully relied on (Selley 1999). Unlike in traditional financial audit engagements, non-accountant GHG assurance practitioners are included as part of an integrated assurance team with collective professional competencies (IFAC 2012a, para. 16b).

The complexity of GHG statements varies considerably from engagement to engagement. When an engagement is relatively complex, specialist competence in quantifying and reporting emissions is likely to be required, including information systems expertise (e.g., to help understand how emissions information is generated) and scientific and engineering expertise (e.g., to identify emissions sources and analyse chemical and physical relationships between inputs, processes, and outputs) (IFAC 2012a, para. A19). Consequently, ISAE 3410 requires the integration of accountant and non-accountant practitioners into various stages of the engagement. For example, ISAE 3410 requires the involvement of key experts in planning and in discussions regarding the susceptibility of the entity's GHG statement to material misstatements, whether due to fraud or error, and the application of applicable criteria to the entity's facts and circumstances (IFAC 2012a, para. 29).

Given that GHG statements cover a wide range of circumstances, the benefit of MDTs in GHG assurance engagements lies in accountant practitioners' unique and complementary skills in assessing the risks of material misstatements in the entity's GHG statement combined with non-accountant practitioners' subject matter and technical expertise in understanding GHG quantification and measurement methodologies (Huggins et al. 2011). However, the effectiveness of the assurance function hinges on the optimum functioning of MDGHGTs, that is, how well accountant and non-accountant practitioners work together to utilise their diverse knowledge and perspectives. To improve MDTs' performance, it is important to understand the potential consequences the multidisciplinary nature of this type of team has on performance. Therefore, the next section explores literature highlighting the potential benefits of adopting MDTs and potential problems faced by such teams.

2.3 Multidisciplinary Teams: Distinctive Features

MDTs are teams composed of individuals with different educational or functional backgrounds, who bring together collective knowledge and expertise to bear on a complex problem or issue (Van der Vegt and Bunderson 2005). Given that the problem or issue is considered complex when it "lies across or at the intersection of various disciplines" (Van Asselt 2000, p. 2), the adoption of MDTs to deal with complex tasks is well recognised in the psychology and management literature (Jehn et al. 1999; Pelled et al. 1999; Bowers et al. 2000; Van der Vegt and Bunderson 2005; Kearney et al. 2009). In such tasks, MDTs are expected to benefit from their diverse knowledge and perspectives because they have access to a broad range of knowledge and perspectives, which yield insights for decision making that transcend the boundaries of each discipline (Van Der Vegt and Bunderson 2005; van Knippenberg and Schippers 2007).

The adoption of MDTs is evident in many studies and in different contexts, such as top management teams (Bantel and Jackson 1989; Wiersema and Bantel 1992; Bunderson and Sutcliffe 2002), research and development teams (Van der Vegt and Bunderson 2005), health care teams (Poulton and West 1999; Fay et al. 2006), financial audit teams (Selley 1999; Bortiz et al. 2014; Griffith 2014), and sustainability assurance teams (O'Dwyer 2011). Although these teams comprise members from different disciplines, they have different natures than MDGHGTs.

Top management teams comprise senior managers or executive managers responsible for each division/unit and are typically involved in important decision making and negotiation to ensure the success of their organisation (Cohen and Bailey 1997). Research and development teams are project teams that are usually cross-functional team comprising members from various disciplines and functional units such as engineering, manufacturing, and marketing (Brown and Eisenhardt 1995; Fong 2003). Thus, their members may need to rotate in and out of the team and return to their functional units after the project ends (Cohen and Bailey 1997). In contrast to top management teams and research and development teams, health care teams can be classified as work teams with memberships that are more permanent because they usually come from similar functional units (Cohen and Bailey 1997). However, members of health care teams (e.g., teams of psychiatrists, surgeons, and nurses) may not have as much diversity in educational and functional backgrounds as the MDTs mentioned earlier (i.e., they share at least some common ground in medical or biological science). Because multidisciplinary financial audit teams normally use experts from different disciplines for consultation purposes and operate as MDTs on an ad hoc basis (Selley 1999; Griffith 2014), they are similar to research and development teams in that they can be classified as project teams. However, non-financial assurance teams, such as MDGHGTs, differ in several ways from the other teams mentioned previously. First, MDGHGTs comprise team practitioners from distinctive disciplines rather than related educational backgrounds, such as engineering and/or science (e.g., chemical engineering and environmental science). As such, they have relatively more diverse knowledge bases than health care teams. Second, unlike top management teams and research and development teams, MDGHGTs do not have different functional areas, i.e., both accountant and non-accountant practitioners in assurance firms work as assurors or consultants in designated assurance divisions. As such, engineers or scientists in the MDGHGTs do not work in an engineering or science department. Third, MDGHGTs are different from multidisciplinary financial audit teams because non-accountant practitioners are integral parts of the assurance team rather than ad hoc consultants.

Given that MDT members are likely to bring different knowledge and perspectives to complex tasks, the exposure to ideas members may not have thought of on their own and the need to discuss and integrate these diverse ideas may stimulate their creativity, thereby leading to more innovative ideas and solutions (van Knippenberg et al. 2004). However, while MDTs seem to be an attractive option when tasks require consideration of multiple perspectives, working in diverse teams can be challenging. The research findings in the work team diversity literature suggest that including members with diverse educational backgrounds in the team does not necessarily lead to effective team performance (Milliken and Martins 1996; Williams and O'Reilly 1998; Pelled et al. 1999; De Dreu and Weingart 2003). On one hand, MDTs benefit from their members' diverse knowledge and perspectives (Guzzo and Dickson 1996; Williams and O'Reilly 1998). On the other hand, the differences in knowledge bases and mindsets make it difficult for MDT members to communicate and coordinate effectively with each other (Williams and O'Reilly 1998; van Knippenberg et al. 2004; Dahlin et al. 2005). A number of studies on MDTs find that team members with different educational/professional backgrounds use different language or terminology only understood by people in the same profession or that have different meanings in other fields (Carlile 2004; Sheehan et al. 2007). MDT members also have different frames of reference, which impedes MDTs' understanding, communication, and effectiveness (van Someren et al. 1998; van Asselt 2000). Further, the members' different functional and professional backgrounds may elicit social categorisation processes, such as out-group vs. in-group identification[3] (van Knippenberg and Schippers 2007), which leads to less communication (Bhappu et al. 1997) and cooperation (Chatman and Flynn 2001) between subgroups. Randel and Jaussi (2003) provide empirical evidence relating to these problems showing that in cross-functional teams, team members identified as coming from a functional background minority are less likely to contribute to their team because others may not value their opinion.

The difficulties associated with MDTs are also evident in assurance teams, in which accountant and non-accountant practitioners work together on sustainability assurance engagements. O'Dwyer (2011) conducts in-depth interviews with practitioners in two Big Four audit firms and find evidence that tensions arise from interactions between accountant and non-accountant practitioners because of their distinct concepts and mindsets. The non-accountant practitioners in this study criticise the accountant practitioners for bringing habits and mindsets from their experience with financial audit engagements to the sustainability engagement, which lead to constraints in thinking about and approaching data. In particular, the non-accountants thought the accountant practitioners follow standard substantive and compliance testing procedures too strictly,

[3] The way group members categorise themselves and others into groups based on the similarities and differences between group members (van Knippenberg et al. 2004). For example, group of male versus group of female, or group of accountants versus group of non-accountant practitioners.

focus heavily on assessing quantitative data accuracy, and do not have sufficient knowledge of the subject matter to evaluate the non-accountant practitioners' work. Different materiality concepts adopted by the accountant and non-accountant practitioners also introduce frustration in the MDTs. The tensions arising from different mindsets are not only recognised by the non-accountant practitioners but also by the accountant practitioners. From the accountant practitioners' perspective, bringing "financial audit" mindsets and approaches to non-financial assurance engagements is important to ensure that all the testing conducted by the non-accountant practitioners is defendable and complied with the firm's standard procedures. However, the accountant practitioners note that it could be difficult for them to avoid bringing financial audit habits to non-financial assurance engagements because, unlike non-accountant practitioners, they worked on both types of engagements. Results from the O'Dwyer (2011) study suggest that if the potential problems resulting from the use of MDTs to provide GHG assurance services are not properly addressed, the quality of the judgments, decision-making, and assurance provided may be affected.

In addition to the problems associated with MDGHGTs' multidisciplinary nature, other factors that are inherent and innate to the team could be associated with the effectiveness of MDGHGTs. Thus, the next section discusses different team effectiveness frameworks suggested by the literature in psychology and management to form a theoretical basis for examining factors that may affect MDGHGTs' effectiveness.

2.4 Team Effectiveness Framework

More than 100 model frameworks of team effectiveness have been developed over the past twenty years (see Salas et al. 2007 and Salas et al. 2008 for reviews). Of these, the most classic input-process-outcome (IPO) framework (McGrath 1964) primarily shapes the conceptualisation of team effectiveness (Kozlowski and Ilgen 2006). Figure 2.1 depicts an adapted version of this framework. In an attempt to illustrate the IPO framework, Mathieu et al. (2000, 2008) define "inputs" as factors facilitating and inhibiting interaction among team members (e.g., team member characteristics, task type and interdependence, environmental complexity), "processes" as interactions among team members that turn team inputs into outputs, and "outcomes" as products of the team activity that are useful to the organisation or other parties. However, more recent models (e.g., Cohen and Bailey 1997; Ilgen et al. 2005; Mathieu et al. 2008) focus on different processes that also mediate the relationship between team input and output, such as the social process (e.g., idea sharing) and the cognitive process (e.g., shared mental model, elaborating on different information and perspectives).

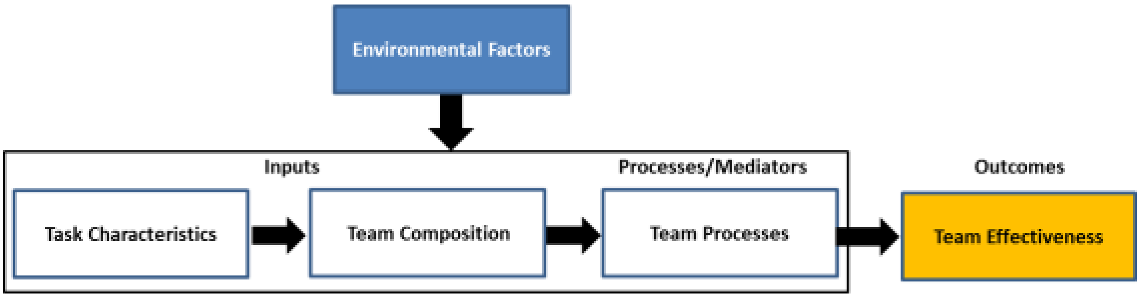


FIGURE 2.1 Modified Team Effectiveness Framework (Adapted from the team effectiveness frameworks developed by Cohen and Bailey 1997, Kozlowski and Ilgen 2006, and Mathieu et al. 2008)

The model develops by Cohen and Bailey (1997) distinguishes group psychological traits, such as shared understandings, beliefs, or emotional tone (e.g., shared mental model and cohesiveness), from the internal and external processes and also introduces environmental factors as drivers of team design. This model suggests that team environment factors, such as the nature of the business, can influence the composition of the team, which indirectly affects outcomes through team processes and psychological traits. In addition, team processes (e.g., communication) and psychological traits (e.g., shared mental models) are influenced by design factors such as team composition, which in turn affects team effectiveness.

More recently, Ilgen et al. (2005) introduce a new framework known as the input-mediator-outcome (IMO) model. In this model, “process” is replaced by the term “mediator” to cover other variables (e.g., psychological traits) that could mediate the relationship between team input variables and team effectiveness. Further, Ilgen et al. (2005) extend the IMO model to an IMOI (input-mediator-outcome-input) model by considering the changes in team performance that are influenced by various inputs and mediators over time (i.e., the time factor).

Kozlowski and Ilgen (2006) suggest that the environment dynamics and complexity influence the team task demand, which shapes the team design. The team members’ resources are then shared and integrated through team processes and emergent states[4] to yield team outcomes. Consistent with Ilgen et al. (2005), Kozlowski and Ilgen (2006) also consider the time factor by proposing that team outcomes may eventually affect the environment and result in a cycle.

In a comprehensive review of recent team effectiveness models, Mathieu et al. (2008) propose a succinct model adapted from the IPO and IMO frameworks. This model involves multilevel team input factors, i.e., team members are nested in teams that are nested in the organisation and environment. The model suggests that environmental factors influence the team task design and other team features, which in turn drive the demand for members with specific knowledge and expertise and affects the role of these members in the team. The mediation factors, including team processes and emergent states, then mediate the effect multilevel input factors have on team effectiveness. As teams mature, the model also suggests feedback loops in which teams may carry out tasks differently after seeing the outcomes, which can potentially affect the team design, task demand, organisation, and environment.

In summary, the team effectiveness frameworks discussed above suggest that task characteristics, team composition, and team processes affect team effectiveness and that these components can be influenced by contextual conditions such as environmental factors (Figure 2.1). Task characteristics can also affect team composition, and team composition can affect team effectiveness through team processes.

To understand how these input and process components could affect the effectiveness of teams, particularly MDTs, the following subsections discuss the research findings for each component. First, a description of various team outcomes used to measure team effectiveness in different team contexts will be presented. Second, research findings and particular variables related to team inputs, including environmental factors, task characteristics, and team composition, are discussed. Third, different team processes examined in both psychology and auditing research are discussed.

[4] Emergent states are constructs that emerge over time as team members interact and the team develops, such as team mental models, and team learning (Kozlowski and Ilgen 2006).

2.5 Team Outcomes

A wide range of outcomes have been used to measure effectiveness in team research (Sundstrom et al. 2000), and these measures have been considered at various dimensions and levels (Cohen and Bailey 1997; Mathieu et al. 2008). For example, Cohen and Bailey (1997) categorise team outcomes into three dimensions: performance effectiveness (e.g., quantity and quality of ideas, time spent, customer satisfaction, supervisor-rated performance), members' attitudes (e.g., team member satisfaction, commitment, team climate), and behavioural outcomes (e.g., employee turnover, safety). Given that different outcomes are achieved by teams at different levels of analysis, Mathieu et al. (2008) classify team outcomes into three levels: organisational, team, and individual.

Organisational-level outcomes (e.g., firms' economic performance indicators) are related to top management teams (TMTs) because TMTs' effectiveness is usually measured by firm performance. Individual-level outcomes (e.g., members' performance and satisfaction with the team) and team-level outcomes (e.g., quantity and quality of outcomes) are related to work teams and project teams (Cohen and Bailey 1997).

At the individual level, various performance measures have been captured, including members' role-fulfilment outcomes (Chen et al. 2007); attitudinal and behavioural measures, such as team climate and team commitment (Balkundi and Harrison 2006); members' satisfaction with the team (Shaw et al. 2011); and members' affective reactions (Kaplan et al. 2013),.

At the team level, studies on work teams generally adopt task-specific objective outcome measures (e.g., the percentage of budget allocated to replacement expense and machine breakdown times [Mathieu et al. 2006]) and objective behaviour measures, such as process indicators (e.g., reduction in team response time [Kirkman et al. 2004] and the elaboration of task-relevant information [Homan et al. 2007]). These studies assess performance effectiveness and subjective measures, such as members' perceptions of overall team performance (Sparrowe et al. 2001; Shaw et al. 2011) and supervisor-rated team performance (Hu and Liden 2011; Cole et al. 2013). For most project teams, effectiveness is measured using subjective measures, such as the perceptions of managers', members', and those outside the team (Cohen and Bailey 1997). For example, Haas (2010) measure project team effectiveness using independent expert panel ratings on strategic and operational effectiveness, which are part of firms' project evaluation process.

As demonstrated by the literature, team outcomes vary based on the type of team and the level of analysis because the natures of tasks performed by different types of teams and in different contexts are different. Therefore, one method of measuring team effectiveness could be to evaluate the outcomes most related to the types of tasks the team performs. The following subsections outline different tasks commonly performed by teams, the risk assessment task, and its outcomes; MDGHGTs are required to perform this type of task.

2.5.1 Team Tasks

According to McGrath's task circumplex (1984), team tasks can be categorised into four different categories: generating, choosing, negotiating, and executing (Figure 2.2). McGrath's task circumplex is illustrated in two dimensions. The horizontal dimension indicates whether the tasks are conceptual-oriented or behavioural-oriented. The vertical dimension indicates whether the tasks require team members to cooperate or resolve conflicts. Within each category, two types of tasks have been identified.

The first category, shown in Quadrant I, is generating. This category comprises Type 2 creativity tasks (generating ideas) and Type 1 planning tasks (generating plans). Although these two tasks are similar in that they require team members to cooperate rather than resolve conflicts, creativity tasks are more conceptual-oriented and planning tasks are more behavioural-oriented.

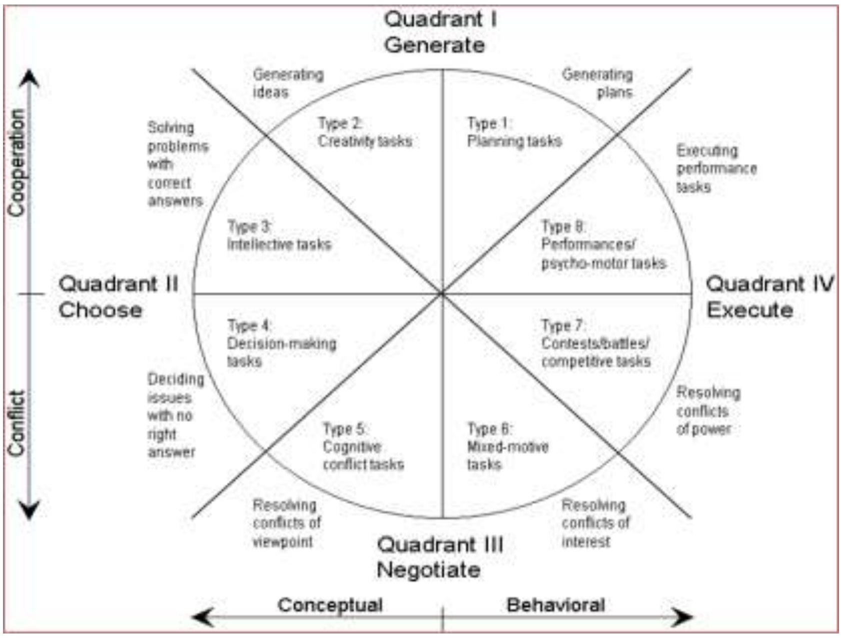


FIGURE 2.2 McGrath's (1984) Group Task Circumplex (Source: Graetz 2011)

The second category, shown in Quadrant II, is choosing. This category comprises Type 3 intellective tasks (solving problems that have correct answers) and Type 4 decision-making tasks (deciding issues that do not have any correct answers). These tasks are both conceptual-oriented tasks, but teams performing intellective tasks are more likely to engage in cooperative interaction because they must help each other select correct answers. Those performing decision-making tasks are more likely to engage in competitive interaction because they must select a preferred answer based on the team consensus.

The third category, shown in Quadrant III, is negotiating. This category is also divided into two tasks: Type 5 cognitive conflict tasks (resolving conflict viewpoints), which are conceptual tasks in which members have different representations or mindsets, and Type 6 mixed-motive tasks (resolving conflicts of interest), which are behavioural tasks that refer to other negotiation tasks such as management negotiations, bargaining, social dilemmas, and the allocation of payoffs.

The last category, shown in Quadrant IV, is executing. This category is heavily involved with physical behaviour or action. This category also includes two types of tasks: Type 7 competitive tasks, which involve competition against an opponent with a winner and a loser, and Type 8 performance tasks, which involve attempting to meet some standards rather than competing with an opponent. Therefore, the former involves engaging in resolving conflicts while the latter involves engaging in cooperative interactions.

In the auditing context, various tasks are performed ranging from generating (e.g., risk or hypothesis generation); selecting and assessing (e.g., risk or evidence evaluation); negotiating (e.g., auditor-client negotiations); and executing (e.g., substantive testing). Given the wide range in these tasks, the effectiveness of audit teams has been measured using various team outcomes according to the nature of tasks performed. For example, the quantity and quality of risks generated by the team (e.g., Trotman et al. 2009) can be used to evaluate the risk generation performance. Assessments of the significance of potential misstatements (e.g., Low 2004) can be used to evaluate the risk assessment performance. The amount of negotiated writedown (e.g., Trotman et al. 2005) can be used to evaluate the auditor–client negotiation performance. Perceptions of the reviewer's and preparer's performances (e.g., Brazel et al. 2010) can be used to evaluate the effectiveness of audit review teams.

In addition to the overall effectiveness of MDGHGTs, this dissertation focuses particularly on "risk assessment tasks", which are the tasks ISAE 3410 requires MDGHGT members to effectively complete together. Therefore, the next section further explains the nature of risk assessment tasks and outcomes that could be used to assess the effectiveness of teams performing such tasks.

2.5.2 Risk Assessment Task and Outcomes

This section describes the nature of risk assessment tasks and various team outcomes that can be adopted to assess MDGHGTs' effectiveness in performing risk assessment tasks. Risk assessment procedures are clearly defined by International Standard on Auditing (ISA) 315 ("Identifying and Assessing the Risks of Material Misstatement Through Understanding the Entity and Its Environment") as "the audit procedures performed to obtain an understanding of the entity and its environment, including the entity's internal control, to identify and assess the risks of material misstatement, whether due to fraud or error, at the financial statement and assertion levels" (IFAC 2012b, para. 4).

The importance of risk assessment procedures is demonstrated by the risk-based approach applied in auditing standards (e.g., ISA 315 [IFAC 2012b], ISAE 3000 "Assurance Engagements Other Than Audits or Reviews of Historical Financial Information" [IFAC 2013], and ISAE 3410 [IFAC 2012a]). These standards require auditors to perform risk assessment procedures to assess the risks of material misstatements (RMM). This diagnostic task allows auditors to understand the nature of their clients' business process, policies, and control environment, which then forms a basis for the design of substantive tests and resource allocation. Therefore, the failure to identify significant RMM or to effectively discuss/communicate information and perspectives during risk assessments, particularly during the planning stage, can lead to ineffective audit results (Low 2004; Fukukawa and Mock 2011).

To identify appropriate outcomes for risk assessment procedures, it is important to examine what is required at each stage of the task, including the risk generation, risk selection/evaluation, and plan generation stages. The nature of the task and the outcomes used to assess performance in each stage are discussed below.

2.5.2.1 Risk Generation

Once practitioners gather sufficient knowledge, understand the clients' business, and look for unexpected changes in account balances or ratios (i.e., perform analytical procedures), they are first required to generate a list of potential RMMs (IFAC 2012a, 2012b). The nature of this task is an idea generation task under McGrath's (1984) task circumplex. Idea generation tasks normally occur at the beginning of the problem-solving or decision-making processes, in which potential alternatives or solutions to problems are created (Osborn 1953). Prior studies, conduct mainly on the analytical procedures (i.e., hypotheses generation) and fraud brainstorming (i.e., fraud risks generation) setting, measure the effectiveness of audit teams performing idea generation tasks using both quantitative and qualitative measures. A summary of outcomes used by prior audit research to assess the effectiveness of teams performing risk assessment tasks is provided in Table 2.1.

TABLE 2.1 Summary of Outcomes Used to Measure the Effectiveness of Audit Teams Performing Idea Generation, Idea Selection or Evaluation, and Plan Generation Tasks

Research Study	Idea Generation	Idea Selection/Evaluation	Plan Generation
Ismail and Trotman (1995)	• Number of plausible hypotheses generated • Number of implausible hypotheses generated • Number of plausible hypotheses added to the reviewed list	• Number of hypotheses deducted from the reviewed list	
Asare et al. (2000)			• Number of tests generated • Number of hypotheses tested (breadth of testing) • Average number of tests conducted per hypothesis examined (depth of testing)
Green and Trotman (2003)	• Number of correct causes generated	• Number of correct causes selected	• Number of audit tests selected • Budgeted hours utilised • Number of hypothesis categories tested (breadth of testing) • Number of tests per hypothesis category (depth of testing)
Asare and Wright (2004)			• Number of identified benchmark procedures
Low (2004)		• Assessments of potential misstatements' significance	• Quality of audit procedure changes, final audit programs, and final time budgets (based on the expert panel's assessments)
Carpenter (2007)	• Number of fraud risks listed • Number of quality fraud risks listed	• Assessments of the likelihood of fraud	

TABLE 2.1 (Continued) Summary of Outcomes Used to Measure the Effectiveness of Audit Teams Performing Idea Generation, Idea Selection or Evaluation, and Plan Generation Tasks

Research Study	Idea Generation	Idea Selection/Evaluation	Plan Generation
Hoffman and Zimbelman (2009)			• Quantitative measures based on audit-planning decisions recommended by the experts • Qualitative measures to assess participants' objectives for the procedures
Lynch (2009)	• Number of relevant fraud risks identified • Number of quality fraud risks listed	• The change in fraud risk assessment before and after brainstorming	
Trotman et al. (2009)	• Number of total misstatements • Number of expert-identified misstatements • Number of expert-identified frauds • Proportion of rarely identified frauds to commonly identified frauds	• Assessments of the likelihood of fraud	
Carpenter et al. (2011)	• Number of fraud risks listed • Number of quality fraud risks listed	• Assessments of the likelihood of fraud	
Hammersley et al. (2011)	• Number of risk factors related to the seeded fraud		• Fraud detection program quality score • Audit program effectiveness score • Audit program inefficiency score

TABLE 2.1 (Continued) Summary of Outcomes Used to Measure the Effectiveness of Audit Teams Performing Idea Generation, Idea Selection or Evaluation, and Plan Generation Tasks

Research Study	Idea Generation	Idea Selection/Evaluation	Plan Generation
Chen et al. (2014)	• Number of unique valid fraud risk factors generated • Number of expert-identified frauds	• The changes in each individual auditor's fraud risk assessments	

While the number of ideas generated (Carpenter 2007; Trotman et al. 2009; Chen et al. 2014) is mainly used to capture the quantitative aspect of risk generation performance, various criteria are used to capture the qualitative aspect of ideas generated. In the fraud risk assessment setting, the rarity or uniqueness of risks generated (Trotman et al. 2009); expert-identified risks (Trotman et al. 2009); potential risks suggested by auditing standards or identified by regulators (Carpenter 2007; Lynch 2009; Carpenter et al. 2011); and seeded risks identified (Hammersley et al. 2011) have been used to measure the quality of ideas generated. Brainstorming studies in psychology examining idea generation performance have also used the originality and feasibility of ideas (Rietzschel et al. 2006); breadth of ideas (the number of idea categories generated); and depth of ideas (the number of ideas generated per category) (Nijstad et al. 2002; Dahlin et al. 2005; Kohn and Smith 2011) to measure the quality of ideas generated. Although these views on quality are broad, the choice of quality criterion depends on the ideation session's goal and what is most useful in the context of a given task (Reinig et al. 2007).

2.5.2.2 Risk Selection/Evaluation

Because the limited resources available are allocated to the most important audit areas, not all RMMs listed by practitioners will be addressed. Consequently, the second stage requires practitioners to exercise their judgment and make prioritising decisions to focus attention on risks that are more significant. ISA 315 states that "the auditor shall determine whether any of the risks identified are, in the auditor's judgment, a significant risk"[5] (IFAC 2012b, para. 27). According to McGrath's (1984) task circumplex, this task is categorised as an idea selection/evaluation task. Idea selection/evaluation tasks occur after ideas have been generated. Since it is usually not possible to implement all the generated ideas, these ideas should be evaluated and selected for further

[5] "Significant risks" are defined as "an identified and assessed risk of material misstatement that, in the auditor's judgment, requires special audit consideration" (IFAC 2012b, para. 4). To decide which risks are significant, auditors should consider whether the risks are related to fraud' recent significant economic, accounting or other developments; complexity of transactions; significant transactions with related parties; the degree of measurement uncertainty; and whether the risks involve significant transactions outside the normal course of the entity's business (IFAC 2012b, para. 28). In the GHG assurance context, auditors should also consider the likelihood of non-compliance with the provisions of laws and regulations directly affecting the content of the GHG statement, the omission of a potentially significant emission, the nature of quantification methods, the degree of complexity in determining the organisational boundary, whether Scope 3 emissions are included in the GHG statement, whether the entity makes significant estimates, and the data on which these estimates are based (IFAC 2012a, para. 34).

implementation (Reiter-Palmon et al. 2012). As shown in Table 2.1, prior research on audit judgment and decision making has suggested different outcomes that could be used to assess the quality of ideas selected/evaluated. These outcomes include consensus (Trotman and Yetton 1985), accuracy of judgments (Trotman 1985), the number of hypotheses deducted from the reviewed list (Ismail and Trotman 1995), number of correct causes selected (Green and Trotman 2003), assessment of potential misstatements' significance (Low 2004), and assessment of the likelihood of fraud (Carpenter 2007, Lynch 2009; Trotman et al. 2009, Carpenter et al. 2011, Chen et al. 2014). In addition, brainstorming studies in psychology focusing on idea selection measured the quality of selected ideas using originality (Faure 2004; Rietzschel et al. 2006; Putman and Paulus 2009) and the feasibility of the ideas selected (Reitzschel et al. 2006).

2.5.2.3 Plan Generation

After assessing RMM, the last stage is identifying appropriate audit procedures to address the assessed RMM by considering the level of assurance (IFAC 2012a, 2012b). Because this task involves generating audit plans, it can be categorised as a plan generation task under McGrath's (1984) task circumplex. Plan generation is considered the last of the cognitive processes associated with creative problem solving, in which the implementation plan is generated to address ideas selected in the previous stage (West 2002; Reiter-Palmon et al. 2012). Therefore, similar to idea generation tasks, quantity and quality measures are used to assess team effectiveness in plan generation tasks. In the auditing context, the quantity of audit plans generated is measured by the number of audit procedures generated or selected (Asare et al. 2000; Green and Trotman 2003), while the quality of audit plans generated is measured by the number of benchmark/relevant audit procedures listed (Asare and Wright 2004), the number of hypotheses tested (breadth of tests), the number of tests conducted per hypothesis (depth of tests) (Asare et al. 2000; Green and Trotman 2003), changes in the quality of audit procedures (Low 2004), participants' objectives for the procedures (Hoffman and Zimbelman 2009), and audit program quality scores (Hammersley et al. 2011).

In summary, team effectiveness is defined differently depending on the context and the task performed, and different outcomes are used to measure team effectiveness depending on the nature and the goal of the task. A wide range of outcomes for teams performing risk generation, risk selection/evaluation, and plan generation tasks have been outlined. According to the team effectiveness framework shown in Figure 2.1, these team outcomes could be affected by various team input and process variables. To provide an understanding of the effects these variables have on team effectiveness, the next sections review the literature examining the relationships between team inputs (i.e., environmental factors, task characteristics, and team composition), team process, and team effectiveness.

2.6 Team Inputs

The following sections discuss three important inputs that contribute to effectiveness based on the following frameworks: environmental factors, task characteristics, and team composition. Research findings for particular variables are discussed for each team input. Environmental factors include client characteristics and risks (i.e., size, complexity, type of company, quality of client's internal control) and client–assurer relationships (i.e., familiarity with the client, client importance). Task characteristics include task interdependence and task type. Team composition variables include team size and team diversity.

2.6.1 Environmental Factors

Environmental factors have been defined as "characteristics of the external environment in which the organisation is embedded, such as industry characteristics or turbulence" (Cohen and Bailey 1997, p. 243). The importance of broader organisation systems and the task environment are also highlighted by Kozlowski and Ilgen (2006). They suggest the importance of identifying the primary context in which teams are embedded, such as whether they are more tightly linked to the organisational system (e.g., firm policies) or task environment (e.g., patient condition, client risks). They note that the primary context is the key driver of the team task's difficulty and complexity, which in turn shapes the way teams work together. Accordingly, environmental factors are likely to vary with context. For auditing and assurance services teams, such as financial audit teams and GHG assurance teams, client characteristics and risks associated with the client's business, environment, and internal control (hereafter, client characteristics and risks) and the relationship between the client and assurer (hereafter, client–assurer relationship) can be identified as the primary context that sets team task demands and activity.

2.6.1.1 Client Characteristics and Risks

Client characteristics and risks are associated with audit effort and staffing decisions (e.g., O'Keefe et al. 1994; Hackenbrack and Knechel 1997; Gibbins and Trotman 2002; Asare et al. 2005; Knechel et al. 2009). Client characteristics are referred to as the client firm's structure (including size, geographic dispersion of operations, product, and industry), while client risks are referred to as the pre-audit likelihood that a company's financial statements are materially misstated (Asare et al. 2007). Based on the audit risk model (IFAC 2010), client risk stems from two sources: (1) inherent risk factors, such as highly complex transactions, the need for a required high degree of judgment and estimation, changes in laws and regulations, or unstable economics; and (2) control risk factors, such as the poor quality of the client's internal control and preparers.

Numerous studies have examined audit team effectiveness by focusing on the effect of client characteristics and risks. O'Keefe et al. (1994) examine the production of audit services by testing the relationship between client characteristics and risks, audit effort (labour hours), and team composition (i.e., mix of labour used: partner, manager, senior, and staff). They find that client size is positively related to audit effort and the use of audit staff. Inherent risks, such as client complexity, also increase audit effort; however, the mix of labour used does not. Further, audit effort and the proportion of partner and manager hours required when undertaking audit engagements for public companies are greater than those in private companies. O'Keefe et al. (1994) conclude that all aspects of risks related to clients may pose greater audit and litigation risks to auditors and their firms, thus prompting auditors to be more careful, improve the quality of audit documentation, and perform a more defensible audit. However, while audit effort is sensitive to client characteristics and inherent risks, O'Keefe et al. (1994) find no significant relationship between control risks—the quality of the client's internal control (degree of auditor's reliance on internal control)— and labour hours disaggregated by rank.

Hackenbrack and Knechel (1997) extend O'Keefe et al.'s (1994) findings by investigating the effect of client characteristics and risks on the number of hours charged to different audit tasks (e.g., planning, internal control evaluation, substantive testing) by different auditor ranks (e.g., partner, manger, staff). By considering both rank and task allocations, Hackenbrack and Knechel (1997) find that more labour resources are required to undertake audit engagements for large clients than for small clients and that more substantive testing and review of critical issues are required for larger clients. This study also finds that more labour resources are required to undertake audit engagements for more complex clients and find a greater demand for non-critical substantive tests than planning or client-interaction tasks. Further, audits of public companies require more labour than audits of private companies and result in a higher demand for reviews of critical substantive tests or client-interaction tasks than non-critical substantive tests or internal control tasks. When client accounting systems are highly centralised and automated, fewer auditors and less substantive testing and review of non-critical issues are required. However, no relationship is found between the degree of auditor's reliance on internal control and labour or task allocations.

Knechel et al. (2009) develop a modified audit production framework based on previous studies (O'Keefe et al. 1994; Hackenbrack and Knechel 1997) by treating client characteristics and risks as exogenous factors that are expected to affect audit production as a whole (i.e., labour inputs, audit process, evidence-gathering activities, and audit evidence). As such, the client characteristics and risks in their study are expected to affect audit production and audit efficiency. By examining the effect of client characteristics and risks on audit efficiency, the study finds that efficiency increases when the client is large and when the client extensively utilises automated systems. However, efficiency decreases when clients have subsidiaries (i.e., are more complex) and when auditors rely on internal controls.

In terms of audit review teams, Gibbins and Trotman (2002) investigate the effect of client characteristics and risks on the audit review effort and find that company size is positively related to the extent of the manager's review (i.e. pages and hours). They suggest that larger clients not only have higher audit risks but also require more time to perform the review. They also test the relationship between the quality of the file preparer and review effort and find that the review pages and hours decrease as the preparer's quality increases. Inherent risk, which is proxied by whether the client is a public company, increases the number of review pages written but not the amount of review hours spent. Gibbins and Trotman (2002) suggest that this result could be explained by the fact that public companies are closely monitored by regulators and other stakeholders, which thus increases the audit risk. Moreover, these companies are subject to a range of regulatory requirements, which then increase the amount of work and review effort. In the voluntary environmental disclosures setting, Brammer and Paveline (2008) also note that larger companies and public companies are under greater pressure to provide a higher quality of voluntary environmental disclosure because larger companies have greater economic significance and attract more public attention, which then induces greater political scrutiny and regulatory pressures.

Brazel et al. (2010) find that audit teams in the fraud brainstorming team setting tend to alter the extent of testing and acquire additional specialists for the engagement when the client size increases. Further, when the client complexity increases, audit teams are likely to adjust the nature of testing and bring in more competent staff.

These studies suggest that client characteristics and risks can influence the way teams are composed and thereby the team effectiveness. The majority of these studies focus on the audit effort and suggest that increasing audit effort leads to more effective audits (Francis 2004); therefore, client characteristics (i.e., client size), inherent risks (i.e., client complexity and whether the client is a public company), and control risks (i.e., the quality of the client's internal control) could all potentially affect MDGHGTs' effectiveness.

In addition to client characteristics and risks, the next section explores the client–assurer relationship's effect on team effectiveness; this relationship is another important environmental factor that is unique to the audit/assurance setting.

2.6.1.2 Client–Assurer Relationship

Numerous audit studies have investigated whether audit quality is affected by the client–assurer relationship, i.e., familiarity with the client (e.g., Dies Jr and Giroux 1996; Arruñada and Paz-Ares 1997; Johnson et al. 2002; Myers et al. 2003; Favere-Marchesi and Emby 2005; Carey and Simnett 2006) and the client importance (e.g., DeAngelo 1981; Reynolds and Francis 2001; Craswell et al. 2002; Chung and Kallapur 2003; Carcello and Nagy 2004; Larcker and Richardson 2004; Li 2009).

Familiarity with the client has both negative and positive effects on audit quality. On one hand, there is a perception that audit quality may be compromised as the length of the audit engagement increases (Carey and Simnett 2006). Empirical research shows that prior involvement on an audit engagement can actually impair auditor's judgments (Favere-Marchesi and Emby 2005) and audit quality (Dies Jr and Giroux 1996; Arruñada and Paz-Ares 1997; Carey and Simnett 2006). Deis Jr and Giroux (1996) investigate the effect of auditor changes on audit fees, audit effort, and audit quality and find that audit hours and audit quality are higher for first-year engagements than for the second year. A comprehensive analysis by Arruñada and Paz-Ares (1997) suggests that working with the same client for a long time could lead auditors to become too familiar with or place too much confidence in their previous work, which may result in auditors putting less effort into the current audit engagement. Carey and Simnett (2006) examine the effect of audit partner tenure and audit quality and find that a long audit partner tenure is associated with a lower propensity to issue a going-concern opinion and to just beating earnings benchmarks. These findings support the notion that audit quality decreases as the length of the audit engagement increases.

On the other hand, auditors and clients can develop client-specific expertise and knowledge with regard to the industry over time (Johnson et al. 2002; Myers et al. 2003), which then enhances audit quality. Johnson et al. (2002) investigate whether audit firm tenure—the length of the audit firm–client relationship—affects audit quality. This study find no evidence that a long audit firm tenure (nine years or more) reduces audit quality, whereas a lower audit quality is found for a short audit firm tenure (two to three years). By exploring the effect of auditor tenure on earnings quality, Myers et al. (2003) find that a longer auditor tenure is associated with less discretionary and current accruals. Further, they find that auditors limit management's ability to use accruals to increase current period earnings or to create reserves to manage future earnings. These findings suggest that familiarity with clients does not reduce audit and earnings quality.

Another stream of literature examines the effect of client importance on audit quality (e.g., DeAngelo 1981; Reynolds and Francis 2001; Craswell et al. 2002; Chung and Kallapur 2003; Carcello and Nagy 2004; Larcker and Richardson 2004; Li 2009). The seminal research by DeAngelo (1981) argues that an auditor's financial dependence on clients generates incentives for the auditor to compromise their independence to retain their clients. However, prior studies fail to find a negative association between audit quality and client importance either at the national audit firm level (e.g., Chung and Kallapur 2003; Larcker and Richardson 2004) or at the local office level (e.g., Li 2009). Reynolds and Francis (2001) find a positive association between client importance and audit quality. Their results reveal that large clients are more likely to receive a going concern audit report than smaller clients. Their findings show that the greater litigation risk posed by larger clients leads to auditors' reporting conservatism.

In summary, the previous audit literature suggested various environmental factors that are unique to the audit/assurance context. Client characteristics, including client size, inherent risks (including client complexity and type of company [public or private]), and control risks (including internal control quality), are associated with both team effectiveness and team composition. Client–assurer relationships, including familiarity with the client and client importance, are also associated with team effectiveness. However, these variables are only tested in the financial audit team setting. Given the differences in the nature of the engagements and team composition between financial audits and GHG assurance, the way in which these environmental factors affect MDGHGTs' effectiveness is unknown.

Task characteristics are another set of team input factors identified by the team effectiveness frameworks to affect team effectiveness. The research findings related to the effect of task characteristics on team effectiveness is discussed in the next section.

2.6.2 Task Characteristics

The literature identifies task characteristics as an important determinant of task distribution, authority, and interpersonal interactions within the team (Kirkman et al. 2004; Kozlowski and Ilgen 2006). In a review of factors correlated with team effectiveness, Kozlowski and Ilgen (2006) point out that task characteristics shape the workflow design and the extent of coordination needed to successfully deal with the task. This is achieved by determining how the team members' knowledge and expertise should be utilised and by prioritising team members' action. To test the effect of task characteristics on team effectiveness, researchers focus on two aspects: the degree of task coordination within the team, which is known as "task interdependence" (Kiggundu 1981; Campion et al. 1993; Stewart and Barrick 2000; Van Der Vegt et al. 2000), and task type (McGrath 1984; Jehn et al. 1999; Stewart and Barrick 2000). The research findings associated with these two task characteristics are discussed in detail in the following subsections.

2.6.2.1 Task Interdependence

The first aspect of task characteristics discussed in this section is task interdependence. Task interdependence refers to the degree to which group members must rely on one another to perform a task effectively given the design of their jobs (Kiggundu 1981; Brass 1985; Campion et al. 1993). The degree of task interdependence typically increases as the work becomes more difficult and personnel require greater assistance from others to perform their jobs (Van Der Vegt et al. 2000).

Empirical research on task interdependence shows a positive relationship between task interdependence and team effectiveness (e.g., Wageman 1995; Campion et al. 1996; Stewart and Barrick 2000; Sparrow et al. 2001). Wageman (1995) examines the effect of task interdependence on team effectiveness in an experiment involving 150 teams of technicians. Wageman finds that task interdependence facilitates cooperation, improves member satisfaction, and results in group processes of a higher quality. Wageman also discovers a U-shaped interdependence-performance relationship, which suggests that when interdependence among team members is high or low, the team performance is better than when such interdependence is moderate. Using teams in a financial services firm, Campion et al. (1996) find that teams work more effectively together when they have better team processes, more motivational job design, higher task interdependence, and higher team diversity. Stewart and Barrick (2000) examine the relationship between task interdependence and the performance of production teams and find that team members collectively work together when task interdependence is high to complete a task while sharing information and resources. In contrast, team members tend to operate more independently in tasks requiring low interdependence, thereby reducing the need for coordination and collaboration among members. In a meta-analysis, Humphrey et al. (2007) find that task interdependence is positively related to behavioural outcomes, supervisor satisfaction, organisational commitment, job involvement, and internal work motivation.

Overall, these findings suggest that task interdependence is associated with team effectiveness. In particular, team members are likely to work more effectively together when the task is highly interdependent (as opposed to when it is moderately interdependent). Despite the important role of task interdependence, another aspect of task characteristics—task type—is also associated with team effectiveness. Different types of tasks and prior research examining the relationship between task type and team effectiveness are discussed in the next section.

2.6.2.2 Task Type

As previously mentioned, another aspect of task characteristics, i.e., task type, could affect team effectiveness. Task type refers to “those aspects of interaction that relate directly to a group’s work on its task” (Hackman 1987, p. 321). Team tasks have been classified in many different ways (McGrath 1984); however, previous literature has suggested that McGrath’s (1984) sophisticated classification scheme is one of the most practical approaches for classifying the tasks performed by real teams operating in organisations (Goodman 1986; Stewart and Barrick 2000). According to McGrath’s (1984) task circumplex (Section 2.5.1), team tasks can be classified into four main categories: generating plans and ideas (e.g., brainstorming tasks); solving problems and deciding issues (e.g., decision-making tasks); resolving conflicts of viewpoint and interest (e.g., negotiation tasks); and executing work (e.g., performance tasks). Based on the task circumplex, these tasks also differ in two dimensions: the extent to which they are either conceptual or behavioural and the extent to which they involve either cooperation or conflict. For example, idea generation tasks are classified as conceptual tasks (as opposed to behavioural tasks), and such tasks involve cooperation (as opposed to conflict).

To test the relationship between task type and team performance, Stewart and Barrick (2000) examine the tasks performed by production teams utilising McGrath's classification scheme. Although that study considered work execution a primary task for production teams, some teams spent a larger proportion of their time performing conceptual tasks (i.e., planning, deciding, and negotiating). Thus, Stewart and Barrick (2000) identify the team tasks each team engaged in based on the relative amounts of time each team spent on behavioural tasks (i.e., execution tasks) and conceptual tasks. According to this study, behavioural tasks require little interaction among team members and thus face less interpersonal difficulties. On the other hand, conceptual tasks involve a high level of interactions and different team processes, such as discussing and negotiating. The study's results show that the curvilinear relationship between task interdependence and team effectiveness may exist for teams primarily engaged in conceptual tasks but not in behavioural tasks. These findings indicate that the relationship between task interdependence and team performance is significantly moderated by the type of task.

While the McGrath (1984) approach to classifying tasks has been adopted to examine the effect of task type on team effectiveness, the teams used in the previous study are production teams likely to have similar educational/professional backgrounds (i.e., teams of engineers/technicians). To examine the effect of task type on MDTs' effectiveness, studies on work team diversity classify tasks based on their complexity (e.g., Jehn et al. 1999; Bowers et al. 2000). Jehn et al. (1999) examine the effect of informational diversity (i.e., differences in knowledge and perspectives resulting from differences in educational backgrounds and work experience), task type, and task interdependence on team performance and find that task type and task interdependence moderate the relationships between informational diversity and team efficiency. In particular, they find that informational diversity is more likely to enhance team performance when tasks are more complex (i.e., non-routine), whereas it has little effect on performance when tasks are less complex (i.e., highly routine). They suggest that when MDTs work on complex tasks, different knowledge bases and perspectives may introduce some disagreement on the tasks. This task conflict could increase the amount and degree of discussion occurring among team members, increase the team effort to carefully process task-relevant information, and enhance team performance (Jehn 1995; Jehn et al. 1999).

In financial audit teams, complex tasks also affect audit judgment and decision making. Simnett and Trotman (1989) is the first study to examine the effect of task complexity on information selection and information processing by focusing on one particular aspect of complexity, i.e., environmental predictability. They find that higher task complexity results in a lower judgment performance, although it does not interact with information selection and processing. Simnett (1996) extends this study by crossing information selection and information processing with another aspect of complexity, i.e., information load. This study finds that information processing limits auditors' predictive accuracy when the task is more complex (i.e., higher information load). Therefore, the findings from both studies suggest that task complexity negatively affects auditors' judgment performance.

In summary, task type is associated with team performance, regardless of whether the tasks are classified based on McGrath's (1984) approach or based on the level of complexity. Tasks requiring more interaction and coordination between team members (i.e., conceptual tasks) are more likely to enhance team performance when the task interdependence is high. Tasks that are more complex enhance the performance of educationally diverse teams (MDTs) and inhibit the performance of educationally homogeneous teams (i.e., financial audit teams).

The research discussed above on task characteristics used real work teams in different organisations and generally employed production, technician, financial services, and financial audit teams. Very few studies (e.g., Jehn et al. 1999) use MDTs with highly distinctive areas of expertise to test the effect of task characteristics on team effectiveness. Given that team composition makes MDTs different from teams in the previous studies, the next section focuses on how differences in team composition can affect team effectiveness.

2.6.3 Team Composition

Based on the team effectiveness frameworks, team composition is an important team inputs that could affect team processes and team effectiveness. Team composition is typically examined in terms of size and diversity (Guzzo and Dickson 1996; Stewart 2006). Both the size and diversity of the team can positively or negatively affect team effectiveness (see Cohen and Bailey 1997 and van Knippenberg and Schippers 2007 for reviews). The following subsections review findings from previous studies examining the effect of team size and team diversity on team effectiveness.

2.6.3.1 Team Size

Past research suggests that team size (as measured by the number of members in the team) can be both beneficial and detrimental to team processes and effectiveness. On one hand, larger teams could be more effective because of the larger pool of resources, including the different knowledge, skills, and perspectives members bring to the team to solve team tasks (Bantel and Jackson 1989; Wiersema and Bantel 1992). West et al. (2003) investigate the relationship between team size and innovation in MDTs using primary health care teams, community mental health care teams, and breast cancer care teams. The results show that, across the three samples, larger teams are more innovative than smaller teams (with the team average size ranging from 9 to 21 members). This study concludes that larger teams tend to process a broader range of information and perspectives, which leads them to generate more creative ideas. Fay et al. (2006) also investigate the effect of team size on MDTs. Using two types of multidisciplinary health care teams (with the average team size ranging from 13 to 20 members), this study finds that team size is positively related to both the quantity and quality of ideas generated by the teams. These relationships are strengthened by the quality of the team processes, i.e., team size becomes positively related to the quantity and quality of ideas generated when the quality of the team processes is high. Consistent with previous studies, Fay et al. (2006) conclude that the effect of size found in their study could be contributed to the broader set of knowledge, skills, and perspectives that can only be fully utilised when team processes are sufficient to facilitate information sharing and integration between MDT members.

On the other hand, larger teams lead to communication (Smith et al. 1994), participation (Poulton and West 1999), and coordination challenges (LePine et al. 2008) and are less likely to work effectively together (Curral et al. 2001). Smith et al. (1994) investigate the role of social integration and communication in multidisciplinary top management teams and the effect of team size on team processes and team outcomes. This study finds a negative relationship between team size (five members on average) and communication between team members. Team size is also negatively associated with social integration through communication between team members, which in turn increases team effectiveness. This study suggests that larger teams may introduce more distance between team members and thus hinder team interaction. Poulton and West (1999) examine the factors associated with effectiveness in primary health care teams and find that members of larger MDTs perceive participation levels to be lower than do members of smaller MDTs. However, this study finds no significant relationship between team size and team effectiveness. Curral et al. (2001) study the relationship between two team inputs—task type and team size—and team processes using MDTs involved in product or service innovation. The findings reveal that larger teams experience poorer team processes than smaller teams. However, when considering the task type, team size negatively affects team processes when tasks require high levels of innovation (are more complex) rather than when tasks require low levels of innovation (are less complex). Curral et al. (2001) conclude that larger teams may have difficulty reaching consensus, which thereby impedes information sharing and team interaction, specifically when the task is complex. Lepine et al. (2008) conduct a meta-analysis of the effect of teamwork processes and relationships on team effectiveness and find that team size moderates the relationship between teamwork processes and team effectiveness. These findings indicate that team outcomes hinge more on effective team work processes when teams are larger compared to when teams are smaller.

Aside from the negative and positive effect of team size on team processes and team effectiveness, some studies find that team size is neither beneficial nor detrimental to team effectiveness. Bantel and Jackson (1989) and Wiersema and Bantel (1992) find no relationship between team size and the performance of multidisciplinary top management teams (i.e., teams involved in innovation and change in diversification strategy).

In summary, previous studies suggest that team size is related to both team processes and team effectiveness. While results vary on the effect of team size on team effectiveness, a larger team consistently hinders team processes. However, size is only one dimension of team composition. To fully understand the effect of team composition on team processes and team effectiveness, another dimension of team composition—team diversity—should be considered. Thus, the next section discusses findings from social psychology literature on work team diversity to explore the effect of team diversity on team processes and team effectiveness.

2.6.3.2 Team Diversity

Diversity is usually referred to as differences between individuals on any aspect that may lead to different perceptions (Triandis et al. 1994; Williams and O'Reilly 1998). To understand how diversity affects team processes and performance, research on work team diversity mainly categorises diversity into two types: social category diversity (e.g., age, gender, ethnicity and other social identity) and informational diversity (e.g., functional and education backgrounds) (Williams and O’Reilly 1998; van Knippenberg and Schippers 2007).

Among the different types of diversity, informational diversity consistently shows a positive effect on team performance (see van Knippenberg and Schippers 2007 and Jackson and Joshi 2011 for reviews). Informational diversity is also referred to as cognitive diversity. Cognitive diversity refers to the differences in team members’ knowledge bases, perspectives, attitudes, values, and beliefs (e.g., Kilduff et al. 2000 and Milliken et al. 2003). These differences could result from members’ different educational backgrounds because the curriculum of study not only indicates one’s personality and cognitive style (Holland 1973) but also shapes the way that person thinks or believes (Dahlin et al. 2005). An individual’s cognitive style could also be influenced by the functional background, such as work experience (Milliken et al.

2003). However, because differences in the functional background could prompt an in-group/out-group identification, which is referred to as social categorisation diversity, the differences in educational background are suggested to be a “purer indicator” of cognitive diversity (Dahlin et al. 2005, p. 1108). Therefore, educational diversity is the most salient and important source of creative thinking and reasoning (Nijstad and Paulus 2003).

Cognitive diversity is expected to lead to a broader set of knowledge and perspectives in a given task, which explains why team compositions incorporating different educational background are being increasingly adopted by organisations facing complex tasks (van Knippenberg and Schippers 2007). However, empirical research has found mixed results on the benefits of such diverse teams (e.g., Williams and O'Reilly 1998; van Knippenberg and Schippers 2007).

On one hand, previous studies show that cognitive diversity leads to improved decision making and creative problem solving (Bantel and Jackson 1989; Wiersema and Bantel 1992; De Dreu and West 2001). Bantel and Jackson (1989) examine the relationship between innovations in banking and the various characteristics of top management teams, including educational and functional backgrounds. The results of this study show that educational and functional diversity are positively related to innovation. Similarly, Wiersema and Bantel (1992) find that higher educational diversity in top management teams is linked to a team propensity to change corporate strategy. De Dreu and West (2001) also find that cognitively diverse team members lead the team to be more innovative, but only when they participate extensively in the team decision making. They conclude that participation among cognitively diverse members facilitates the exchange of different information and perspectives, which then stimulates creativity and divergent thoughts.

The positive effect of cognitive diversity on team performance can also be attributed to task conflict among team members. A number of studies find that cognitive diversity introduces task conflict, i.e., disagreements among team members on task-related issues (Jehn 1995; Jehn et al. 1999; Pelled et al. 1999). Jehn (1995) examines the benefits and detriments of conflict in work teams and management teams and finds that task conflicts negatively affect team members’ satisfaction and intention to stay on the team as well as the affection of other team members. However, the results reveal that task conflicts are beneficial to team functioning when teams perform non-routine tasks. The interviews and observations of teams in this study show that effective teams that perform non-routine tasks experience a high level of disagreement and norms promoting open discussion on task-relevant issues. Both task conflict and open discussions about tasks require teams to thoroughly process and critically evaluate all information and perspectives, thereby enhancing team performance. Jehn et al. (1999) conduct a survey to investigate the relationship between diversity, conflict, and team performance. This study finds that cognitive diversity is positively associated with task conflict in work teams. The results of this study also show that task conflict mediates the positive relationship between cognitive diversity and team performance. Further, cognitive diversity enhances team performance when tasks are complex. Pelled et al. (1999) also find that cognitive diversity is positively related to task conflict and that task conflict is positively related to team performance. However, they do not find any relationship between cognitive diversity and team performance. Thus, task conflict does not found to mediate the relationship between cognitive diversity and team performance. Through a meta-analysis of the associations between task conflict, team performance, and team member satisfaction, De Dreu and Weingart (2003) find that task conflict is negatively correlated with team performance and team member satisfaction, which is inconsistent with previous studies' findings. This negative relationship is stronger when teams perform highly complex tasks compared to less complex tasks. The conflicting findings in the previous research on task conflict indicate that task conflict is not the only key driver for the positive relation between cognitive diversity and the performance of cognitively diverse teams, such as MDTs.

Although cognitive diversity among team members is expected to improve team effectiveness, it can also be deleterious to team performance (van Knippenberg and Schippers 2007). Team members with different professional backgrounds often use different language or terminology that are only understood by people in the same profession or that have different meanings in other fields, which leads to misunderstandings and misinterpretations of the communicated information (van Someren et al. 1998; Van Asselt 2000). This argument is supported by a number of studies on MDTs, which find that knowledge-sharing difficulties are rooted in differences in team members' language (e.g., Carlile 2004; Sheehan et al. 2007). This is exacerbated by the fact that team members' different functional and professional backgrounds may elicit social categorisation processes, such as out-group and in-group identification (van Knippenberg and Schippers 2007). This out-group/in-group bias can lead to less communication (Bhappu et al. 1997) and less cooperation (Chatman and Flynn 2001; Randel and Jaussi 2003) between the subgroups. Therefore, if the inclusion of members with different expertise and educational backgrounds is maximised at the expense of the shared understanding of the team task, the team's ability to work effectively together is threatened. The question that arises from these findings is how diverse teams should be to truly benefit from the different expertise offered by MDT members.

In attempt to explain the relationship between knowledge diversity and the team integration process, West (2000) proposes that this relationship can be described using an inverted-U shape. That is, team members tend to follow other members' views when diversity is too low, and they do not have enough overlapping mental representation to integrate, communicate, and coordinate with other team members when diversity is too high. Dunbar (1997) examines teams of scientists with different areas of expertise; his findings support this relationship by showing that teams of experts from slightly different areas outperform teams of experts from the same area in terms of problem solving. Together, both studies suggest that some degree of diversity in knowledge and expertise could be advantageous. However, too much diversity in knowledge and expertise may create misunderstandings and thus inhibit information sharing and integration between team members.

Although previous studies have suggested that size and diversity could have affect team effectiveness, the studies focus on top management teams, project teams, or work teams with slightly different areas of expertise. Very limited research has been conducted on the effect of different team compositions on the performance of MDTs in which members are highly diverse in their educational backgrounds (e.g., Curral et al. 2001; Randel and Jaussi 2003). However, the research findings show that the true benefit of a larger pool of resources lies in the integration of different perspectives and how well team members deal with communication and coordination problems arising from the team's diverse nature. This emphasises the important role of team processes. Therefore, the next section discusses various team processes suggested by studies on both psychology and auditing to enhance team effectiveness.

2.7 Team Processes

The majority of team effectiveness frameworks in psychology and management (Cohen and Bailey 1997; Ilgen et al. 2005; Mathieu et al. 2008) suggest that team processes are one of the most important factors related to team effectiveness. Team processes are defined as mechanisms to integrate different knowledge and expertise possessed by team members and to coordinate effort to resolve task demands (Kozlowski and Ilgen 2006). The importance of team processes becomes even more salient for teams comprising members with diverse educational and expertise backgrounds given that the true benefit of MDTs not only lies in a larger pool of knowledge and skill-sets but also the integration of diverse information and the exchange of different perspectives (van Knippenberg et al. 2004). Therefore, effective team processes provide a means of achieving the potential advantages of diverse expertise.

Given that team processes vary depending on the type of team format used (Kerr and Tindale 2004), previous psychology literature has investigated various forms of teams, including decision making by individuals, nominal teams, and interacting teams (see Kerr and Tindale 2004 for a review). These team formats have also been examined in the auditing context (see Rich et al. 1997b and Nelson and Tan 2005 for reviews) because many audit decisions are made using a team-based approach. The review process, which is another form of the decision-making process and is an essential part of the audit (Trotman 1985), has also been examined (e.g., Trotman 1985; Trotman and Yetton 1985; Ismail and Trotman 1995). Although these team formats are normally tested in teams comprising members with similar educational backgrounds (i.e., students from the same class or teams of auditors with an accounting background), MDGHGTs could adopt one of several team formats to maximise the advantages of educational diversity. In particular, this dissertation is interested in three types of team format, i.e., nominal, interacting, and review teams. Therefore, this section focuses exclusively on studies examining the effect of these three team formats on team performance, particularly on idea generation and idea selection performance.

Table 2.2 provides a summary of experimental audit research investigating the effect of different types of team processes on audit effectiveness. This table shows the various forms of team interactions examined by audit researchers, including working alone as an individual, working in teams without any interaction (nominal teams), working in teams with verbal interaction (interacting, brainstorming, and review with discussions teams), and working in teams with non-verbal interactions (electronic brainstorming and review teams). In addition, team research in auditing varies in terms of the participants used, including audit students, audit practitioners in hierarchical teams, and audit practitioners in non-hierarchical teams. Because this dissertation focuses on three particular team formats, that is, nominal, interacting, and review teams, study findings examining and comparing the performances of these team formats are reviewed and discussed further in the following subsections.

2.7.1 Nominal Teams

In order to fully understand the effect of different team formats on team performance, previous studies (e.g., Steiner 1972; Hill 1982; Diehl and Stroebe 1987, 1991) employ at least one baseline model to help predict the team's outcome under a control condition. In a comprehensive review on group performance and decision making, Kerr and Tindale (2004) conclude that, among many baseline models, Steiner's (1972) potential productivity baseline is the most frequently adopted. By assuming the optimal coordination and integration of group members' knowledge and ability, this baseline predicts the ideal level of group performance, which helps determine if groups perform better or worse than the anticipated level (Kerr and Tindale 2004). Accordingly, a number of group decision making researchers in both psychology (e.g., Diehl and Stroebe 1987, 1991) and auditing (e.g., Trotman et al. 1983; Trotman 1985; Carpenter 2007; Hoffman and Zimbelman 2009; Carpenter et al. 2011) have adopted a statistical pooling team, known as a "nominal team",[6] to establish this baseline performance.

[6] The nominal team process differs from the Nominal Group Technique (NGT). NGT is a structured decision-making process developed by Van de Ven and Delbeco (1971) in which group members work separately to list ideas in the early stage but share ideas through some verbal communication in later stages. NGT is different from the nominal team process in that the nominal team process does not allow any communication between team members.

TABLE 2.2 Summary of Previous Audit Judgment and Decision Making Studies Examining the Effects of Different Team Processes on Audit Team Effectiveness

Participants	Individuals working independently	Individuals working in nominal teams	Teams interacting or brainstorming in face-to-face settings	Teams interacting or brainstorming in electronic settings	Teams involved in a hierarchical review process (with and without interaction)	Teams involved in a hierarchical review process (electronic review)	Teams involved in a non-hierarchical review process
Financial audit practitioners	Trotman and Yetton (1985)	Trotman and Yetton (1985)	Trotman and Yetton (1985)		Trotman and Yetton (1985)		
	Trotman (1985)	Trotman (1985)	Trotman (1985)		Trotman (1985)		
	Ismail and Trotman (1995)				Ismail and Trotman (1995)		
		Brazel et al. (2004)			Brazel et al. (2004)	Brazel et al. (2004)	
	Carpenter (2007)	Carpenter (2007)	Carpenter (2007)				
					Agoglia et al. (2009)	Agoglia et al. (2009)	
	Hoffman and Zimbelman (2009)		Hoffman and Zimbelman (2009)				
Auditing students		Lynch et al. (2009)	Lynch et al. (2009)	Lynch et al. (2009)			
			Trotman et al. (2009)				
Financial audit practitioners					Payne et al. (2010)		
		Chen et al. (2014)		Chen et al. (2014)			
Internal audit practitioners	Carpenter et al. (2011)	Carpenter et al. (2011)	Carpenter et al. (2011)				
Multidisciplinary assurance practitioners		**Present Study**	**Present Study**				**Present Study**

A nominal team can be formed by pooling the efforts of members who work separately without any communication (e.g., to generate ideas). All members' ideas are then combined, and redundant ideas are eliminated to ensure that each idea is counted only once. Therefore, the number of ideas generated by nominal teams represents the expected level of productivity when team interaction has no effect on performance (Diehl and Strobe 1987).

When comparing the performances of nominal teams and other team formats, one factor that is usually of interest is whether other factors, such as interactions between team members, facilitate or inhibit team performance. If team interactions facilitate team performance, the team should exceed the potential productivity baseline, which means they exhibit process gains. However, if such interactions inhibit team performance, they should fall behind the baseline and exhibit process losses (Steiner 1972; Diehl and Strobe 1987). Since nominal teams do not exhibit any process gains or losses, the ability of other team formats to reach the same level of productivity as nominal teams hinges on the balance between these process gains or losses (Dennis and Valacich 1993).

Because nominal teams are treated as a baseline, nominal teams' performance is usually compared with the performances of other team formats, specifically interacting teams. Therefore, the next section discusses research findings related to nominal teams in conjunction with research findings related to interacting teams.

2.7.2 Interacting Teams

Teams are increasingly being used by organisations to perform various tasks required for the complex problems faced (Paulus 2000). In conducting such tasks, teams need to interact to exchange ideas and perspectives, especially when they are involved in idea generation and idea selection. The following discussion on the effect of interaction on team performance will first consider the type of task performed and consider idea generation performance and idea selection performance.

2.7.2.1 The Effect of Team Interaction on Idea Generation Performance

A cognitive process, such as exchanging ideas, involves two or more individuals sharing their unique knowledge and perspectives based on their experiences, skill-sets, and educational backgrounds (Paulus 2008). Therefore, people usually believe that one method of generating better ideas is to stimulate creativity through team interaction (Paulus et al. 1993). The assumption that interacting teams are more effective than non-interacting teams in terms of idea generation has led to a substantial amount of research comparing the performance of different types of interacting teams to nominal teams. Osborn (1957), a proponent of the brainstorming interacting group technique, suggests that the quantity and quality of ideas generated by teams can be increased through the stimulation and integration of ideas. He posits that communication among brainstorming group members can enhance productivity in idea generation tasks by allowing team members to build on others' ideas. Although Osborn (1957) suggests that brainstorming groups can outperform an equal number of individuals who work independently, empirical research in psychology has consistently found conflicting results, i.e., brainstorming groups generate fewer and lower quality ideas than nominal groups (e.g., Hill 1982; Diehl and Stroebe 1987; Mullen et al. 1991; Argote and Kane 2003; Dennis et al. 1999; Rietzschel et al. 2006).

Diehl and Stroebe (1987) review 22 experiments conducted to test Osborn's claim. They show that 80 percent of these studies find that nominal groups generate more ideas than interacting groups, while the other 20 percent of the studies report no difference between nominal and interacting groups. In an attempt to explain this failure of interacting groups to outperform nominal groups, Deihl and Stroebe (1987) investigate three possible explanations for process losses: production blocking, evaluation apprehension, and free riding.[7] Among the three potential causes, they find that production blocking explains most of the process loss in interacting groups. Therefore, these results suggest that verbal communication among members inhibits, rather than facilitates, idea generation in groups.

[7] Production blocking occurs because group members have to take turns verbalising their ideas and thus have to listen to others' ideas while thinking, which could interfere with their own thoughts. Free riding occurs because group members rely on others to complete tasks because of the perception that their inputs are unidentifiable or dispensable. Evaluation apprehension occurs because group members are afraid of being evaluated by other group members and thus withhold their ideas.

In the audit setting, auditors usually communicate with their team members in various stages of the audit engagement (Nelson and Tan 2005). Interest in the use of interacting teams has increased because of the requirements in auditing standard SAS No. 99 (AICPA 2002) to perform a "brainstorming session" on every engagement to detect potential material misstatements due to fraud. The rationale behind this requirement is to increase audit effectiveness through collaboration since most auditors never confront material fraud during their careers and thus may be less effective when assessing fraud risk individually (Beasley and Jenkins 2003). Accordingly, brainstorming groups have received increasing attention from auditing researchers (Carpenter 2007; Hoffman and Zimbelman 2009; Lynch et al. 2009; Trotman et al. 2009; Carpenter et al. 2011; Chen et al. 2014).

Previous studies examine two aspects of brainstorming teams that are of interest: comparisons of performance in such teams with the performance of individuals and/or nominal teams (e.g., Carpenter 2007; Lynch 2009; Hoffman and Zimbelman 2009; Carpenter et al. 2011) and comparisons across different brainstorming techniques, such as brainstorming with guidelines and brainstorming without guidelines (Trotman et al. 2009), face-to-face discussions and electronic interactive discussion (Lynch et al. 2009; Chen et al. 2014), and strategic reasoning and brainstorming (Hoffman and Zimbelman 2009).

Given that idea generation performance is central to the success of fraud risk assessment, the quantity and quality of fraud risks listed have been widely used to measure audit teams' effectiveness. Carpenter (2007) is the first to examine brainstorming teams in the fraud assessment setting by comparing the performances of nominal teams and face-to-face brainstorming teams. Audit professionals from the Big Four firms were grouped into three-person teams comprising a staff auditor, a senior, and a manager. In the first phase, all participants were asked to assess the risk of fraud individually. After collecting the individuals' answers, the team members began brainstorming sessions. The study finds evidence that brainstorming teams generate fewer risks than the nominal teams, but the quality (i.e., the number of fraud ideas that are a clearly identified fraud specific to the case) of those risks are higher than those generated by the nominal teams. These results suggest that although the interactions between hierarchical team members in this study lead to some losses in the number of ideas generated, it is effective in screening out non-quality ideas.

In the electronic brainstorming setting, Lynch et al. (2009) examine the effect of computer-mediated communication on auditors' fraud assessment performance by comparing three different team formats: nominal, face-to-face brainstorming, and electronic brainstorming. Although they find that electronic brainstorming teams are superior to face-to-face brainstorming teams, they still do not outperform nominal teams. The findings indicate that, in terms of the quality of ideas generated (i.e., higher number of relevant fraud risks generated), nominal teams outperform face-to-face brainstorming teams, while no differences are found between nominal and electronic brainstorming teams.

While Lynch et al. (2009) use undergraduate student subjects, Chen et al. (2014) examine hierarchical audit teams' performance in relation to nominal and interacting electronic brainstorming. This study finds that nominal teams generate more unique fraud risk factors, more fraud hypotheses, and fraud hypotheses of a higher quality (i.e., number of unique expert-identified fraud hypotheses) than interacting teams. The findings from this study show that interacting teams cannot outperform nominal teams due to the social loafing of less experienced auditors in the team.

Carpenter et al. (2011) extend the earlier research by investigating whether the process gains (i.e., quality of ideas) found for hierarchical external audit teams in Carpenter (2007) could be generalised to internal audit teams, which are much less hierarchical in nature. They find that nominal teams are superior to brainstorming teams in terms of the quantity of risks generated (i.e., number of risks listed) but not for the quality (i.e., number of risks that matched the frauds identified as actual frauds by the SEC). The brainstorming teams generate much higher number of quality risks than the nominal teams, which indicates that the interaction of team members result in process losses in the quantity of ideas and process gains in the quality of ideas even when the team members have a similar number of years of working experience.

The findings from previous audit brainstorming studies indicate that nominal teams are superior to interacting teams regardless of whether the teams are of a high hierarchical nature. However, it is important to note that although these teams have different hierarchical natures, they have similar educational backgrounds, i.e., they typically have an accounting/auditing background.

In the internal control evaluation context, Trotman et al. (1983) compare interacting teams to nominal teams and find that interacting teams exhibit significantly less consensus than nominal teams. They conclude that the interacting teams' failure to outperform nominal groups may stem from the use of audit students with a limited range of knowledge and the use of routine tasks, which makes it difficult for the participants in this study to differentiate their relative expertise. To address the issues raised by Trotman et al. (1983), Trotman (1985) conducts an experiment using senior auditors and a more complex task to allow team members to demonstrate their expertise. This study uses participants who had already gained some knowledge about their team members' relative expertise through working experiences, utilises a non-structured task, and allows participants to show their calculations to other team members. Based on these factors, the study expects that the team members' ability to recognise differential expertise would be higher than in the previous study. Consistent with the expectations, the interacting teams outperform the nominal teams in terms of the accuracy of judgments, which are measured by the absolute difference between an auditor's estimate of the expected dollar error in the system and the mean of the distribution of simulated errors. Trotman (1985) and Libby et al. (1987) suggest that interacting teams will outperform nominal teams if enough variation occurs in the team members' performance and if they are able to recognise the differences in their expertise. Further, they suggest that the team task must be complex enough to detect differences in expertise. In line with these suggestions, Nijstad and De Dreu (2002) note that the majority of brainstorming studies find that nominal groups are superior to interacting groups because the participants are typically students with similar educational backgrounds. They suggest that these participants have a low cognitive diversity and thus are less likely to bring different information and perspectives to the given task and to generate different ideas when they interact.

Among the extensive studies on group brainstorming, very few studies have tested the effect of cognitive diversity on group idea generation. Stroebe and Diehl (1994) test this effect and find that interacting groups with heterogeneous members (in terms of dominant associations regarding environmental concerns) generate almost the same number of ideas as nominal groups. They suggest that a broad range of ideas and perspectives shared within cognitively diverse groups stimulate the team members' creativity, which in turn outweighs the productivity losses, such as production blocking, usually observed in interacting groups. These findings suggest that interacting teams could outperform nominal teams when the teams are cognitively diverse.

In addition to the quantity of ideas generated, the benefits of interaction between cognitively diverse members are likely to be captured in the qualitative aspects of their performance, particularly how these teams utilise their diverse knowledge and perspectives. Cognitively diverse teams, such as MDTs, have advantages over teams comprising members with similar educational backgrounds in terms of the breadth of ideas generated because of their opportunity to be stimulated by a broader range of knowledge, skills, and perspectives (Stroebe and Diehl 1994; Nijstad et al. 2002; Dahlin et al. 2005). Stroebe and Diehl (1994) find that interacting groups with heterogeneous members (in terms of dominant associations regarding environmental concerns) generate almost the same number of categories (breadth) as nominal groups. They suggest that the broad range of ideas and perspectives shared within cognitively diverse groups triggers group members to explore different categories of ideas, which thus outweighs the productivity losses usually observed in interacting groups.

Nijstad et al. (2002) also find that groups exposed to semantically heterogeneous ideas generate more categories of ideas (more breadth), while groups exposed to semantically homogeneous ideas generate more new ideas within the same category (more depth). They conclude that diverse cognitive stimulation increases the breadth of ideas generated, while homogeneous cognitive stimulation increases the depth of ideas generated.

Dahlin et al. (2005) investigate information use in MDTs and find that teams use and generate a wider range of information and analyse ideas in greater depth when the team is more educationally diverse. They emphasise that exploring the breadth and depth of ideas generated is important because increasing the breadth of ideas allows many possible alternatives to be analysed, while increasing the depth of ideas allows important issues to be focused upon and thereby more completely explored. This study suggests that MDTs possession of a broader set of knowledge and frameworks allows team members to analyse familiar information in both a breadth and depth approach while leaving more time for members to deeply process unfamiliar information; however, that these benefits only hold up to a certain point—once the level of educational diversity in the team is too high, the breadth and depth of information use decreases. These findings suggest that too much diversity in education makes it difficult for team members to understand each other and thus inhibits information sharing, exploring, and integrating (West 2002; Dahlin et al. 2005).

In summary, previous studies in both psychology and auditing find that nominal teams usually outperform interacting teams in idea generation tasks. This finding could be explained by the lack of cognitive diversity among participants in the previous brainstorming studies. Therefore, it is interesting to examine the effect of interaction in cognitively diverse teams on idea generation performance, not only in terms of the quantity of ideas generated but also how these teams utilise their diverse knowledge and perspectives (i.e., the breadth and depth of ideas generated).

2.7.2.2 The Effect of Team Interaction on Idea Selection Performance

While the majority of brainstorming studies focus heavily on the number of ideas generated (West 2002; Paulus 2008), it is usually not possible or practicable to implement all these ideas, even if most of them are good ideas. Thus, these ideas should be evaluated and selected for further implementation (Reiter-Palmon et al. 2012).

Idea selection is a part of the decision-making process (Mumford et al. 1991) and may be performed by groups or individuals that generate ideas or by other groups or individuals not involved in the idea generation process (Paulus 2008). Groups may be better at selecting ideas than individuals because they have a larger pool of knowledge and perspectives, which in turn helps screen out inappropriate alternatives (Laughlin and Hollingshead 1995; Paulus 2008). However, research has shown that neither individuals nor groups perform idea selection well (e.g., Rietzschel et al. 2006; Rietzschel et al. 2010).

Faure (2004) finds that nominal groups generate a larger quantity and a higher quality of ideas than interacting groups. However, these two groups do not differ in the quality of risks they select. Rietzschel et al. (2006) find that nominal groups generate better ideas than interacting groups, both in terms of the number of quality ideas and the originality of ideas. However, when it comes to idea selection, no difference is found between nominal and interacting groups with regard to the originality and feasibility of the ideas selected. Putman and Paulus (2009) also find that nominal groups generate a larger number of ideas and more original ideas than interacting groups. Interestingly, groups who generate ideas individually select more original ideas than groups who generate ideas interactively. The findings from these studies indicate that interactions between group members do not enhance idea selection performance. Rietzschel et al. (2010) note that the most interesting finding that these studies discovered is that "participants' idea selection is in fact not better than chance" (p.49) because participants in these studies do not select ideas of a higher quality when they work as a group than the ideas they generate individually.

In contrast to the previous studies presented above, previous audit studies find that interactions between auditors lead to higher levels of judgment quality in idea evaluation/selection tasks. For example, Carpenter (2007) and Lynch (2009) find that when fraud is present, the risk assessments given by audit teams are much higher after they have the opportunity to interact in a brainstorming session than when they are asked to assess risks individually (i.e., nominal groups). Carpenter et al. (2011) also report that nominal internal audit teams respond differently between qualitative and quantitative risk assessment scales, while this response mode bias is not present in brainstorming teams. They suggest that brainstorming itself does not lead to this process gain but rather that the interactions between members in brainstorming teams lead to this gain.

Although a number of audit research studies emphasise that selecting good ideas is very important in diagnostic tasks, such as analytical procedures and risk assessment, auditors have difficulties performing these tasks (e.g., Bedard and Biggs 1991; Hirst and Koonce 1996; Asare and Wright 1997; Green and Trotman 2003; Moreno et al. 2007; Hammersley et al. 2011; Luippold and Kida 2012; Pike et al. 2013). For example, Green and Trotman (2003) examine the processes underlying auditors' success in analytical procedures by considering both hypothesis generation and evaluation. The results show that unsuccessful auditors have difficulties selecting the correct cause, even though they generate the correct cause in their hypothesised causes.

However, the teams used in the previous brainstorming literature in auditing are homogenous in nature. Few research studies have examined the idea selection performance of assurance teams comprising practitioners from different disciplines. Although there is reason to believe that MDTs may be better than homogeneous teams at generating and selecting ideas, especially when a task requires that diverse knowledge and perspectives be shared and integrated, the cognitive diversity among MDTs inhibit the effectiveness of these teams. Thus, the next section explores additional team process strategies suggested by previous literature to enable diverse information perspectives to be fully utilised while reducing the negative effect of cognitive diversity and thereby enhance MDTs' performance.

2.7.2.3 Enhancing the Performance of Multidisciplinary Interacting Teams

Although the review of previous studies on brainstorming suggests that teams could benefit from the interaction process when members are cognitively diverse (Nijstad and De Dreu 2002), the social psychology literature provides mixed evidence on the benefits and detriments of MDTs on team performance (as discussed in section 2.6.3.2). On one hand, cognitive diversity enhances team performance by bringing greater knowledge and skill-sets to a given task (Guzzo and Dickson 1996; Williams and O'Reilly 1998). On the other hand, cognitive diversity is deleterious to team performance because individuals from diverse functional backgrounds may have different frames of reference, professional language, and problem-solving styles that impede the optimum sharing and recognition of diverse ideas and information (van Knippenberg and Schippers 2007). As such, the literature identifies a number of team process strategies for optimising MDTs' performance, including extending the discussion time in the start-up phase, information elaboration, and cross-understanding.

Extension of Discussion Time in the Start-Up Phase

Research has shown that group members interacting in face-to-face discussion often fail to exchange unique knowledge and focus instead on the knowledge everyone has in common (e.g., Stasser and Titus 1985; Larson et al. 1994; Stasser et al. 1995). Consequently, groups cannot take advantage of the expert members' unique knowledge and expertise. To address this problem, Larson et al. (1994) examine discussions of shared and unshared information in three-person groups and find that the information the groups have in common tends to be discussed earlier than the unique information. They also find that information discussed at later stages frequently has less effect on the decisions made by the groups. Therefore, they suggest that extending the discussion time in the start-up phase provides a greater likelihood of diverse information and perspectives being shared and considered.

Information Elaboration

Diversity could have both positive and negative effects on team performance (e.g., Guzzo and Dickson 1996; Williams and O'Reilly 1998; van Knippenberg and Schippers 2007), which explains why research on work team diversity tends to find mixed results when examining diversity's effect on team performance (see van Knippenberg and Schippers 2007, for a review). In an attempt to address inconsistent results for the effect of diversity on team performance, van Knippenberg et al. (2004, p. 1009) propose a "categorisation-elaboration model (CEM)" positing that the elaboration of task-relevant information is "the primary process underlying the positive effects of diversity on performance" (p.1012). Elaboration is defined as "the exchange of information and perspectives, the process of feeding back the results of this individual-level processing into the group, and discussion and integration of its processes" (van Knippenberg et al. 2004, p. 1011). This model proposes that diversity among team members is positively related to the elaboration of task-relevant information and perspectives and that this information elaboration affects team performance. Therefore, the potential beneficial effects of diversity are collected through the elaboration process, specifically when the task requires the combination, reconciliation, and integration of different knowledge, skills, and perspectives (van Knippenberg et al. 2004).

Van Ginkel and van Knippenberg (2008) provide the first evidence of elaboration's role in informationally diverse groups. This study's results show that diverse groups engage in more elaboration of task-relevant information and make decisions of a higher quality when they share a mental model emphasising elaboration than groups that hold such shared task representation to a lesser extent. The positive effect of elaboration on the effectiveness of diverse groups is also confirmed in further studies testing the CEM model (van Ginkel and van Knippenberg 2009; van Ginkel et al. 2009). Van Ginkel et al. (2009) examine the effect of team reflexivity (i.e., the process of discussing ideas about the task, task goals, and possible strategies) and shared task representations (i.e., the extent to which team members understand the importance of information elaboration) on the levels of information elaboration and team performance. The results from this study show that team reflexivity promotes the development of task representations, which then increases the level of information elaboration and improves decision quality. Van Ginkel and van Knippenberg (2009) find that knowledge about distributed information (i.e., team members know which members hold certain information) improves the quality of group decision making through task representations emphasising elaboration. That is, when diverse teams have knowledge about distributed information, they understand the importance of information elaboration better, which then stimulates information elaboration and enhances the decision-making performance. These studies suggest that information elaboration mediates the relationship between team diversity and team performance.

The findings from previous studies support van Knippenberg et al.'s (2004) proposition that information elaboration drives the positive effect of diversity on team performance. However, information diversity in these prior studies was manipulated by giving different information to each team member rather than via actual discipline diversity. Further, these studies examine the role of elaboration by focusing exclusively on decision-making tasks. The role of elaboration has not yet been examined using MDTs working in an organisation and on idea generation tasks.

Cross-Understanding

Huber and Lewis (2010) propose 'cross-understanding', which is another potential construct, as a means of explaining inconsistencies in the work team diversity literature. Cross-understanding refers to "the extent to which group members have an accurate understanding of one another's mental models" (Huber and Lewis 2010, p. 7). They suggest that team members are more likely to predict other members' behaviours if they understand "what others know, believe, are sensitive to, and prefer" and thus will be able to select their responses more effectively (Huber and Lewis 2010, p. 9).

Huber and Lewis (2010) posit that cross-understanding between team members affects team processes and outcomes through three mechanisms: communication, elaboration, and collaboration. First, cross-understanding allows members to choose concepts and words that members with different educational backgrounds can understand, thereby enhancing communication effectiveness. Second, cross-understanding allows members to be aware of what other members know, believe, and prefer. This could prompt them to ask for clarification, discuss, or elaborate on issues related to others' knowledge, beliefs, or preferences, which thereby increases elaboration effectiveness. Third, cross-understanding allows team members to recognise the differences between their team members' mindsets and their own mindsets. This helps them anticipate other members' behaviours and select their responses to such behaviours more appropriately, thereby enhancing collaboration effectiveness.

Cross-understanding is particularly important for teams engaged in tasks requiring that diverse knowledge and perspectives be shared and integrated as well as some degree of task interdependence and cooperation between diverse team members to complete the task (Huber and Lewis 2010). By encouraging members to share, discuss, and integrate their diverse knowledge and perspectives while increasing the quantity and quality of task-relevant information discussed in the group, cross-understanding between members could counteract the negative effects of cognitive diversity found in MDTs, thus improving MDGHGTs' performance.

In addition to face-to-face interacting teams, auditing research has also suggested another form of team commonly used in practice, i.e., review teams. The next section discusses the nature and effect of the audit review process on the performance of financial audit teams examined by previous audit judgment and decision-making literature.

2.7.3 Review Teams

Interactions among team members can come in different forms. In the financial audit setting, one interaction that occurs between auditors has received a lot of attention from audit researchers (see Rich et al. 1997b and Nelson and Tan 2005 for reviews); this interaction is a process known as the "audit review process". This form of interaction usually involves superior members of audit teams evaluating the work of and providing guidance to subordinate team members. Unlike face-to-face interacting teams (i.e., brainstorming), the review process has consistently been found to enhance team judgments.

Trotman and Yetton (1985) examine the effect of the review process on the consensus of internal control evaluations and find that the review process significantly improves the level of consensus (i.e., reviewers versus individuals). However, no differences are found between the performances of the review teams, nominal teams, and interacting teams. Trotman and Yetton (1985) suggest that the addition of a second opinion, regardless of its form, seems to improve audit effectiveness. More importantly, they suggest that they do not find any differences between the three forms of teams because the auditors may not have been able to differentiate their expertise from others. Alternatively, it could be difficult for the auditors to do so due to the lack of variation in their expertise.

Trotman (1985) uses a more complex task to compare the accuracy of judgments made by an interacting team comprising a manager and a senior with a nominal team and a review team in which the manager reviews the senior's judgment without any discussion. To make it easier for participants to recognise the differences in their expertise, professional auditors who have worked together before are chosen, a more complex task was used, and participants were allowed to present their calculations to other team members. The results show that the review process increases the accuracy of auditor judgment by reducing the systematic bias and variance in an individual auditor's judgment. The comparison between different types of teams reveal the superiority of interacting teams in that the interacting teams outperform the nominal teams in terms of judgment accuracy. However, no difference is found between the interacting teams of equal rank and the review process team.

Using an idea generation task, Ismail and Trotman (1995) examine the effectiveness of the review process on hypotheses generation tasks. The experiment involves teams of six: two seniors (reviewees) working individually to generate a list of hypotheses, two more experienced seniors (senior-reviewers), and two managers (manager-reviewers) reviewing a set of hypotheses prepared by audit seniors. The reviews were conducted either with or without discussion with the reviewees. This study finds that the review process increases the number of plausible hypotheses generated regardless of the team members' experience (i.e., senior or manager) or team interaction (i.e., with or without discussion). Ismail and Trotman (1995) suggest that during the review process, auditors may benefit from a larger pool of information or may be stimulated by other auditors' ideas. The stimulation of ideas when seeing the ideas of others is also evidenced in the electronic brainstorming setting. Kohn et al. (2011) find that participants who first generate ideas individually before joining the brainstorming group tend to use others' ideas to form combinations of ideas, while participants who first generate ideas in groups tend to use their own ideas to form a combination of ideas. Thus, this study concludes that participants are stimulated by the ideas of others especially when they see those ideas for the first time.

More recent studies on the audit review process examine the effect of alternative review formats on reviewers and reviewees' performances (Brazel et al. 2004; Agoglia et al. 2009; Payne et al. 2010). Brazel et al. (2004) examine the effect of face-to-face and electronic reviews on the effectiveness and efficiency of workpaper preparers. This study focuses on the performance of audit workpaper preparers (reviewees) and finds that preparers provide more effective workpapers, make higher quality judgments, feel more accountable, and are less efficient when they expect a face-to-face review (as opposed to an electronic review). This study also compares the two review formats with a control team (no review). No difference is found between the electronic review and control teams, whereas a significant difference is found between the face-to-face and control teams. This study suggests that review teams could benefit from the nature of a face-to-face review because it allows a real-time response, a reviewer to be present, and more effective communication.

While Brazel et al. (2004) focus on the performance of reviewees (workpaper preparers), Agoglia et al. (2009) investigate the effects of face-to-face and electronic review on the quality of the reviewers' judgments. This study finds that reviewers in the electronic review condition make lower quality judgments than reviewers in the face-to-face condition. The quality of the preparers' workpapers is also lower when preparers anticipate an electronic review compared to when they expect a face-to-face review. The mediation analysis reveals that reviewers in the electronic condition have difficulties recognising and mitigating lower-quality workpapers, which then leads to a lower quality of going concern judgments.

Payne et al. (2010) compare the effects of adding discussion after the preparation of written review notes (as opposed to adding no discussion). This study finds that a face-to-face discussion of written review notes increases audit effectiveness because preparers more thoroughly examine the audit evidence compared to when they anticipate written review comments.

Altogether, the findings from previous studies on the audit review process suggest that the review process enhances audit team effectiveness. However, the audit review process has only been tested in a financial audit setting in which superiors review the work of subordinates who have similar educational backgrounds (i.e., an accounting/auditing background). Whether the benefit of the review process also carries through to the MDTs setting (i.e., accountant practitioners reviewing the work of non-accountant practitioners) is yet to be determined.

2.8 Summary

MDTs are increasingly used to make important decisions, develop new products, and solve complex problems. This trend is also evident in the GHG assurance setting, with ISAE 3410 requiring MDTs to be involved in planning GHG assurance engagements, including discussions to assess an entity's potential material misstatements because of fraud or error. The need for MDTs stems from the notion that members from different educational backgrounds are likely to bring diverse knowledge, skills, and perspectives to the team that are beneficial in dealing with complex tasks. In the GHG assurance setting, this need occurs because of the technical nature of the assurance subject matter. However, the cognitive diversity surrounding MDTs could make it difficult for them to communicate, coordinate, and perform the required tasks.

The team effectiveness frameworks in psychology and management literature and the findings in the audit team literature have suggested that various team inputs and processes factors affect MDGHGTs' effectiveness, including environmental factors, task characteristics, team composition, and team processes. However, a limited amount of auditing literature has examined how these factors affect MDTs' effectiveness. Among all the factors, the frameworks suggest that team processes are central to the effectiveness of teams. Different team processes have been examined in the group decision making literature in psychology and auditing through the manipulation of different team formats, i.e., nominal, interacting, and review teams. While a well-developed body of literature exists on group decision making, limited research has been conducted on assurance teams comprising practitioners from distinctive disciplines. Further, previous studies in psychology and management suggested additional team processes strategies that could maintain the benefits of diversity while reducing the difficulties faced by interacting MDTs. These team process strategies include the extension of discussion time in the start-up phase, the elaboration on task-relevant information, and cross-understanding between MDT members. Despite extensive theories in the social psychology literature on strategies to enhance MDT's performance, auditing research addressing the effects of such strategies on MDTs' performance is currently absent.

In sum, while prior research provides some insights into team input and process factors that may affect MDTs' performance, this research has yet to examine how these factors affect teams comprising professionals with diverse educational backgrounds that perform idea generation and selection tasks. In relation to idea generation performance, the majority of the group decision making research in both psychology and auditing has used subjects with similar educational backgrounds (either students or practitioners) rather than practitioners with distinctive educational backgrounds to examine the effect on performance in different team formats. Further, strategies to enhance MDTs suggested by social psychology literature are normally tested using informationally diverse teams (i.e., teams in which members hold different distinct pieces of information) performing decision-making tasks rather than educationally diverse teams performing idea generation tasks. This dissertation aims to fill these identified gaps by testing the findings from prior research in the emerging GHG assurance setting, in which practitioners from distinctive educational backgrounds (i.e., accounting, science, and engineering) are required to work together on idea generation and idea selection tasks to assess the risk of material misstatements.

The main aim of this dissertation, therefore, is to find ways to improve MDGHGTs' effectiveness by understanding how these teams' performance is affected by factors affecting MDTs' performance in other settings. This aim is addressed through two studies. Chapter 3 presents a retrospective recall study investigating factors that could affect MDGHGTs' perceived effectiveness, including environmental factors, task characteristics, team composition, and team processes. This knowledge is then applied in an experimental study, presented in Chapter 4, which aims to examine the effect of different team formats on MDGHGTs' risk assessment performance.

CHAPTER 3

STUDY ONE: FACTORS ASSOCIATED WITH THE EFFECTIVENESS OF MULTIDISCIPLINARY GREENHOUSE GAS ASSURANCE ENGAGEMENT TEAMS

3.1 Introduction

This study aims to provide empirical evidence on factors significantly affect the effectiveness of multidisciplinary GHG assurance teams (MDGHGTs). This evidence is informed by team effectiveness frameworks from the psychology, management, and auditing literatures. Testing important factors (e.g., environmental factors, task characteristics, team composition, and team process variables) suggested by the social psychology, management, and auditing literatures in the context of MDGHGTs is a significant and innovative application of this literature to address important assurance issues.

ISAE 3410 ("Assurance Engagements on Greenhouse Gas Statements") requires practitioners from various disciplines to work together on either reasonable or limited assurance engagements (IFAC 2012a, para. 6 and A42). The implications of assurance requiring multidisciplinary teams (MDTs) to work together on GHG assurance engagements have not been empirically tested. While ISAE 3410 notes that including members with diverse expertise will help in quantifying and reporting emissions, particularly on relatively complex engagements (IFAC 2012a, para. A19 and A42), the social psychology literature discussed in Chapter 2 provides mixed evidence on the benefits and detriments of MDTs on team performance (e.g., Jackson 1996; Jackson et al. 2003; van Knippenberg and Schippers 2007; Jackson and Joshi 2011). A number of team effectiveness frameworks (e.g., Kozlowski and Ilgen 2006; Mathieu et al. 2008) note that a team with diverse members is only one important success factors. These frameworks have suggested various inputs and process attributes (e.g., task characteristics, team processes, and environmental factors) that could also significantly affect team effectiveness.

Although extensive literature examines group decision making in auditing, these studies mainly focus on the effectiveness of financial audit teams, which are hierarchical in nature and are typically composed of members with similar educational/professional backgrounds (i.e., accounting). A very limited amount of research investigates auditing factors associated with the effectiveness of multidisciplinary assurance teams in which the educational diversity of the team members is more of interest than the hierarchical nature between superiors and subordinates in the team.

To add to the limited amount of empirical evidence on MDGHGT effectiveness, this study develops a MDGHGT effectiveness framework to investigate the relationships between team effectiveness and factors that are under the control of the assurance firms, including team composition and team processes. The effect task characteristics and exogenous environmental factors have on team effectiveness are also explored. The study also demonstrates how MDGHGTs are operationalised in practice by presenting descriptive data about GHG assurance engagement characteristics, MDGHGT composition, and team processes, thereby providing new information to the literature. This evidence is then used to inform the experimental design of Study Two.

The study utilises a retrospective recall design (Gibbins and Trotman 2002; Fargher et al. 2005; Gibbins et al. 2007) to examine how MDGHGTs work together in the field as they complete their GHG assurance engagements. Data are reported by GHG assurance professionals who were in a position to report on their own experiences as part of a GHG assurance team on two separate engagements: one in which they thought the team worked more effectively together and one in which they thought the team worked less effectively together. Using GHG assurance team members' retrospective recall in a field setting enables this study to reveal GHG assurance engagement characteristics, team composition, and team process features not currently explored in the literature.

Given the minimal guidance currently available in existing regulations and assurance standards regarding the composition, selection, and evaluation of GHG assurance teams, this study's findings increase the understanding of firms' current practices, which could inform the development of more detailed guidance in this area. As GHG assurance engagements and other assurance engagements (e.g., assurance on corporate social responsibility reports and integrated reports) become a more significant business stream for leading assurance firms (KPMG 2008, 2013), the need for MDTs is emphasised. While it is necessary in some circumstances to engage non-accountant practitioners for some aspects of financial statement audits (such as when an actuary needs to determine appropriate loan loss provisions), non-accountant practitioners are required for many GHG assurance engagements and are indispensable in complex GHG assurance engagements. Further, in contrast to MDGHGTs, such non-accounting practitioners are usually used in financial audits on an ad hoc basis rather than as integral members of the assurance team (Griffith 2014). Assurance firms are currently in the early stages of developing these engagements and evaluating whether an appropriate team has been allocated, whether appropriate team processes have been followed, and whether engagements have been performed effectively. As such, this study provides evidence with the potential to assist firms in these evaluations.

The remainder of the chapter is structured as follows. Section 3.2 discusses team effectiveness frameworks. The MDGHGT effectiveness framework is also developed, and the relationships between each component in the framework and the relevant literature are discussed. Section 3.3 contains the hypotheses development, and Section 3.4 describes the research methodology. Section 3.5 reports the results, including the descriptive results and the tests of the hypotheses. Section 3.6 provides the results of the sensitivity analyses, Section 3.7 reports the secondary analysis, and Section 3.8 reports the additional analyses. Section 3.9 summarises and discusses the implications and limitations of the study.

3.2 Theory and Research Framework

3.2.1 Team Effectiveness Frameworks

The theoretical framework used in this study is developed from recent team effectiveness frameworks suggested in the psychology and management literature, which address various attributes that significantly affect team effectiveness (e.g., Kozlowski and Ilgen 2006; Mathieu et al. 2008). As discussed in the literature review, the most classic input-process-outcome (IPO) framework formulated by McGrath (1964) primarily shapes the way team effectiveness is conceptualised (Kozlowski and

Ilgen 2006). In an attempt to illustrate the IPO framework, Mathieu et al. (2008) define *inputs* as factors that facilitate and inhibit interaction among team members (e.g., task characteristics, team composition), *processes* as interactions among team members that turn team inputs into outputs (e.g., discussions), and *outcomes* as products of the team activity that are useful to the organisation or other parties.

Although team effectiveness can be measured in many ways, its measurement has mostly been classified into three categories: (1) performance effectiveness, (2) members' attitudes, and (3) behavioural outcomes (Cohen and Bailey 1997; Mathieu et al. 2008). Many studies summarise and review measurements of team effectiveness (e.g., Cohen and Bailey 1997; Sundstrom et al. 2000; Mathieu et al. 2008). The measurement metrics used in these studies include quantity, quality, productivity, time spent, supervisor-rated performance, satisfaction with team services, team innovativeness, employee satisfaction, and team commitment. Cohen and Bailey (1997) also reported that most survey studies on team effectiveness focused on team members' and managers' perceptions of the overall team performance.

In terms of input factors, task characteristics and team composition are commonly seen as influencing team effectiveness (West and Anderson 1996; Cohen and Bailey 1997; Kozlowski and Ilgen 2006). To test the effect of task characteristics on team effectiveness, researchers have mainly focused on two aspects: the degree of task coordination within the team (i.e., "task interdependence") (Kiggundu 1981; Stewart and Barrick 2000; Morgeson and Humphrey 2006; Humphrey et al. 2007) and task type (Straus and McGrath 1994; Jehn et al. 1999; Stewart and Barrick 2000; Van der Vegt et al. 2000). Research has also examined the role of team composition by focusing on two salient team composition and structural variables: team size (Curral et al. 2001; West et al. 2003; Fay et al. 2006; LePine et al. 2008) and diversity (Bantel and Jackson 1989; Milliken and Martins 1996; Jehn et al. 1999; van Knippenberg et al. 2004). However, there is a small amount of evidence shows the effect of these task characteristics and team composition variables on the effectiveness of work teams comprising practitioners from highly distinctive areas of expertise (e.g., accounting and engineering/science).

The majority of team effectiveness models also highlight the central role of team processes (Guzzo and Dickson 1996; Cohen and Bailey 1997; Ilgen et al. 2005; Mathieu et al. 2008). More recent team effectiveness models (Ilgen et al. 2005; Mathieu et al. 2008) tend to focus on different processes that also mediate the relationship between team input and output, such as social processes (e.g., idea sharing) and cognitive processes (e.g., shared mental models and elaborating on different information and perspectives). Because team processes are defined as a mechanism to integrate different knowledge and expertise possessed by team members and to coordinate efforts to resolve task demands (Kozlowski and Ilgen 2006), the importance of team processes to MDTs becomes even more salient. When dealing with complex tasks, MDT members' complementary expertise has been found to be advantageous (Jehn et al. 1999; Pelled et al. 1999; van Knippenberg and Schippers 2007). However, the benefits of diverse knowledge and skill-sets will only be utilised effectively through effective team processes (West and Anderson 1996).

Kozlowski and Ilgen (2006) argue that environment dynamics and complexity drive the team task demand and team processes, shape the team design to solve the task demand, and result in increased team effectiveness. They suggest that it is important to identify the primary context in which teams are embedded and whether they are linked more tightly to the organisational system (e.g., firm policies) or the task environment (e.g., client financial condition, client inherent risks). They note that the primary context is the key driver for the difficulty and complexity of the team task and that it influences the way teams work together. Accordingly, environment factors are likely to vary with context.

For teams in the auditing and assurance services, such as financial audit teams and GHG assurance teams, "client characteristics and risks" can be identified as their primary context. In the auditing context, the audit production framework developed by Knechel et al. (2009) suggests that client characteristics and risks should be considered as "exogenous factors that affect audit production as a whole" (p. 1607). This view is supported previous audit literature's extended use of client characteristics and risks as the factors that affect audit effectiveness (O'Keefe et al. 1994; Hackenbrack and Knechel 1997; Gibbins and Trotman 2002; Asare et al. 2005, 2007). Another unique factor in the audit/assurance context is the “client-assurer relationship”. Extensive studies have been conducted on auditors’ familiarity with the client (Tan 1995; Johnson et al. 2002; Myers et al. 2003; Favere-Marchesi and Emby 2005; Carey and Simnett 2006) and the client’s importance (DeAngelo 1981; Reynolds and Francis 2001; Craswell et al. 2002; Chung and Kallapur 2003; Carcello and Nagy 2004; Larcker and Richardson 2004; Li 2009). These factors could also be considered environmental factors that affect audit/assurance team effectiveness.

As previously discussed, the team effectiveness frameworks developed by the psychology and management literatures suggest a number of environmental, inputs, and process factors that could affect MDGHGTs’ effectiveness. The audit production framework in auditing also suggests that client characteristics and risks are an environmental factor that can affect the entire audit/assurance production and thus audit/assurance effectiveness. Further previous audit literature suggests that the relationship between the client and the assurer could affect audit/assurance effectiveness. This study therefore develops a MDGHGT effectiveness framework based on various factors informed by these frameworks and studies in the psychology, management, and auditing literatures. The MDGHGT effectiveness framework will be discussed in detail in the next section.

3.2.2 The MDGHGT Effectiveness Framework

Figure 3.1 shows the six components of the MDGHGT effectiveness framework. The framework demonstrates that the first five components (client characteristics and risks, client–assurer relationship, task characteristics, team composition, and team processes, are proposed to influence the sixth. However, the main interest of this study is on factors that are under the control of the assurance firms (i.e., things that assurance firms can do to improve MDGHGT effectiveness). Therefore, this study develops hypotheses around the relationships between the three key components: 1) team composition; 2) team processes; and 3) team effectiveness (Components 4, 5, and 6 in Figure 3.1). The other three components, which are inherent to the GHG assurance setting (client characteristics and risks, client–assurer relationship, and task characteristics) are treated as control variables in this study (Components 1, 2, and 3 in Figure 3.1).

First, the framework assumes that team process variables directly affect MDGHGT effectiveness (Path 1). Thus, when the team processes better, the diverse knowledge and perspectives of MDGHGT members are shared and integrated better, thus leading to better team effectiveness. Second, team composition factors may have indirect effects on team effectiveness, as shown by the dotted arrow in Figure 3.1 (Path 2), while also having direct effects on team processes (Path 3).

The framework also assumes that environmental factors (client characteristics and risks and the client–assurer relationship) affect the MDGHGT inputs and processes as a whole and thus affect MDGHGT effectiveness, while the task characteristics may affect the team composition (i.e., staffing decisions) and team processes. Therefore, although environmental and task characteristics factors are treated as control variables, the effects of these control variables on team composition, team processes, and team effectiveness are also explored.

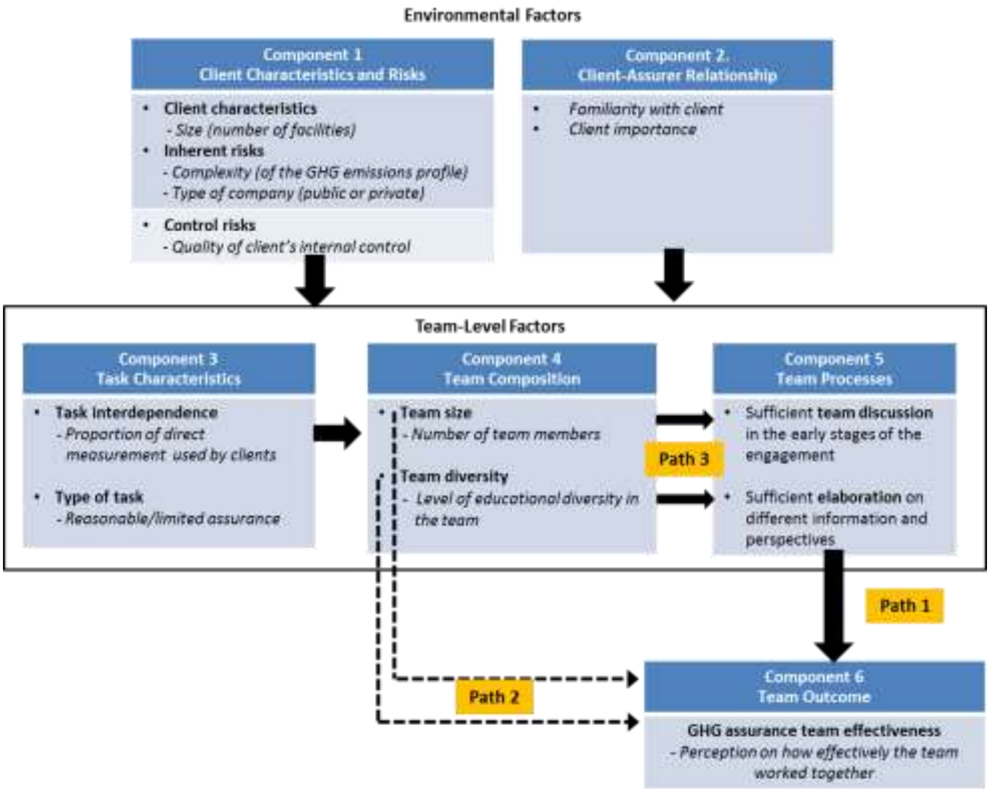


FIGURE 3.1 Research Framework: The Six Components of MDGHGT Effectiveness

Regression analysis is used to examine the relationships between team composition, team processes, and MDGHGT effectiveness (Paths 1, 2, and 3). Regression analysis is considered appropriate in this study because of the nature of the variables of interest. Because the variables in this framework are measured and observable, using alternative methods, such as path analysis, is not an advantage for this study[8]. In addition to regression analysis, this study recognises the potential mediation relationship between team composition, team processes, and MDGHGT effectiveness. However, since the main objective of this study is to explore factors with significant effects on MDGHGT effectiveness, the potential mediation relationship is tested as a secondary analysis.

Figure 3.1 shows the 12 variables examined within the six components. The previously mentioned studies on team effectiveness and audit quality suggested that these variables are potentially related to GHG assurance teams' effectiveness. The variables examined within each component are listed below:

1. Client characteristics and risks
   1.1 Client size
   1.2 Complexity of the client's GHG emissions profile
   1.3 Type of client's company (public or private)
2. Client–assurer relationship
   2.1 Familiarity with client
   2.2 Client importance
3. Task characteristics
   3.1 Task interdependence
   3.2 Task type
4. Team composition
   4.1 Team size
   4.2 Team diversity
5. Team processes
   5.1 Sufficient team discussion in the early stages of engagement
   5.2 Sufficient elaboration on different information and perspectives

[8] Path analysis is more useful to examine relationship between 'unobservable' rather than 'observable' variables (Iacobucci 2009; Urbach and Ahlemann 2010). As such, there is no incremental benefit for using path analysis if the variables are all observed.

The first seven variables (1.1 to 3.2) are treated as control variables, while the last four variables (4.1 to 5.2) are the independent variables of interest. The dependent variables are the perceived team effectiveness, sufficient team discussion, and sufficient elaboration. These variables will be discussed in the next sections.

3.2.3 Environmental Factors

3.2.3.1 Client Characteristics and Risks

Client characteristics and risks are considered to be highly related to MDGHGTs' effectiveness because ISAE 3410 adopts a risk-based approach for the conduct of assurance on GHG emissions statements (IFAC 2012a). An important client characteristic that could affect MDGHGT effectiveness is client size. *Client size* is found to increase audit effort (O'Keefe et al. 1994; Hackenbrack and Knechel 1997), audit review time (Gibbins and Trotman 2002), and audit efficiency (Knechel et al. 2009). Client size also influences the task characteristics and team composition. O'Keefe et al. (1994) find that more audit staff are included in the team when the client size increases. Similarly, Hackenbrack and Knechel (1997) find that audits of large entities consume more labour and increases the demand for substantive testing and review of critical objectives compared with audits of small entities. Knechel et al. (2009) also find that client size is positively related to audit efficiency.

In addition to client characteristics, the effect of client risks on audit quality have also been examined (O'Keefe et al. 1994; Krishnan and Schauer 2000; Gibbins and Trotman 2002; Asare et al. 2005; Knechel et al. 2009). According to the audit risk model, client risks can affect the effectiveness of assurance teams through two elements: inherent risks and control risks (IFAC 2010). Inherent risks, such as *client complexity*, increase the effort auditors give to the client (O'Keefe et al. 1994; Hackenbrack and Knechel 1997; Knechel et al. 2009). *Type of company* (public or private) also increases the audit effort (O'Keefe et al. 1994) because it attracts more public attention, which then induces greater political scrutiny and regulatory pressures for the firm (Brammer and Pavelin 2008). Consequently, the amount of work and audit effort required when undertaking assurance engagements for public companies is greater than for private companies (O'Keefe et al. 1994). Another element of client risk is control risk, which includes the *quality of the client's internal control* (i.e., client's accounting systems, quality of the reports, and quality of the report preparers). O'Keefe et al. (1994) find no association between the client's internal control system (i.e., perceived quality of the overall internal control system) and audit effort, while Hackenbrack and Knechel (1997) find that auditors put less effort into substantive testing and review of noncritical issues when the client adopts highly centralised and automated accounting systems. Moreover, Knechel et al. (2009) find that audits are more efficient for clients with highly automated systems and less efficient when auditors rely on the clients' internal control. In the context of audit review, Gibbins and Trotman (2002) find that the quality of the preparer and the preparer's work are associated with less review effort. Given that increasing audit effort leads to more effective audits (Francis 2004), factors such as client size, complexity, type of company, and the quality of the client's internal control system could affect team effectiveness.

Client characteristics and risks also influence audit team composition. O'Keefe et al. (1994) find that more audit partners and managers are required when the client's financial risks become higher and that more audit seniors and staff are required when the client's inherent risks become higher. Hackenbrack and Knechel (1997) find that more auditors are required when client complexity increases and when the client is a public company, while less auditors are required when the client uses highly centralised and automated accounting systems.

As noted in ISAE 3410, the complexity of GHG assurance engagements varies depending on client characteristics (e.g., number of GHG emissions facilities and industry), scope of emissions (e.g., Scope 1 or Scope 2 emissions)[9], and GHG quantification methods (e.g., direct measurement or estimation)[10]. Further, different

[9] The Greenhouse Gas Protocol defines Scope 1 emissions as direct GHG emissions from sources that are owned or controlled by the company, such as emissions from combustion in boilers and furnaces and emissions from chemical production. Scope 2 emissions are defined as indirect GHG emissions from electricity generation that are purchased, transferred, and consumed by the company (WBCSD and WRI 2004).

[10] Direct measurement (or direct monitoring) may be adopted to measure GHG concentration and flow rates; it uses continuous emissions monitoring or periodic sampling (IFAC 2012a, para. A22a). Such measurement systems can be established, for example, in flues, stacks, pipes, or ducts and may be applicable in a number of industries, although not all industries (Australian DCCEE 2010, p. 46). However, emissions are most often estimated by references to readily observable variables that are closely related to GHG emissions, such as the quantity of electricity, gas, and fossil fuels consumed (Australian DCCEE 2010, p. 25). These estimation procedures involve the use of designated emission

types of expertise are required to deal with quantifying and reporting emissions, particularly when the engagement is relatively complex (IFAC 2012a). Thus, the client characteristics and risks are likely to influence the entire GHG assurance process and thus affect MDGHGT effectiveness.

3.2.3.2 Client–Assurer Relationship

Another important environment factor affecting assurance teams' performance is the client–assurer relationship. Numerous audit studies investigate whether *familiarity with the client* affects audit quality (e.g., Deis Jr and Giroux 1996; Johnson et al. 2002; Myers et al. 2003; Favere-Marchesi and Emby 2005; Carey and Simnett 2006). Empirical research, however, has suggested two competing views on this issue. On one hand, prior involvement in audit engagement has been found to impair auditor judgments (Tan 1995; Favere-Marchesi and Emby 2005) and audit quality (Carey and Simnett 2006). On the other hand, client-specific expertise and knowledge with regard to the industry can be developed by auditors over time, which in turn increases audit quality (Johnson et al. 2002; Myers et al. 2003). However, these studies focus on the audit partner tenure rather than the audit team tenure.

The effect of *client importance* on audit quality has also been widely examined (e.g., DeAngelo 1981; Reynolds and Francis 2001; Craswell et al. 2002; Chung and Kallapur 2003; Carcello and Nagy 2004; Larcker and Richardson 2004; Li 2009). DeAngelo (1981) argues that an auditors' financial dependence on their clients generates incentives for the auditor to compromise their independence to retain their clients. However, prior studies do not find a negative association between audit quality and client importance either at the national audit firm level (e.g., Chung and Kallapur 2003; Larcker and Richardson 2004) or at the local office level (e.g., Li 2009). There is insufficient evidence that economic dependence posed by important clients (i.e., client fees) compromises audit quality (Craswell et al. 2002). Moreover, auditors report more conservatively for more important clients (Reynolds and Francis 2001).

---

factors in estimating emissions. For example, an emissions factor will be applied to electricity consumption to estimate the emissions.

In summary, various environmental factors have been suggested to affect MDGHGT effectiveness including client size, complexity, type of company (public or private), quality of the client's internal control, familiarity with the client, and client importance.

In addition to environmental factors, the team effectiveness frameworks suggest that other input and process factors, such as task characteristics, team composition, and team processes, play important roles in determining team effectiveness. Thus, these factors' effects on the effectiveness of MDGHGTs are addressed in the following section.

3.2.4 Task Characteristics

Task interdependence refers to the degree to which team members depend on one another to accomplish their task effectively (Kiggundu 1981; Brass 1985; Campion et al. 1993; Humphrey et al. 2007). Therefore, the degree of task interdependence usually varies depending on the complexity of the task and other team members' need for help to execute the tasks (Van der Vegt et al. 2000). As suggested by previous literature, task interdependence facilitates open communication, better cooperation, and less conflict among team members and results in higher quality group processes (e.g., Wageman 1995; Campion et al. 1996; Stewart and Barrick 2000). However, these benefits depend largely on the level of interdependence. Wageman (1995) finds that task interdependence enhances team effectiveness when the task interdependence among team members is high or low but not when such interdependence is moderate. This curvilinear relationship is also confirmed in later studies (e.g., Stewart and Barrick 2000). Campion et al. (1996) find that teams work more effectively together when the members rely more on one another when working on the task, the teams are more diverse, and the team processes are more effective. Stewart and Barrick (2000) suggest that when task interdependence is high, team members collectively work together to complete a task while sharing information and resources. In contrast, for a task requiring low interdependence, team members tend to operate more independently, thereby reducing the need for coordination and collaboration among members. In a meta-analysis, Humphrey et al. (2007) find that task interdependence is positively related to behavioural outcomes, supervisor satisfaction, organisational commitment, job involvement, and internal work motivation.

In the GHG assurance context, the degree of task interdependence in GHG assurance engagements could vary based on the complexity of the GHG emissions profile. ISAE 3410 recognises the need for scientific and engineering expertise, particularly when a client's GHG emissions profile involves significant Scope 1 emissions that result in a high proportion of direct measurement used to quantify the emissions (IFAC 2012a, para. A19). In such complex technical tasks, team members may need to rely more on the expertise of non-accountant practitioners in the team. However, when an entity has Scope 2 emissions, the designated emission factors determined by regulators will be used to quantify the emissions[11] (IFAC 2012a, para. A19), e.g., emissions from the generation of purchased electricity consumed by a service company. In this case, the engagement may be less complex and less dependent on non-accountant practitioners' competence. Therefore, the degree of task interdependence may be lower in the latter case.

Task type is another aspect of task characteristics that has been examined in the team literature (e.g., McGrath 1984; Straus and McGrath 1994; Jehn 1995; Jehn et al. 1999; Bowers 2000; Stewart and Barrick 2000). This literature finds that different types of tasks have different effects on team effectiveness. For example, Stewart and Barrick (2000) find that compared to behavioural tasks (e.g., execution tasks), intellective tasks (e.g., generating ideas, decision making, and negotiating) weaken the relationships between task interdependence and team performance. They suggest that intellective tasks, such as idea generation and decision making, require a higher degree of interaction and coordination among team members compared to behavioural tasks (e.g., execution tasks). Therefore, a high degree of task interdependence improves team effectiveness when teams perform tasks requiring a high degree of interaction and coordination. Teams performing complex tasks also perform more effectively than those who perform simple tasks (Bowers et al. 2000), particularly when team members possess a different knowledge base (Jehn et al. 1999). When dealing with a complex task that is not well understood, team members may disagree about the task, which then

[11] This method involves calculating emissions by applying, for example, mass balance equations, entity-specific emissions factors, or average emissions factors for a region, source, industry, or process to surrogate activities (IFAC 2012a, para. A22b). For example, an emissions factor will be applied to electricity consumption to quantify emissions. In Australia, emission factors are national average factors determined by the Department of Climate Change and Energy Efficiency using the Australian Greenhouse Emissions Information System (Australian DCCEE 2010, p.25).

forces them to discuss and reconcile their diverse knowledge and perspectives (Jehn 1995; Jehn et al. 1999). In contrast, when dealing with a simple, well understood task, it is unnecessary for team members to discuss their disagreements because they can adopt standard procedures to perform the task (Jehn et al. 1999).

With regard to GHG assurance, ISAE 3410 notes that practitioners will be working on either a reasonable assurance or a limited assurance engagement (IFAC 2012a, para. 6 to 8). ISAE 3410 states: "Because the level of assurance obtained in a limited assurance engagement is lower than in a reasonable assurance engagement, the procedures the practitioners will perform in a limited assurance engagement will vary in nature from, and are less in extent than for, a reasonable assurance engagement" (IFAC 2012a, para. 8). Accordingly, reasonable assurance engagements could be more complex and require more interaction and coordination between accountant and non-accountant practitioners in MDGHGTs than limited assurance engagements. However, given the communication and coordination difficulties found in MDTs (Bhappu et al. 1997), how the differences in task complexity and interaction required between the two types of GHG assurance engagements will affect MDGHGT effectiveness is unknown. Further, the different amounts and natures of work required by reasonable and limited assurance engagements may influence the size of the team and the combination of accountant and non-accountant practitioners included.

In conclusion, task characteristics, including task interdependence and task type, could determine MDGHGTs' effectiveness because they may affect not only the way teams should be composed but also how the team should operate. Apart from the task characteristics, two additional factors (team composition and team processes) also likely affect MDGHGTs' effectiveness. These factors' effects on team effectiveness are discussed further in the next sections.

3.2.5 Team Composition

Research on team composition mainly focuses on two salient team composition and structural variables: *size* and *diversity* (Guzzo and Dickson 1996; Stewart 2006). Team size and diversity have positive and negative effects on team effectiveness (see Cohen and Bailey 1997 and van Knippenberg and Schippers 2007, for reviews). These variables are discussed below.

Team size is considered an important determinant of team effectiveness. However, empirical research shows that the effect of team size on team effectiveness is somewhat mixed. On one hand, larger teams are at an advantage because of the higher amount of cognitive resources available to the team (Bantel and Jackson 1989; Wiersema and Bantel 1992; West et al. 2003; Fay et al. 2006), which then increases team effectiveness. West et al. (2003) investigate the relationship between innovation and team size (9 to 21 members) in multidisciplinary health care teams and find that the levels of innovation are higher for larger teams. Fay et al. (2006) find that team size (13 to 20 members) is positively associated with innovation. These studies conclude that large teams are more creative than small teams because their members represent skills that are more diverse and thus may be better at processing large amounts of diverse information.

However, a number of research studies on social psychology (e.g., Smith et al. 1994; Curral et al. 2001; LePine et al. 2008) find a negative relationship between team size and team effectiveness. For multidisciplinary top management teams, Smith et al. (1994) find a negative relationship between team size (5 members on average) and information communication, which in turn affects social integration and team effectiveness. They suggest that when a team is larger, the distance between team members is greater. This lack of social integration thereby impedes team interaction and indirectly affects team performance (Smith et al. 1994). Curral et al. (2001) suggest that larger teams (5 members on average) face difficulties in reaching consensus, having sufficient participation, and agreeing on objectives. Their results show that larger MDTs suffer from poorer team processes, which then make them work less effectively together, particularly when the task requires a high level of innovation. In a meta-analysis of teamwork processes, LePine et al. (2008) find that team size moderates the relationship between teamwork processes and team effectiveness. They suggest that larger teams are more likely to experience motivation and coordination losses than smaller teams. These findings are consistent with the process losses in larger teams observed in the early social psychology literature (e.g., Steiner 1972). In multidisciplinary health care teams, Poulton and West (1999) find that team members' perceived levels of participation in larger teams is lower than in smaller teams. Some evidence also shows that team size is neither beneficial nor detrimental to team effectiveness (e.g., Bantel and Jackson 1989; Wiersema and Bantel 1992). Therefore, while the effect of team size on team effectiveness is somewhat mixed, team size mainly hinders effective team processes.

Prior studies suggest that the *diversity* of team members is a “double-edged sword”, i.e., it can simultaneously enhance and reduce team effectiveness (Milliken and Martins 1996; van Knippenberg and Schippers 2007; Jackson and Joshi 2011). Among the different types of diversity, diversity in members’ educational backgrounds (i.e., cognitive diversity) consistently shows a positive effect on team performance (see van Knippenberg and Schippers 2007 and Jackson and Joshi 2011, for reviews). In multidisciplinary top management teams, Bantel and Jackson (1989) and Wiersema and Bantel (1992) find a positive relationship between cognitive diversity and organizations’ technical and administrative innovation. In cross-functional teams (i.e., teams comprised of members from different disciplines and functional units in the organisation), De Dreu and West (2001) find that having cognitively diverse members on a team stimulates creativity and divergent thoughts. A number of researchers (Jehn 1995; Jehn et al. 1999; Pelled et al. 1999) find that cognitive diversity introduces “task conflict”, or disagreements among team members on task-related issues. Debating and reconciling different viewpoints promotes thorough information processing and forces team members to be more critical in evaluating their problems and alternatives (Jehn 1995; van Knippenberg et al. 2004). Consistent with this, Jehn (1995) finds that task conflict benefits MDTs during non-routine tasks. However, they find that on routine tasks, task-related disagreements disturb MDTs’ regular functioning. Jehn et al.'s (1999) findings show that task conflict mediates the positive effect of cognitive diversity on team performance. However, although Pelled et al. (1999) find a positive relationship between cognitive diversity and task conflict, they do not find any relationship between cognitive diversity and team performance. Finally, De Dreu and Weingart (2003) find that task conflict is detrimental to team performance. Given the conflicting results from this prior research, task conflict may not be the only important driver for the positive relation between cognitive diversity and MDT effectiveness.

Despite the potential beneficial effects of cognitive diversity on MDTs' effectiveness, dissimilarity among team members with different educational backgrounds can be detrimental to team effectiveness. A number of studies on MDTs find that team members with different educational/professional backgrounds use different language/terminology that is only understood by people in the same profession or that have different meanings in other fields (Carlile 2004; Sheehan et al. 2007). These members have a different frame of reference that then impedes the understanding, communication, and effectiveness of MDTs (van Someren et al. 1998; van Asselt 2000). Further, different functional and professional backgrounds between members may elicit social categorisation processes, such as out-group vs. in-group identification (van Knippenberg and Schippers 2007), which leads to less communication (Bhappu et al. 1997) and cooperation (Chatman and Flynn 2001) between the subgroups. Randel and Jaussi (2003) support this view by showing that in cross-functional teams, team members who are identified as being from a functional background minority are less likely to contribute to their team because their opinion may not be valued by others.

Using the setting of assuring sustainability information, O'Dwyer (2011) finds that multidisciplinary teams comprised of "accountant" and "non-accountant" experts suffer from different mindsets when working together on a task, particularly with respect to the way each type of expert approaches the judgment of data. O'Dwyer (2011) shows that non-accountant experts are uncomfortable working with financial auditors because these auditors usually bring financial audit mindsets and habits to sustainability assurance by strictly following standard testing procedures, which in turn restricts their ability to deal with non-financial data. These differences can lead to a lack of understanding or misunderstandings among team members, less cooperation, and less team effectiveness. The problems associated with distinct mindsets between practitioners could be more salient in GHG assurance engagements because of the complex nature of the subject matter and the highly diverse expertise needed for this type of assurance.

Altogether, as discussed above, the previous literature on team size and team diversity suggests that these two team composition factors can positively or negatively affect MDT effectiveness. The inconsistent results indicate that the effect of team composition on team effectiveness may be subject to other factors, specifically team processes. Thus, the effect of team processes on team effectiveness is discussed in the next section.

3.2.6 Team Processes

Recent team effectiveness models (e.g., Ilgen et al. 2005; Mathieu et al. 2008) suggest that *team processes* mediate the relationship between team inputs (e.g., team composition) and outcomes (e.g., team performance). In line with these models, Fay et al. (2006) find that cognitive diversity among multidisciplinary health care team members results in better quality of innovation but only when the quality of team processes is high. Given the conflicting evidence obtained from the previous literature, the effects of team size (e.g., Fay et al. 2006; LePine et al. 2008) and diversity (e.g., van Knippenberg and Schippers 2007) on team performance may be contingent on team processes. Thus, having more members or more educationally diverse members in the team per se does not guarantee better team performance.

A number of team process strategies are identified in the literature to optimise MDTs' performance. As discussed earlier in the literature review, previous studies show that team members interacting through face-to-face discussions often fail to exchange unique knowledge and focus instead on the knowledge everyone has in common (e.g., Stasser and Titus 1985; Larson et al. 1994; Stasser et al. 1995). Consequently, teams cannot always take advantage of their team members' unique knowledge and expertise. However, Larson et al. (1994) find that information team members have in common tends to be discussed earlier by the team than unique information and that information discussed in later discussions frequently affects the decision less. This study suggests that extending MDTs' discussion time in the start-up phase provides more opportunities for diverse information and perspectives to be shared and integrated. Amason and Sapienza (1997) also finds that encouraging teams to have open discussion promotes effective sharing among team members and thus helps them make the most of their cognitive diversity.

In an attempt to address the inconsistent results reported in the previous literature regarding the effect of cognitive diversity on team performance, van Knippenberg et al. (2004) argue that different information and perspectives can be fully utilised only when all team members are willing to exchange and elaborate on task-relevant information. Accordingly, they propose that "information elaboration" is a crucial team process that underlies the true benefit of cognitive diversity on team effectiveness. Elaboration is defined as "the exchange of information and perspectives, the process of feeding back the results of this individual-level processing into the group, and discussion and integration of its processes" (van Knippenberg et al. 2004, p. 1011). Empirical studies in psychology have therefore begun to examine the role of information elaboration and have found that it improves the quality of decisions made by informationally diverse groups (van Ginkel and Van Knippenberg 2009; van Ginkel et al. 2009). Homan et al. (2007) show that the level of information elaboration is higher in informationally diverse groups than in informationally homogeneous groups. However, this result does not hold when informationally diverse groups experience a strong subgroup categorisation (i.e., male members hold information A and female members hold information B). Despite the promising results from the existing body of literature, prior studies manipulate cognitive diversity by giving different pieces of information to each group member, and participants in these studies were students with similar educational backgrounds and experience as opposed to educationally diverse members. Very little research has been conducted on the effect of having sufficient discussion time to share diverse information and sufficient information elaboration on the performance of teams comprising members from various disciplines.

3.3 Hypotheses Development

3.3.1 The Effect of Team Processes on MDGHGT Effectiveness

According to the information processing perspective, multiple knowledge and skills and a wider breadth of perspectives possessed by MDT members should enhance team effectiveness (van Knippenberg and Schippers 2007; Jackson and Joshi 2011). The previous literature (Jehn 1995; Jehn et al. 1999; Pelled et al. 1999) suggests that cognitively diverse teams benefit from disagreements on task-related issues (i.e., task conflict), which leads MDT members to thoroughly process information by debating and reconciling their different knowledge and perspectives (Jehn 1995; van Knippenberg et al. 2004). However, different cognitive styles (Holland 1973), frames of reference (van Asselt 2000), and languages used (Sheehan et al. 2007), which could result from or be accentuated by different educational backgrounds (Dahlin et al. 2005), can complicate communication and understanding between MDT members (Bhappu et al. 1997; van Someren et al. 1998). Therefore, MDTs may be unable to take advantage of their cognitive diversity without a team process that encourages and supports sharing and integrating different knowledge and perspectives.

Larson et al. (1994) find that teams with sufficient discussion time in the start-up phase share and consider more unique information when making decisions. In the MDT context, research on work team diversity (van Knippenberg et al. 2004; Homan et al. 2007; van Ginkel et al. 2009) suggest that diverse knowledge and perspectives within MDTs can enhance team effectiveness through information elaboration. By thoroughly discussing, reconciling, and integrating different information, viewpoints, or disagreements on the task, MDTs can overcome the difficulties associated with different educational backgrounds (van Knippenberg et al. 2004). Therefore, sufficient elaboration on different information and perspectives could result in MDTs working more effectively together. This elaboration process can be particularly important for tasks that transcend the knowledge of separate disciplines because the combination and integration of information from different knowledge domains is necessary (e.g., van Asselt 2000).

The need for effective team processes in which diverse knowledge and perspectives are exchanged and integrated becomes even more important in the GHG assurance context because accountant practitioners (with accounting/financial audit expertise) and non-accountant practitioners (with science and engineering expertise) are required to work together to conduct assurance engagements on GHG statements. Because they have different knowledge and mindsets (e.g., O'Dwyer 2011) and are likely to experience different social categorisations into subgroups such as accountant versus non-accountant (e.g., van Knippenberg and Schippers 2007), accountant and non-accountant practitioners are likely to experience task conflict and difficulties sharing, communicating, and understanding each other’s ideas. Consequently, they may need to take time to share, explain, reconcile, and integrate their different information and perspectives. Having sufficient discussion time in the early stages of engagement and sufficient elaboration on different perspectives, therefore, are expected to enhance MDGHGTs' effectiveness. Thus, the following hypotheses are formed:

Hypothesis 1a**:** *Multidisciplinary GHG assurance teams that perceive that they have sufficient discussion time to share diverse information and perspectives in the early stages of engagement will work more effectively together*.

Hypothesis 1b: *Multidisciplinary GHG assurance teams that perceive that they have sufficient information elaboration of different information and perspectives will work more effectively together.*

3.3.2 The Effect of Team Size on MDGHGT Effectiveness and Perceived Sufficiency of Discussion in the Early Stages of Engagement

Although larger teams are found to be more effective because additional members add more knowledge, skills, and perspectives to the team, increasing the MDT size introduces more coordination challenges (LePine et al. 2008); communication and social integration difficulties (Smith et al. 1994); and time pressure (West and Anderson 1996); these obstacles then decrease team effectiveness. However, these process losses are found in the context of multidisciplinary health care teams comprising members with diverse functional backgrounds (i.e., doctors, nurses, and medical technicians) but similar task and educational backgrounds (i.e., all are from medical science). Coordination, communication, and integration problems can be more salient in the MDGHGT context, in which members have diverse educational backgrounds (i.e., accounting versus science/engineering). Differences in the team members' common ground may lead them to interpret the same thing differently and use different professional language (van Someren et al. 1998; O'Dwyer 2011), which then restricts communication and understanding between MDT members. Given the coordination and communication barriers, larger MDGHGTs are expected to be less effective and less likely to perceive that they have sufficient discussion time to share different information and perspectives. Therefore, the following hypotheses are formed:

Hypothesis 2a**:** *The effectiveness of multidisciplinary GHG assurance teams will decrease when the number of team members increases.*

Hypothesis 2b: *Multidisciplinary GHG assurance teams will be less likely to perceive that they have sufficient discussion time to share diverse information and perspectives in the early stages of engagement when the number of team members increases.*

3.3.3 The Effect of Team Diversity on MDGHGT Effectiveness and Perceived Sufficiency of Elaboration on Different Information and Perspectives

The social psychology literature provides mixed evidence on the positive and negative effects of multidisciplinarity on team performance. From the social categorisation perspective, diversity is deleterious to team performance because it leads to an in-group/out-group bias (van Knippenberg and Schippers 2007). This is worsened by the fact that team members from diverse educational backgrounds (e.g., accounting and environmental science) may have different frames of reference and professional language that hinder the optimal sharing and integration of diverse ideas (van Someren et al. 1998; van Asselt 2000; Carlie 2004; Sheehan et al. 2007; O'Dwyer 2011). However, from the information processing perspective, MDT members are predicted to bring a broad range of task-relevant knowledge, skills, and perspectives to a given task (van Knippenberg and Schippers 2007; Jackson and Joshi 2011). This not only gives MDTs a larger pool of cognitive resources but also facilitates the true benefit of this type of diversity, which lies in the integration of diverse information and reconciliation of different perspectives; these in turn help improve teams' creativity and decision making (van Knippenberg et al. 2004).

Unlike in other MDTs examined in the literature, in which teams comprise members with slightly different expertise (e.g., health care teams with a medical background), MDGHGTs comprise two distinctive main areas of expertise: assurance expertise (primarily financial accounting and audit backgrounds) and subject matter expertise (e.g., engineering/science backgrounds). Consequently, the problems found in MDTs can be more salient in MDGHGTs. That is, accountant and non-accountant practitioners are more likely to have difficulties understanding each other's reasoning and perspectives and are more likely to have conflicting views on the task. These difficulties could force them to thoroughly elaborate on different task-relevant information and perspectives (van Knippenberg et al. 2004; Homan et al. 2007). By having sufficient elaboration on task-relevant information to reconcile team members' understanding and explain the reasoning underlying their thoughts, such high level of diversity may help improve rather than impede MDGHGTs' effectiveness. Based on this argument, the last set of hypotheses to be tested is formed:

Hypothesis 3a**:** *The effectiveness of multidisciplinary GHG assurance teams will increase when the level of diversity in the team increases.*

Hypothesis 3b**:** *Multidisciplinary GHG assurance teams will be more likely to perceive that they have sufficient elaboration on different information and perspectives when the level of diversity in the team increases.*

3.4 Research Methods

3.4.1 Research Instrument Design

A retrospective recall study using the technique developed by Gibbins and Newton (1994) and involving the use of a repeated-measures design to isolate individual differences is utilised. This approach has also been used to address issues in the review process (Gibbins and Trotman 2002; Fargher et al. 2005) and in auditor–client negotiations (Gibbins et al. 2001; Gibbins et al. 2007). The approach used in this study builds on the approach used by Gibbins and Trotman (2002) that elicits factual information and allows GHG assurance professionals to describe their experiences as part of a GHG assurance team.

A comprehensive research instrument was developed to investigate factors identified by the team effectiveness frameworks that may affect the perceived team effectiveness of MDGHGTs. The research instrument was pre-tested with professionals who perform GHG assurance engagements. The instrument included 10 sections of questions on 10 pages. Respondents were asked to answer all questions in the order they were presented and to not change any answer once they had written it. The research instrument is provided in Appendix 1.

To operationalise a repeated-measure design, participants were required to recall two recent GHG assurance engagements they were involved in: one in which they thought the assurance team worked effectively together and one in which they thought the team worked less effectively together. In both cases, participants were asked to select engagements in which at least one engagement team member was from a financial audit background and at least one team member had a background that was not in financial auditing.

Participants were asked a series of questions about the first engagement chosen (hereafter, the effective engagement), including their role in this assurance engagement, client size (estimated annual revenue, estimated annual GHG emissions, and number of facilities), client industry, type of company, reason for the assurance engagement, type of engagement (reasonable or limited assurance), relative complexity of the client's GHG emissions profile, the client–assurer relationship, and the client's systems to capture and record GHG data. Two questions also asked for the percentage of direct measurement methodologies the client used to quantify GHG data and the percentage of each type of scope of emissions in the client's assured GHG report. Consistent with previous studies (e.g., Gibbins and Trotman 2002; Fargher et al. 2005), participants were then asked to indicate on a nine-point scale (1 = low/much lower; 9 = high/much higher) their views on four aspects related to the quality of the client's report preparer: the availability of the preparer(s) to the assurance team, the capabilities of the report preparer(s) compared to similar engagements, the quality of work of the report preparer(s) compared to similar engagements, and the quality of the report preparer(s) documentation compared to similar engagements.

The next section focused on the GHG assurance engagement team's background. Participants were asked to provide details for the GHG engagement team members for the chosen client, including team role titles, educational/professional backgrounds (e.g., financial audit, engineering, and science), degree of overall involvement in the engagement, and familiarity with team members. This was followed by two sections evaluating the GHG assurance engagement team and team processes on a nine-point scale, including how well the GHG assurance team worked together (main dependent variable; 1 = did not work well together; 9 = worked very well together); sufficiency of discussion time in the early stages of engagement (1 = not enough time; 9 = more than enough time); sufficiency of elaboration and integration of different information and perspectives from different team members (1 = not enough; 9 = more than enough). Further, questions relating to the evidence gathering and evaluation stages of the assurance engagements were asked, including the extent of assurance procedures used to gather evidence, the extent of a clear separation between the information search stage and the information processing stage (1 = no separation; 9 = clear separation), and the extent of team discussion on the information collected before final evaluations and decisions were made (1 = our team did not do this; 9 = our team did this). The research instrument also contained open-ended questions to obtain further insights on factors perceived to contribute to the GHG assurance team working well together, factors inhibiting the GHG assurance team's ability to work well together, other mechanisms used to share and integrate different information and perspectives from different team members, missing factors that would have made the team work together better, and other comments on any issues the participant wished to raise regarding the team for the chosen GHG assurance engagement. All questions were then repeated for the second case (hereafter, the less effective engagement).

In the last section of the questionnaire, participants were asked to provide demographic details, including designated title within the firm, tertiary educational background, experience in conducting GHG assurance engagements, and training hours on assurance for GHG emissions.

3.4.2 Respondents

Initially, 35 respondents from three of the Big Four accounting and assurance service firms participated: 18 from Firm A, 3 from Firm B, and 14 from Firm C[12]. However, six respondents were excluded from the analysis because they had undertaken only one GHG assurance engagement and thus were not able to recall a second engagement to compare and contrast an effective team engagement and a less effective one. Panel A in Table 3.1 presents demographic information for the respondents' job responsibilities and educational background. The 29 respondents consisted of eight partners and directors, nine managers, and seven staff and seniors. The remaining five respondents did not indicate their position in the firm. For the respondents' tertiary education backgrounds, an equal percentage (21 percent) of respondents had either an accounting or an environmental science background, 17 percent had engineering or science backgrounds, an equal percentage (14 percent) had both accounting and engineering/environmental science backgrounds or engineering/science and others (e.g., MBA), and seven percent had both accounting and others (e.g., finance and economics). Two respondents did not indicate their tertiary education backgrounds.

Panel B of Table 3.1 contains demographic information for the respondents' working and training experience in GHG assurance. The average working experience for environmental/GHG assurance was 4.8 years (range 1 to 15 years), and participants had undertaken 19.3 environmental/GHG assurance engagements (range 4 to 80 engagements). Half of the respondents had experience leading GHG assurance teams and had led 15.3 engagements on average (range 2 to 60 engagements). The 13 respondents who stated they had attended training courses on assurance for GHG statements reported having 22.5 training hours on average (range 5 to 60 hours).

[12] These three firms conducted nearly all the GHG assurance engagements undertaken by the accounting profession in Australia at the time the research instrument was administered in 2011. The total number of assurers that had worked on GHG/sustainability assurance engagements was estimated at approximately 60 in Australia at that time. All of them were invited to participate in this study. Although these engagements are also undertaken by specialist engineering/environmental firms, their assurance and engagement methodology is very different from that undertaken by firms in the accounting profession; as such, they were excluded from the study. In addition, the responses were dominated by firm A and C. The fourth Big Four accounting and assurance services firm had no GHG assurance engagements at that time.

TABLE 3.1 Respondent Demographic Information

Panel A: Respondents' job responsibilities and education background

	Number of Respondents	Percent of 29 Respondents
1. Respondent's present job responsibilities:		
Executive director/Director/Associate Director/Partner	8	28%
Senior Manager/Manager	9	31
Senior Consultant/Consultant	7	24
Not answered	5	17
Total	29	100
2. Respondent's tertiary education background:		
Accounting	6	21%
Non-accounting	11	38
Combined	10	34
Not answered	2	7
Total	29	100

Panel B: Respondents' working and training experience

	Number of Respondents	Percent of 29 Respondents
1. Respondent's experience as a GHG assurance team leader:		
Yes	13	45%
No	13	45
Not answered	3	10
Total	29	100
2. Respondent's training experience for assurance of GHG statements:		
Yes	13	45%
No	14	48
Not answered	2	7
Total	29	100

	Mean	(S.D.)	Median	Range Low	Range High
3. Respondent's working and training experience on GHG assurance engagements:					
Years of environmental/GHG assurance experience	4.82	(3.71)	4	1	15
Number of environmental/GHG assurance engagements undertaken	19.26	(18.35)	10	4	80
Number of environmental/GHG assurance engagements undertaken as an assurance team leader (n = 13)	15.31	(15.90)	10	2	60
Training hours on assurance for GHG emissions (n = 13)	22.75	(17.46)	20	5	60

3.4.3 Administration of Research Instrument

A contact person in each of the three assurance firms in Australia was approached to request their participation. The contact person in Firms A and B were sent 20 research instruments each and were asked to distribute the questionnaires among employees at all levels who were involved in GHG assurance engagements. Participants were asked to mail the completed research instruments directly to the researcher via reply-paid envelopes. For Firm C, 14 research instruments were distributed by the researcher to participants attending a GHG assurance training session. After these participants completed the research instruments, they sealed the research instrument in an envelope and returned the envelopes directly to the researcher. Respondents were asked to answer each question frankly and anonymously in a cover letter. They were asked not to identify themselves, their firms, or their clients. All participants were assured that their responses would be kept completely confidential and that only aggregate results would be reported.

3.4.4 Analysis of Research Instrument

The first two parts of the research instrument consisted of repeated measures for the two cases: more effective team engagement and less effective team engagement. Both univariate and an ordinary least squares regression clustered by respondent and industry were used to test the hypotheses for this study. Because each respondent in this study provided two GHG assurance engagement cases and these engagements could be categorised into the same industry, two-way standard error clustering (Petersen 2009) was used to account for two dimensions within cluster correlation: respondent and industry sector. This technique was used in recent auditing studies (e.g., Lim and Tan 2010; Chen et al. 2013). Adopting this technique instead of including dummy variables to represent each respondent (e.g., Gibbins and Trotman 2002; Fargher et al. 2005) prevents the degrees of freedom from reducing by about half.

3.4.4.1 Research Model

Three ordinary least square regressions were employed to test the research question and hypotheses for the present study. Figure 3.2 illustrates the relationships of interest that were examined through six hypotheses.

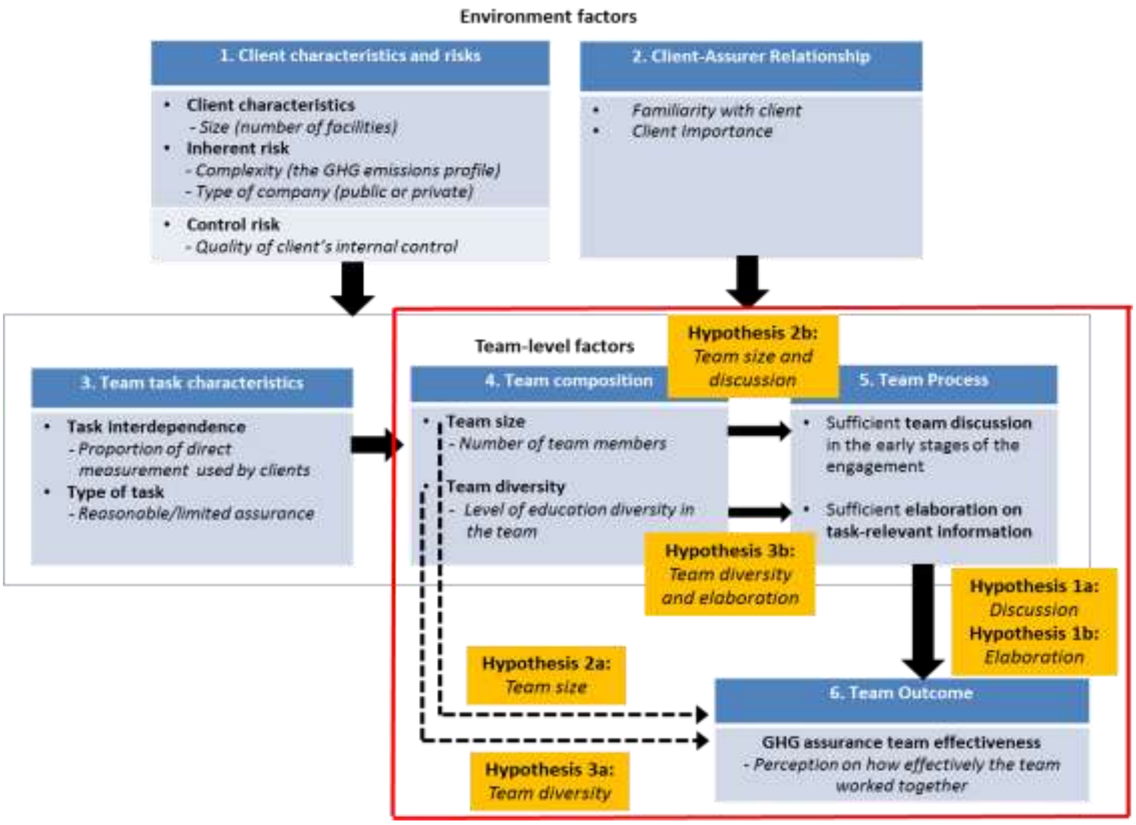


FIGURE 3.2 The Relationships of Interest for the Testable Hypotheses

Model 1 was employed to address four hypotheses: H1a, H1b, H2a, and H3a. Hypotheses 1a and 1b predict that team processes (i.e., *DISCUSS*[13] and *ELABORATE*) positively affect MDGHGTs' effectiveness (denoted as *TEAMEF*). Hypotheses 2a and 3a predict that team composition affects *TEAMEF*, with *TEAMSIZE* negatively affecting *TEAMEF* and *DIVERSITY* positively affecting *TEAMEF*. The model shown below included these variables and additional control (i.e., *DIRECT*, *TASK*, *SIZE*, *COMPLEX*, *PUBLIC*, *AVGIC*, *FAMILIAR*, and *IMPORTANCE)* and demographic (i.e., *GHGYEAR* and *TRAINING*) variables considered relevant to the GHG assurance context. The model tested was as follows:

[13] All variables are defined in Section 3.4.4.2

TEAMEF = f(***DISCUSS, ELABORATE, TEAMSIZE, DIVERSITY***, *DIRECT, TASK, SIZE, COMPLEX, PUBLIC, AVGIC, FAMILIAR, IMPORTANCE, GHGYEAR, TRAINING)* (1)

The second model was employed to address Hypothesis 2b, which predicts that team composition, specifically *TEAMSIZE*, negatively affects *DISCUSS*. The third model was employed to address Hypothesis 3b, which predicts that another team composition variable, *DIVERSITY*, positively affects *ELABORATE*. The models shown below include these variables and additional control and demographic variables considered relevant to the GHG assurance context. The following regression models were used:

DISCUSS = f(***TEAMSIZE***, *DIVERSITY, DIRECT, TASK, SIZE, COMPLEX, PUBLIC, AVGIC, FAMILIAR, IMPORTANCE, GHGYEAR, TRAINING)* (2)

ELABORATE = f(*TEAMSIZE*, ***DIVERSITY***, *DIRECT, TASK, SIZE, COMPLEX, PUBLIC, AVGIC, FAMILIAR, IMPORTANCE, GHGYEAR, TRAINING)* (3)

3.4.4.2 Variables

TEAMEF is an overall rating by the study participant of how effectively the GHG assurance team worked together on the engagement measured on a nine-point scale (1 = did not work well together; 9 = worked very well together). This measure therefore indicates *perceived team effectiveness*.

Team Process Variables

DISCUSS is a rating of the sufficiency of team discussion time to share diverse information in the early stages of engagement measured on a nine-point scale (1 = there was not enough time; 9 = there was more than enough time). Prior literature on teams suggest that extending the discussion time of diverse groups in the start-up phase provides greater opportunity for diverse information to be shared (Larson et al. 1994; van Knippenberg et al. 2004). Therefore, having sufficient discussion time in the early stages of engagement is expected to be positively related to GHG assurance team effectiveness.

ELABORATE is a rating of the sufficiency of elaboration and integration of different information from different team members measured on a nine-point scale (1 = there was not enough; 9 = there was more than enough). Recent studies suggest that the elaboration of task-relevant information moderates the effects of informational diversity on MDT performance (van Knippenberg et al. 2004; van Ginkel et al. 2009). The effectiveness of MDGHGTs is expected to increase when the sufficiency of elaboration and integration of different perspectives increases.

Team Composition Variables

TEAMSIZE is the number of members on the team. The empirical evidence is mixed regarding the relationship between team size and team effectiveness (e.g., West et al. 2003; Fay et al. 2006; Curral et al. 2001; LePine et al. 2008) While large MDTs may benefit from the broader knowledge and expertise their members bring to the team, they may face more communication problems compared to small MDTs. Because MDT members are likely to have different mindsets and use different professional language, the communication problem could be accentuated in large MDTs. Therefore, MDGHGT effectiveness and the level of perceived sufficiency of discussion time are expected to decrease as the team size increases. Further, team size usually depends on the client and task characteristics. For example, more team members will be required when undertaking large assurance engagements (O'Keefe et al. 1994; Hackenbrack and Knechel 1997).

DIVERSITY is the level of educational diversity in the team. Educational diversity refers to the extent to which a team comprises members with different majors or disciplines (Dahlin et al. 2005; Shin and Zhou 2007). It is measured using Blau's (1977) index of heterogeneity: $1 - \Sigma(\mathrm{Pi})^2$, where P*i* is the proportion of team members with discipline *i* (e.g., Wiersema and Bantel 1992; Shin and Zhou 2007). When the index is higher, educational diversity on the team will be higher. Nine areas of educational background are represented in the sample. However, the majority of team members have accounting, engineering/science, or both accounting and engineering/science degrees. Therefore, the Blau's index value is calculated based on these three categories. To account for each team members' contribution to the GHG assurance engagement, the proportion of team members within each category (Pi) is calculated by weighting the number of team members by their degree of involvement. The degree of involvement is an overall rating of how much each team member was involved in the engagement and is measured on three levels: low (coded 1), medium (coded 2), or high (coded 3). The level of educational diversity is expected to increase the level of information elaboration and thus enhance MDGHGTs' effectiveness (van Knippenberg et al. 2004).

Control variables

Two variables control for task characteristics factors, four variables control for client characteristics and risk factors, and two variables control for the client–assurer relationship. These control variables are discussed below.

Task Characteristics Variables

DIRECT is the proportion of direct measurement of emissions compared to other estimation techniques the client used to quantify their GHG data. The direct measurement is normally used in Scope 1 emission engagements; thus, significant technical/scientific knowledge is required to measure or approximate the reported and assured levels of GHG emissions from various emission sources. Therefore, a higher proportion of direct measurement used by clients may increase the degree to which accountant practitioners depend on non-accountant practitioners to execute the task. Although the high degree of task interdependence enhances the effectiveness of teams comprising educationally homogeneous members (Stewart and Barrick 2000) and educationally heterogeneous members (e.g., Jehn et al. 1999), the relationship between task interdependence and team composition is not well established in the literature.

TASK is a dummy variable that is coded 1 if the engagement is a reasonable level of assurance and 0 if it is a limited assurance engagement. Prior literature suggests that different types of tasks performed by teams require different degrees of interaction and coordination among team members and thus may affect team effectiveness (e.g., Jehn et al. 1999; Stewart and Barrick 2000). Since reasonable and limited assurances have different levels of assurance and types of evidence obtained, the levels of interaction and coordination required between accountant and non-accountant practitioners could be different. Thus, whether the teams are engaged in a reasonable or limited assurance may affect the team composition and the MDGHGTs' effectiveness. However, the relationships between types of assurance engagements and team effectiveness are not well established in the literature.

Client Characteristics and Risks Variables

SIZE is the size of the client. Client size is measured by the number of facilities the client has. The number of facilities is measured in three categories: single facility, two to five facilities, and more than five facilities. Clients with a single facility can be considered "small" clients, while clients with two to five and more than five facilities can be considered "medium" and "large" clients, respectively. Therefore, this variable has been treated as a dummy variable coded 1 if the client is large and 0 if the client is small or medium. Unlike in financial audit studies in which the total revenue or total assets are used to proxy size, the number of client facilities is a more appropriate measure for client size in the GHG assurance context. While clients are required to report all GHG emissions from facilities under their operational control, ISAE 3410 requires GHG assurance teams to perform various procedures at the facility level (IFAC 2012a). Therefore, more facilities may pose more work for the engagement, which thus increases audit effort (O'Keefe et al. 1994; Hackenbrack and Knechel 1997; Gibbins and Trotman 2002; Knechel et al. 2009) and increases the demand for interaction and cooperation among team members (Jehn et al. 1999).

COMPLEX is a participant rating of the relative complexity of a client's GHG emissions profile compared to similar GHG assurance engagements measured on a nine-point scale (1 = much lower profile complexity than others; 9 = much higher profile complexity than others). Because the complexity of the GHG assurance depends heavily on the scope of emissions and the methods used to quantify such emissions (IFAC 2012a), this variable should capture the overall complexity of GHG assurance engagements. Complexity increases the effort auditors give the client (O'Keefe et al. 1994; Hackenbrack and Knechel 1997), and decreases auditors' performance (Simnett and Trotman 1989; Simnett 1996). Studies on work team diversity (Jehn et al. 1999;

Bowers et al. 2000) find that teams with members who possess different information perform better on more complex tasks. However, the different representations possessed and different languages used by members from different disciplines may cause a lack of understanding or misunderstandings among team members, which then inhibit team effectiveness. Thus, complexity is expected to be negatively associated with the MDGHGTs' effectiveness.

PUBLIC is a dummy variable coded 1 if the client is a public company and 0 otherwise. Whether the client is publicly held has been found to increase the audit effort (e.g., O'Keefe et al. 1994). In the voluntary environmental disclosures setting, public companies face greater political scrutiny and regulatory pressures on the firm (Brammer and Pavelin 2008), and the amount of work and audit effort required when undertaking assurance engagements for public companies is greater than those for private companies (O'Keefe et al. 1994). Therefore, when the client is publicly held, GHG assurance teams are expected to work more effectively together.

AVGIC is a composite score of the five internal control quality ratings assessed by participants measured on a nine-point scale. The five quality ratings are as follows:

Quality 1: Rating of the client's systems development to capture and record GHG data (1 = not at all developed; 9 = very well developed);

Quality 2: Rating of the client's report preparer(s) availability to the GHG assurance team for the engagement (1 = low availability; 9 = high availability);

Quality 3: Rating of the client's report preparer(s) capability compared to similar GHG assurance engagements (1 = much lower capabilities; 9 = much higher capabilities than others);

Quality 4: Rating of the client's report preparer(s) quality of work compared to similar GHG assurance engagements (1 = much lower quality work than others; 9 = much higher quality work than others); and

Quality 5: Rating of the client's report preparer(s) quality of documentation compared to similar GHG assurance engagements (1 = much lower quality work than others; 9 = much higher quality work than others).

Because the five ratings of internal control quality are strongly correlated with each other (Spearman's correlation coefficient of 0.462 to 0.744, all p = 0.000), it is considered appropriate to combine these ratings together as one variable. A factor analysis is conducted on these five ratings, which yielded one factor. To combine these five quality ratings, the individual scores for each rating are multiplied by its factor loading (factor loadings = 0.109, 0.092, 0.296, 0.296, and 0.208 for items 1 to 5, respectively). Since all factor loadings are scaled by 1, the five ratings can then be combined into *AVGIC*. The prior evidence on the effect of internal control quality on team effectiveness is somewhat mixed. O'Keefe et al. (1994) and Hackenbrack and Knechel (1997) do not find any evidence that the client's internal control affects audit effort, while Gibbins and Trotman (2002) find that the quality of preparers and preparer's work negatively affect audit effort.

Client-Assurer Relationship Variables

FAMILIAR is the familiarity with the client measured by the number of previous GHG assurance engagements undertaken for the client. Prior involvement with the client on an audit engagement, particularly long audit tenure, could lead auditors to become too familiar with their previous work, thus putting less effort into subsequent audit engagements (Tan 1995; Favere-Marchesi and Emby 2005; Carey and Simnett 2006). Therefore, familiarity with the client is expected to be negatively associated with MDGHGT effectiveness.

IMPORTANCE is a rating of the client's relative importance to the respondent's assurance firm measured on a nine-point scale (1 = low importance; 9 = high importance). Although economic dependence posed by important clients is expected to reduce audit quality (DeAngelo 1981), the prior literature finds no evidence that dependence on client fees compromises audit quality (Chung and Kallapur 2003; Larcker and Richardson 2004; Li 2009). Moreover, auditors report more conservatively for more important clients (Reynolds and Francis 2001). Therefore, client importance is expected to be positively related to MDGHGTs' effectiveness.

Demographic variables

The dependent variables, including team effectiveness (*TEAMEF*), sufficiency of team discussion (*DISCUSS*), and sufficiency of elaboration and integration of different perspectives (*ELABORATE*), are subject to the respondents' varying interpretations and experiences. Therefore, four demographic variables are considered to assess whether any of the demographic information collected is a significant determinant of the ratings of effectiveness or sufficiency of team discussion and elaboration: number of years involved in conducting GHG assurance (*GHGYEAR*), number of GHG assurance engagements undertaken (*GHGNUM)*, number of GHG assurance engagements undertaken as a team leader *(LEADNUM)*, and number of training hours on GHG assurance *(TRAINING)*. However, after testing for multicollinearity, *GHGNUM* and *LEADNUM* are dropped from all models. Tests for multicollinearity indicate that a high level of multicollinearity is present when these two variables are included in the models (*VIF* = 7.120 to 7.583 for *GHGNUM* and 5.155 5.421 for *LEADNUM*). Therefore, only *GHGYEAR* and *TRAINING* have been included in the research models.

3.5 Results

The results are presented in the following order: Section 3.5.1 presents the descriptive statistics, correlation matrix (Spearman's rank correlation coefficients), and frequencies for each team effectiveness component. This section is organised by the analysis framework shown in Figure 3.1, beginning with client characteristics and risks, then the client–assurer relationship, task characteristics, team composition, and team processes. The results are presented for the full sample and the two subsamples for more effective and less effective teams. Section 3.5.2 presents the regression results for Hypotheses 1 to 3.

3.5.1 Descriptive Statistics

3.5.1.1 Client Characteristics and Risks

Table 3.2 presents the client characteristics and risk factors, including the client's number of facilities, industry sector, and type of company. The table reports descriptive statistics for the full sample and the sub-sample of engagements in which teams worked more effectively together (hereafter, more effective teams) and the sub-sample of engagements in which teams worked less effectively together (hereafter, less effective teams)[14]. As shown in this table, 62 percent of the clients have more than five facilities, with exactly the same percentages found in both more and less effective teams. The un-tabulated chi-square test of independence shows no difference in the frequencies of client size between the two subsamples. Again, no differences are found between more and less effective teams for the frequencies of clients' industry sectors ($p = 0.317$, un-tabulated) and type of company ($p = 0.517$, un-tabulated). The majority of the clients are in mining (31 percent) and production sectors (29 percent), and a majority are public companies (78 percent).

Table 3.3 shows that the mean number of the participants' ratings for the client's GHG emissions profile complexity compared to similar GHG engagements they had undertaken is 6.448 and the average composite score for the quality of client's internal control (i.e., systems to capture and record GHG data and the quality of GHG inventory preparers; AVGIC) is 5.568, both on a nine-point scale (1 = low; 9 = high).

While no differences are found between the two sub-samples in the complexity of the client's GHG emissions profile, the quality of client's internal control between more effective (6.481) and less effective (4.655) teams is significantly different ($p = 0.000$). The Spearman's rank correlations shown in Table 3.4 confirm that the quality of client's internal control is positively and significantly correlated with the perceived team effectiveness ($r = 0.666$, $p = 0.000$). Other than internal control quality, none of the client characteristics and risk factors are significantly correlated with perceived team effectiveness.

[14] *TEAMEF* mean scores for the effective and less effective teams are 7.500 (range from 5.0 to 9.0) and 4.052 (range from 1.0 to 7.0), respectively. The difference in perceived team effectiveness between these two teams is highly significant ($z = -6.321$, $p = 0.000$, two-tailed).

TABLE 3.2 Frequencies for Client Characteristics

1. Client's number of facilities

	Full Sample		More Effective Teams		Less Effective Teams	
	Number of Responses	**Percent of 58 Responses**	**Number of Responses**	**Percent of 29 Responses**	**Number of Responses**	**Percent of 29 Responses**
Single Facility	6	10%	3	10%	3	10%
2–5 Facilities	16	28	8	28	8	28
5+ Facilities	36	62	18	62	18	62
Total	58	100	29	100	29	100

2. Client's industry sector

	Full Sample		More Effective Teams		Less Effective Teams	
	Number of Responses	**Percent of 58 Responses**	**Number of Responses**	**Percent of 29 Responses**	**Number of Responses**	**Percent of 29 Responses**
Mining	18	31%	10	34%	8	28%
Production	17	29	9	31	8	28
Utilities	9	16	5	17	4	14
Transport	5	9	1	3	4	14
Property	3	5	3	10	0	0
Finance	2	3	1	3	1	3
Services	2	3	0	0	2	7
Government	1	2	0	0	1	3
Not answered	1	2	0	0	1	3
Total	58	100	29	100	29	100

3. Client's type of company

	Full Sample		More Effective Teams		Less Effective Teams	
	Number of Responses	**Percent of 58 Responses**	**Number of Responses**	**Percent of 29 Responses**	**Number of Responses**	**Percent of 29 Responses**
Public	45	78%	24	83%	22	76%
Private	13	22	5	17	7	24
Total	58	100	29	100	29	100

TABLE 3.3 Descriptive Statistics

	Full Sample (N = 58)			More Effective Teams (N = 29)			Less Effective Teams (N = 29)			Prob. of Mean Difference [a]
	Mean	**Median**	**Std. Deviation**	**Mean**	**Median**	**Std. Deviation**	**Mean**	**Median**	**Std. Deviation**	
TEAMEF	5.776	6.000	2.090	7.500	7.000	0.991	4.052	4.000	1.325	0.000***
DISCUSS	4.526	4.750	1.879	5.224	5.000	1.962	3.828	4.000	1.525	0.001***
ELABORATE	4.879	5.000	1.834	6.052	6.500	1.560	3.707	4.000	1.250	0.000***
TEAMSIZE	5.120	5.000	1.728	5.379	5.000	1.916	4.862	5.000	1.505	0.372
DIVERSITY	0.417	0.469	0.169	0.450	0.469	0.166	0.385	0.469	0.170	0.042**
DIRECT	0.529	0.700	0.384	0.559	0.750	0.390	0.500	0.600	0.382	0.285
COMPLEX	6.448	7.000	1.632	6.397	7.000	1.423	6.500	7.000	1.842	0.957
AVGIC	5.568	5.896	1.556	6.481	6.405	0.976	4.655	4.690	1.500	0.000***
FAMILIAR	1.160	0.000	1.642	1.170	0.000	1.891	1.140	1.000	1.382	0.979
IMPORTANCE	6.870	7.000	1.856	7.207	7.500	1.745	6.534	7.000	1.932	0.112
GHGYEAR	4.815	4.000	3.714	4.815	4.000	3.750	4.815	4.000	3.750	1.000
TRAINING	10.500	0.000	16.208	10.500	0.000	16.369	10.500	0.000	16.369	1.000
Categorical variables:										
TASK	0.430	0.000	0.500	0.480	0.000	0.509	0.380	0.000	0.494	
SIZE	0.620	1.000	0.490	0.620	1.000	0.490	0.620	1.000	0.490	
PUBLIC	0.780	1.000	0.421	0.830	1.000	0.384	0.720	1.000	0.455	
MIXBG [b]	0.190	0.000	0.395	0.280	0.000	0.455	0.100	0.000	0.310	
REG [b]	0.530	1.000	0.503	0.660	1.000	0.484	0.410	0.000	0.501	

Notes:
*, **, *** Probability of difference is significant at the 0.10, 0.05, and 0.01 levels (2-tailed), respectively. Definitions of variables are presented in Appendix 2.
[a] Probability of difference using a paired-samples t-test (Wilcoxon Signed Ranks Test) for difference of means.
[b] Alternative/additional variables tested in the sensitivity/additional analysis.

TABLE 3.4 Correlation Matrix

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1.TEAMEF	1																
2.DISCUSS	**.480**	1															
3.ELABORATE	**.692**	.596	1														
4.TEAMSIZE	.095	.094	.226	1													
5.*DIVERSITY*	.136	.037	.120	.261	1												
6.DIRECT	.115	-.152	.031	-.196	.052	1											
7.TASK	.171	.006	.100	.242	.139	.086	1										
8.SIZE	.014	.111	.000	.060	-.204	-.259	-.181	1									
9.COMPLEX	-.134	-.122	-.071	-.095	-.126	-.009	.072	-.028	1								
10.PUBLIC	.042	.170	.077	.051	.110	-.040	-.033	.091	.243	1							
11.AVGIC	**.666**	.232	**.439**	.126	-.107	.053	.174	.025	-.064	-.016	1						
12.FAMILIAR	-.138	.091	-.003	.018	-.101	-.086	-.261	.335	.200	.273	.117	1					
13.IMPORTANCE	.140	.096	.030	-.204	-.097	-.102	-.167	-.091	.093	**.385**	.252	.109	1				
14.GHGYEAR	-.030	-.115	-.149	-.117	-.097	-.233	-.071	.300	.021	-.033	-.041	-.052	.194	1			
15.TRAINING	-.026	.196	-.070	.235	.307	**-.376**	.199	.046	.067	-.014	.011	.198	-.014	-.036	1		
16.MIXBG	.198	-.039	.061	.033	**.599**	.038	.201	-.075	-.015	.049	.133	-.009	-.142	-.006	.213	1	
17.REG	.275	.068	.165	.169	.160	.264	**.463**	**-.373**	.110	-.087	.172	-.130	-.187	**-.454**	-.036	.275	1

Notes:
Figures in bold are significant at the 0.01 level (two-tailed). Definitions of variables are presented in Appendix 2.

3.5.1.2 Client–Assurer Relationship

Table 3.3 presents descriptive statistics for the client–assurer relationship factors, including familiarity with client (*FAMILIAR)* and client importance (*IMPORTANCE)*. Overall, the average number of previous GHG assurance engagements undertaken for the client is 1.160 engagements, and the relative importance of the client to the assurance firm is 6.870 on a nine-point scale (1 = low importance; 9 = high importance). No significant differences are found for familiarity with the client or client importance between more effective and less effective teams (both $p > 0.10$). The Spearman's rank correlation presented in Table 3.4 shows that neither *FAMILIAR* nor *IMPORTANCE* is significantly correlated with team effectiveness (both $p > 0.10$).

Table 3.5 shows additional information regarding other services the assurance firm provided for the same client. In the full sample, the assurance firm undertaking the GHG assurance engagement for the client also acts in other capacities for this client, mostly as a financial statement auditor (61 percent). The same pattern holds for both more effective and less effective teams, and the frequencies between the two teams are not significantly different ($p = 0.693$).

TABLE 3.5 Frequencies for Client–Assurer Relationship

Other services the assurance firm provided for the client:	**Full Sample**		**More Effective Teams**		**Less Effective Teams**	
	Number of Responses	**Percent of 58 Responses**	**Number of Responses**	**Percent of 29 Responses**	**Number of Responses**	**Percent of 29 Responses**
Financial statement audit	35	61%	19	66%	16	55%
Advisory services	3	5	1	3	2	7
Internal/ regulatory audit	3	5	2	7	1	3
None	17	29	7	24	10	35
Total	58	100	29	100	29	100

3.5.1.3 Task Characteristics

As shown in Table 3.6, the measurement methodologies the clients used to quantify their GHG data are mostly a mix of direct measurement and estimation techniques. A high percentage of direct measurement (75 to 100 percent) is used by approximately half of the clients (43 percent of the full sample). Table 3.3 reveals that the proportion of direct measurement (*DIRECT*), compared to other estimation techniques used by the clients in this study, is around 0.529 with no significant differences between the more effective and less effective teams. Table 3.4 shows that the Spearman's rank correlation between *DIRECT* and team effectiveness is 0.115 (p = 0.393). The proportion of direct measurement used by the client is not significantly correlated with team effectiveness.

TABLE 3.6 Frequencies for GHG Assurance Task Characteristics

1. Percentage of direct measurement methodologies used by the client to quantify their GHG data

	Full Sample		**More Effective Teams**		**Less Effective Teams**	
	Number of Responses	**Percent of 58 Responses**	**Number of Responses**	**Percent of 29 Responses**	**Number of Responses**	**Percent of 29 Responses**
0%	13	22%	6	21%	7	24%
1–25%	7	12	3	10	3	10
26–50%	5	9	3	10	3	10
51–75%	7	12	2	7	2	7
75–100%	25	43	14	48	14	48
Not answered	1	2	1	3	0	0
Total	58	100	29	100	29	100

2. GHG assurance engagement type

	Full Sample		**More Effective Teams**		**Less Effective Teams**	
	Number of Responses	**Percent of 58 Responses**	**Number of Responses**	**Percent of 29 Responses**	**Number of Responses**	**Percent of 29 Responses**
Limited	33	57%	15	52%	18	62%
Reasonable	25	43	14	48	11	38
Total	58	100	29	100	29	100

With regard to the type of engagement, Table 3.6 shows that the majority of GHG assurance engagements recalled by participants in this study are limited assurance engagements (57 percent of the full sample). No significant differences are found in the frequencies of limited and reasonable assurance between more effective and less effective teams (p = 0.426, un-tabulated). The Spearman's rank correlation in Table 3.4 shows that the type of engagement (*TASK*) is not significantly correlated with team effectiveness (r = 0.171, p = 0.201).

3.5.1.4 Team Composition

For each of the cases, respondents provided details relating to the GHG assurance team members. The full sample frequencies for the MDGHGT characteristics in Table 3.7 show that the majority of teams are composed of two to five members (67 percent) and that three to five accountant practitioners (45 percent) and three to five non-accountant practitioners are on the team (51 percent). All accountant practitioners have accounting backgrounds, while the non-accountant practitioners mostly have science (45 percent) or engineering (37 percent) backgrounds. Respondents also reported that 8 percent of the accountant practitioners have both accounting and other backgrounds (i.e., science and engineering). Overall, 44 percent of the GHG assurance team members have accounting backgrounds, 52 percent have other backgrounds (e.g., science and engineering), and 4 percent have mixed backgrounds (i.e., accounting and other).

Table 3.3 shows that the means for the total number of members in each team (*TEAMSIZE)* is 5.120 for the full sample, 5.379 for more effective teams, and 4.862 for less effective teams. No significant difference is found in the team size between effective teams and less effective teams (p=0.372). The mean level of educational diversity within the teams (*DIVERSITY)* is 0.417 for the full sample, 0.450 for more effective teams, and 0.385 for less effective teams. The mean numbers indicate that effective teams are more diverse than less effective teams, and this difference is statistically significant (p = 0.042). As Table 3.4 shows, neither *TEAMSIZE* nor *DIVERSITY* is correlated with team effectiveness on a univariate basis (r = 0.095, p = 0.478 and r = 0.136, p = 0.307, respectively).

TABLE 3.7 Frequencies for GHG Assurance Team Composition

1. Number of members in the team

	Full Sample		More Effective Teams		Less Effective Teams	
	Number of Responses	**Percent of 58 Responses**	**Number of Responses**	**Percent of 29 Responses**	**Number of Responses**	**Percent of 29 Responses**
2–5	39	67%	19	66%	20	69%
6–10	18	31	9	31	9	31
More than 10	1	2	1	3	0	0
Total	58	100	29	100	29	100

2. Number of accountant practitioners in the team

	Full Sample		More Effective Teams		Less Effective Teams	
	Number of Responses	**Percent of 58 Responses**	**Number of Responses**	**Percent of 29 Responses**	**Number of Responses**	**Percent of 29 Responses**
0	5	9%	2	7%	3	10%
1–2	25	43	13	45	12	42
3–5	26	45	12	41	14	48
More than 5	2	3	2	7	0	0
Total	58	100	29	100	29	100

3. Educational background of accountant practitioners (ACC) in the team

	Full Sample		More Effective Teams		Less Effective Teams	
	Number of ACC	**Percent of ACC**	**Number of ACC**	**Percent of ACC**	**Number of ACC**	**Percent of ACC**
Accounting	132	92%	71	89%	61	95%
Accounting and other	12	8	9	11	3	5
Total	144	100	80	100	64	100

4. Number of non-accountant practitioners in the team

	Full Sample		More Effective Teams		Less Effective Teams	
	Number of Responses	**Percent of 58 Responses**	**Number of Responses**	**Percent of 29 Responses**	**Number of Responses**	**Percent of 29 Responses**
0	1	2%	0	0%	1	3%
1–2	26	45	13	45	13	45
3–5	30	51	16	55	14	49
More than 5	1	2	0	0	1	3
Total	58	100	29	100	29	100

5. Educational background of non-accountant practitioners (Non-ACC)

	Full Sample		More Effective Teams		Less Effective Teams	
	Number of Non-ACC	**Percent of Non-ACC**	**Number of Non-ACC**	**Percent of Non-ACC**	**Number of Non-ACC**	**Percent of Non-ACC**
Science	69	45%	36	46%	33	39%
Engineering	57	37	32	41	36	42
Engineering and Science	3	2	2	3	1	1
Other	13	8	5	6	8	9
Unknown	11	7	4	5	7	8
Total	153	100	79	100	85	100

6. Overall educational background of members in the team

	Full Sample		More Effective Teams		Less Effective Teams	
	Number of members	**Percent of members**	**Number of members**	**Percent of members**	**Number of members**	**Percent of members**
Accounting	132	44%	71	45%	61	41%
Other	153	52	79	50	85	57
Accounting and Other	12	4	9	6	3	2
Total	297	100	159	100	149	100

3.5.1.5 Team Processes

To obtain evidence on the effect of different team processes on GHG assurance team effectiveness, respondents provided two sets of ratings: “sufficient discussion time in the early stages of the engagement” and “sufficient elaboration and integration of different information” (each on a nine-point scale: 1 = not enough; 9 = more than enough). The means for these two ratings are reported in Table 3.3. The overall mean for the sufficiency of discussion time (*DISCUSS)* is 4.526, while the means for effective and less effective teams are 5.224 and 3.828, respectively. A highly significant difference is found for the sufficiency of the discussion between the teams that worked more effectively together and the teams that worked less effectively together (p = 0.000). Considering the sufficiency of elaboration of different information and perspectives *(ELABORATE)*, the mean rating is close to the scale midpoint (mean 4.879, refer to Table 3.3). The means for effective and less effective teams are 6.052 and 3.707, respectively. The elaboration rating scores between these two teams are highly significantly different (p = 0.000).

Table 3.4 reports the Spearman correlations between team effectiveness and the two sets of team process ratings and shows that both *DISCUSS* and *ELABORATE* are positively and highly correlated with team effectiveness (r = 0.480 and r = 692, respectively, both p = 0.000). The results indicate that team members work more effectively together when they discuss and elaborate more on different information and perspectives.

Table 3.4 also presents the correlations between *DISCUSS* and *ELABORATE* with other variables. *DISCUSS* and *ELABORATE* are found to be highly correlated with each other (r = 0.596, p = 0.000). They are not found to be correlated with any other variables except for the average quality of clients’ internal control (*AVGIC* and *DISCUSS:* r = 0.232, p = 0.079; *AVGIC* and *ELABORATE*: r = 0.439, p = 0.001).

3.5.2 Regression Results

3.5.2.1 Hypothesis 1: The Effect of Team Processes on Team Effectiveness

Hypothesis 1 examines the effect of team process variables, including the perceived sufficiency of discussion in the early stages of engagement *(DISCUSS)* and elaboration on different information and perspectives *(ELABORATE),* on MDGHGTs' effectiveness. Model 1, as shown in Section 3.4.4.1, is used to address this hypothesis. However, because *DISCUSS* and *ELABORATE* are highly correlated (r = 0.596, p = 0.000), Model 1 was also run with either *DISCUSS* (Model 1a) or *ELABORATE* (Model 1b).

Hypothesis 1a predicts that MDGHGTs perceived to have sufficient discussion time in the early stages of engagement are more likely to work effectively together. Model 1 (Table 3.8) shows that, in the presence of *ELABORATE* and all the other variables, *DISCUSS* has no significant effect on team effectiveness (t=0.42, p=0.340, one-tailed). However, in the absence of *ELABORATE,* Model 1a shows that *DISCUSS* has a significant positive effect on team effectiveness (t = 3.96, p = 0.000, one-tailed) albeit with a significantly lower predictive power (R^2 reduced from 73.28 to 67.41). Therefore, H1a is only supported without *ELABORATE.* This finding indicates that having sufficient discussion in the early stages of engagement could enhance MDGHGTs' effectiveness. However, the fact that *DISCUSS* is no longer significant with the inclusion of *ELABORATE* and the high correlation between these two processes suggests that having sufficient discussion in the early stages of engagement is part of the elaboration process.

Hypothesis 1b hypothesises that MDGHGTs perceived to have sufficient information elaboration of different perspectives are more likely to work effectively together. Table 3.8 shows that either in the presence of *DISCUSS* (Model 1) or in the absence of *DISCUSS* (Model 1b), *ELABORATE* consistently show a highly significant positive effect on team effectiveness (Model 1: t = 2.75, p = 0.005; Model 1b: t = 3.40, p = 0.001, one-tailed). This result provides support for H1b and thus indicates that having sufficient information elaboration on different perspectives helps MDGHGTs work more effectively together.

Two control variables, *AVGIC* and *FAMILIAR*, also show a significant association with team effectiveness. The results reported in Table 3.8 show for all three models that in the presence of team process factors, the quality of the client's internal control is positively associated with team effectiveness while familiarity with the client is negatively associated with team effectiveness.

3.5.2.2 Hypothesis 2: The Effect of Team Size on Team Effectiveness and Team Processes

Hypothesis 2 investigates the effect of team size on team effectiveness and team processes. The relationship between team size (*TEAMSIZE*) and team effectiveness (H2a) and the relationship between team size and the perceived sufficiency of discussion in the early stages of engagement (*DISCUSS*) (H2b) are examined. As mentioned in Section 3.4.4.1, Model 1 is used to address Hypothesis 2a and Model 2 is used to address Hypothesis 2b. The results are shown below.

Hypothesis 2a predicts that the effectiveness of MDGHGTs is likely to decrease when the number of team members increases. In addition to Model 1, two supplementary models—Model 1c and 1d (Table 3.9)—were employed to test the direct relationship between team size and team effectiveness (i.e., in the absence of team processes). For Model 1c, Model 1 was altered by excluding team process variables (*DISCUSS* and *ELABORATE).* The result from Model 1c was used later in the secondary analysis to explore whether there is a potential mediation relationship between team size, sufficient discussion, and team effectiveness. For Model 1d, Model 1c was altered by excluding *DIVERSITY* to see if the result holds in the absence of another team composition variable—the level of educational diversity in the team.

TABLE 3.8 Regression Results for the Team Process Factors on the Effectiveness of MDGHGTs

	Model 1		Model 1a (without *DIVERSITY*)		Model 1b (without *TEAMSIZE*)	
Intercept	-1.313	(-0.90)	-2.199	(-1.91)*	-1.172	(-0.90)
DISCUSS	**0.044**	**(0.42)**	**0.430**	**(3.96)*****		
ELABORATE	**0.484**	**(2.75)*****			**0.513**	**(3.40)*****
TEAMSIZE	**-0.071**	**(-0.66)**	**-0.002**	**(-0.02)**	**-0.078**	**(-0.76)**
DIVERSITY	**0.731**	**(0.59)**	**2.258**	**(1.60)**	**0.633**	**(0.54)**
DIRECT	0.503	(1.38)	0.573	(1.27)	0.480	(1.33)
TASK	-0.167	(-0.54)	0.093	(0.27)	-0.192	(-0.58)
SIZE	0.563	(1.33)	0.596	(1.38)	0.562	(1.37)
COMPLEX	-0.047	(-0.52)	-0.005	(-0.05)	-0.053	(-0.61)
PUBLIC	0.098	(0.26)	0.026	(0.08)	0.108	(0.29)
AVGIC	0.752	(4.48)***	0.867	(6.77)***	0.754	(4.46)***
FAMILIAR	-0.370	(-2.73)***	-0.305	(-2.01)*	-0.378	(-2.91)***
IMPORTANCE	0.035	(0.59)	-0.030	(-0.72)	0.040	(0.63)
GHGYEAR	0.014	(0.50)	0.031	(0.90)	0.011	(0.46)
TRAINING	0.009	(0.85)	-0.006	(-0.61)	0.011	(1.38)
RESPONDENT [a]	YES		YES		YES	
INDUSTRY [a]	YES		YES		YES	
Number of observations[b]	51		51		51	
Regression R^2 (%)	73.28		67.41		73.24	

Notes: *, **, *** Significant at $p < 0.1$, $p < 0.05$, and $p < 0.01$, respectively, one-tailed for the variables of interest (highlighted in bold) and two-tailed for others. Definitions of variables used in the regression are presented in Appendix 2.

[a] The standard errors clustered by respondent and industry sector are used to compute the *t*-statistics. For each variable, the regression coefficient is reported, followed by the *t*-statistics in parentheses.

[b] Missing values are found for *DIRECT(1)*, *GHGNUM(6)* and *LEADEXP(6)*, thus seven observations are deleted.

Overall, team size is not significantly associated with team effectiveness. This lack of significant association holds in the presence and absence of the team process variables *DISCUSS* and *ELABORATE* (Table 3.8, Model 1: t = -0.66, p = 0.256; Table 3.9, Model 1c: t = -0.51, p = 0.308, one-tailed) and in the absence of *DIVERSITY* (Table 3.9, Model 1d: t = -0.24, p = 0.406, one-tailed). Therefore, although team size seems to affect team effectiveness in the expected direction, H2a is not supported. The fact that team size is not related to team effectiveness in the presence or absence of team processes suggests that team processes have no potential mediation effect on the relationship between team size and team effectiveness[15].

Hypothesis 2b predicts that the perceived sufficiency of discussion in the early stages of engagement is likely to decrease when the number of team members increases. In addition to Model 2, Model 2a is employed to test if the result in Model 2 holds in the absence of *DIVERSITY*. The regression results in Table 3.10 show that the association between *TEAMSIZE* and *DISCUSS* is of marginal significance and is negative with and without *DIVERSITY* included in the model (Model 2: t = -1.67, p = 0.051; Model 2a: t = -1.60, p = 0.059, respectively, one-tailed). Therefore, H2b is marginally supported, which indicates that MDGHGT members are less likely to think they have enough discussion in the early stages of engagement when teams are larger.

In examining the control factors again in the presence of the team composition factors (and in the absence of the team process factors), *AVGIC* is found to be significantly and positively associated with team effectiveness, while *FAMILIAR* is found to be significantly and negatively associated with team effectiveness (Table 3.9). *AVGIC* is also positively associated with the perceived sufficiency of discussion in the early stages of engagement as well as *TRAINING* (Table 3.10). However, *GHGYEAR* is found to be negatively associated with the perceived sufficiency of discussion in the early stages of engagement.

[15]According to the three-variable path diagram developed by Baron and Kenny (1986), four conditions must be met to confirm the mediation relationship (see Section 3.7 for more details). One of the four conditions is that the independent variable must be significantly associated with the outcome variable in the absence of the presumed mediator. Because no significant relationship is found between *TEAMSIZE* and team effectiveness (in neither the presence nor absence of the team processes), the four conditions are not met.

TABLE 3.9 Regression Results for the Team Composition Factors on the Effectiveness of MDGHGTs

		Model 1c	Model 1d (without *DIVERSITY*)	Model 1e (without *TEAMSIZE*)
Intercept	β_0	-0.428 (-0.55)	0.752 (0.58)	-0.663 (-1.43)
TEAMSIZE	$\boldsymbol{\beta_1}$	**-0.065 (-0.51)**	**-0.020 (-0.24)**	
DIVERSITY	$\boldsymbol{\beta_2}$	**2.122 (3.58)*** **		**2.007 (4.80)*** **
DIRECT	β_3	0.186 (0.37)	0.259 (0.44)	0.243 (0.55)
TASK	β_4	-0.087 (-0.18)	-0.139 (-0.31)	-0.109 (-0.23)
SIZE	β_5	0.609 (2.17)**	0.397 (0.88)	0.616 (2.26)
COMPLEX	β_6	-0.059 (-0.90)	-0.113 (-1.21)	-0.063 (-1.04)
PUBLIC	β_7	0.141 (0.56)	0.199 (0.65)	0.081 (0.47)
AVGIC	β_8	1.048 (15.44)***	1.037 (11.69)***	1.021 (12.31)***
FAMILIAR	β_9	-0.386 (-3.00)***	-0.389 (-2.32)**	-0.373 (-2.77)***
IMPORTANCE	β_{10}	-0.005 (-0.07)	-0.021 (-0.26)	0.018 (0.48)
GHGYEAR	β_{11}	-0.022 (-0.49)	-0.011 (-0.24)	-0.023 (-0.52)
TRAINING	β_{12}	0.008 (1.40)	0.010 (1.18)	0.006 (0.75)
RESPONDENT [a]		YES	YES	YES
INDUSTRY [a]		YES	YES	YES
Number of observations[b]		51	51	51
Regression R^2 (%)		59.65	57.33	59.47

Notes: *, **, *** Significant at $p < 0.1$, $p < 0.05$, and $p < 0.01$, respectively, one-tailed for the variables of interest (highlighted in bold) and two-tailed for others. Definitions of variables used in the regression are presented in Appendix 2.

[a] The standard errors clustered by respondent and industry sector are used to compute the *t*-statistics. For each variable, the regression coefficient is reported, followed by the *t*-statistics in parentheses.

[b] Missing values are found for *DIRECT(1)*, *GHGNUM(6)* and *LEADEXP(6)*, thus seven observations are deleted.

TABLE 3.10 Regression Results for the Team Composition Factors on the Perceived Sufficiency of Discussion in the Early Stages of Engagement

		Model 2		Model 2a (without *DIVERSITY*)		Model 2b (without *TEAMSIZE*)	
Intercept	β_0	4.122	(3.35)***	3.946	(4.18)***	3.593	(3.98)***
TEAMSIZE	$\boldsymbol{\beta_1}$	**-0.146**	**(-1.67)***	**-0.152**	**(-1.60)***		
DIVERSITY	$\boldsymbol{\beta_2}$	**-0.316**	**(-0.22)**			**-0.575**	**(-0.42)**
DIRECT	β_3	-0.900	(1.39)	-0.911	(-1.41)	-0.772	(-1.28)
TASK	β_4	-0.418	(-0.64)	-0.411	(-0.66)	-0.468	(-0.68)
SIZE	β_5	0.0300	(0.14)	0.0616	(0.36)	0.0459	(0.21)
COMPLEX	β_6	-0.127	(-1.29)	-0.119	(-0.98)	-0.135	(-1.40)
PUBLIC	β_7	0.267	(0.86)	0.258	(0.85)	0.133	(0.59)
AVGIC	β_8	0.421	(6.32)***	0.423	(5.75)***	0.360	(3.80)***
FAMILIAR	β_9	-0.187	(-1.22)	-0.187	(-1.24)	-0.159	(-0.96)
IMPORTANCE	β_{10}	0.0591	(0.61)	0.0615	(-0.64)	0.111	(1.41)
GHGYEAR	β_{11}	-0.124	(-2.36)**	-0.126	(-2.29)**	-0.126	(-2.24)**
TRAINING	β_{12}	0.0326	(2.91)***	0.0323	(3.22)***	0.0282	(3.43)***
RESPONDENT [a]		YES		YES		YES	
INDUSTRY [a]		YES		YES		YES	
Number of observations[b]		51		51		51	
Regression R^2 (%)		30.93		30.85		29.41	

Notes: *, **, *** Significant at p < 0.1, p < 0.05, and p < 0.01, respectively, one-tailed for the variables of interest (highlighted in bold) and two-tailed for others. Definitions of variables used in the regression are presented in Appendix 2.

[a] The standard errors clustered by respondent and industry sector are used to compute the *t*-statistics. For each variable, the regression coefficient is reported, followed by the *t*-statistics in parentheses.

[b] Missing values are found for *DIRECT(1)*, *GHGNUM(6)* and *LEADEXP(6)*, thus seven observations are deleted.

This finding indicates that MDGHGTs are more likely to perceive that they have sufficient discussion when the quality of the client's internal control becomes higher and when they have more training in GHG assurance. However, they perceive that they do not have enough discussion in the early stages when they have more experience conducting GHG assurance engagements.

3.5.2.3 Hypothesis 3: The Effect of Team Diversity on Team Effectiveness and Team Processes

Hypothesis 3 investigates the effect of educational diversity on team effectiveness and team processes. The direct relationship between the level of diversity in the team and team effectiveness (H3a) and the direct relationship between diversity and the perceived sufficiency of elaboration on different information and perspectives (H3b) are tested. As mentioned in Section 3.4.4.1, Model 1 is used to address Hypothesis 3a and Model 3 is used to address Hypothesis 3b. The results are shown below.

Hypothesis 3a hypothesises that the level of educational diversity in the team (*DIVERSITY*) is positively associated with MDGHGTs' effectiveness. In addition to Model 1, Models 1c and 1e (in Table 3.9) were employed to test the direct relationship between team diversity and team effectiveness (i.e., in the absence of team processes). Team process variables (*DISCUSS* and *ELABORATE)* were dropped from Model 1 to form Model 1c. The results from Model 1c were used later in the secondary analysis to explore the potential mediation relationship between team diversity, sufficient elaboration, and team effectiveness. Model 1c was then altered by excluding *TEAMSIZE* to see if the results hold.

Model 1 in Table 3.8 shows that, in the presence of all variables, *DIVERSITY* is not significantly associated with team effectiveness ($t = 0.59$, $p = 0.280$, one-tailed). However, Model 1c in Table 3.9 shows that, in the absence of team process variables, *DIVERSITY* is highly significantly and positively associated with team effectiveness ($t = 3.58$, $p = 0.001$, one-tailed). This result also holds in the absence of *TEAMSIZE* (Table 3.9, Model 1e: $t = 4.80$, $p = 0.000$, one-tailed). Therefore, H3a is conditionally supported—that is, the level of educational diversity in MDGHGTs is only significantly related to team effectiveness when team processes are not considered. This finding suggests a potential mediation effect of team processes on the relationship between team diversity and team effectiveness. The mediation relationship between these variables is further examined in the secondary analysis in Section 3.7.

Hypothesis 3b hypothesises that the perceived sufficiency of elaboration on different information and perspectives (*ELABORATE*) is likely to increase when the level of educational diversity in the team (*DIVERSITY*) increases. In addition to Model 3, Model 3b is employed to test if the results in Model 3 hold in the absence of *TEAMSIZE*. Table 3.11 reveals that *DIVERSITY* is significantly and positively associated with *ELABORATE*, regardless of whether or not *TEAMSIZE* is taken into account (Model 3: $t = 1.79$, $p = 0.040$ and Model 3b: $t = 1.90$, $p = 0.032$, respectively, one-tailed). These results indicate that H3b is supported, that is, MDGHGT members are more likely to perceive that there is sufficient elaboration on different information and perspectives when teams become more educationally diverse.

Table 3.11 also shows that *AVGIC* is the only control variable significantly associated with the perceived sufficiency of elaboration on different information and perspectives. None of the demographic variables are associated with the perceived sufficiency of information elaboration.

The results for all hypotheses are summarised in Table 3.12.

TABLE 3.11 Regression Results for the Team Composition Factors on the Perceived Sufficiency of Elaboration on Different Information and Perspectives

		Model 3		Model 3a (without *DIVERSITY*)		Model 3b (without *TEAMSIZE*)	
Intercept	β_0	1.449	(1.47)	3.062	(2.79)***	1.542	(1.69)*
TEAMSIZE	$\boldsymbol{\beta_1}$	**0.026**	**(0.28)**	**0.086**	**(0.80)**		
DIVERSITY	$\boldsymbol{\beta_2}$	**2.901**	**(1.79)****			**2.946**	**(1.90)****
DIRECT	β_3	-0.572	(-0.75)	-0.472	(-0.57)	-0.594	(-0.85)
TASK	β_4	0.204	(0.31)	0.134	(0.21)	0.213	(0.34)
SIZE	β_5	0.092	(0.19)	-0.198	(-0.75)	0.089	(0.19)
COMPLEX	β_6	-0.013	(-0.13)	-0.086	(-0.82)	-0.012	(-0.12)
PUBLIC	β_7	0.064	(0.18)	0.144	(0.31)	0.0877	(0.20)
AVGIC	β_8	0.573	(6.04)***	0.557	(7.03)***	0.583	(7.55)***
FAMILIAR	β_9	-0.014	(-0.13)	-0.019	(-0.13)	-0.019	(-0.15)
IMPORTANCE	β_{10}	-0.088	(-0.53)	-0.109	(-0.67)	-0.097	(-0.63)
GHGYEAR	β_{11}	-0.064	(-1.33)	-0.049	(-0.98)	-0.064	(-1.33)
TRAINING	β_{12}	-0.005	(-0.43)	-0.003	(-022)	-0.0046	(-0.37)
RESPONDENT [a]		YES		YES		YES	
INDUSTRY [a]		YES		YES		YES	
Number of observations[b]		51		51		51	
Regression R^2 (%)		59.65		57.33		59.47	

Notes: *, **, *** Significant at $p < 0.1$, $p < 0.05$, and $p < 0.01$, respectively, one-tailed for the variables of interest (highlighted in bold) and two-tailed for others. Definitions of variables used in the regression are presented in Appendix 2.

[a] The standard errors clustered by respondent and industry sector are used to compute the *t*-statistics. For each variable, the regression coefficient is reported, followed by the *t*-statistics in parentheses.

[b] Missing values are found for *DIRECT(1)*, *GHGNUM(6)* and *LEADEXP(6)*, thus seven observations are deleted.

TABLE 3.12 Summary of Results

Factor	Hypothesis	Expectation	Outcome
Team Processes	1a	Multidisciplinary GHG assurance teams that **perceive that they have sufficient discussion time** to share diverse information and perspectives **in the early stages** of engagement will **work more effectively together**.	Conditionally Supported (Only significant in the absence of elaboration)
	1b	Multidisciplinary GHG assurance teams that **perceive that they have sufficient information elaboration** of different information and perspectives will **work more effectively together**.	Supported
Team Size	2a	The **effectiveness** of multidisciplinary GHG assurance teams will **decrease when the number of team members increases**.	Not supported
	2b	Multidisciplinary GHG assurance teams will be **less likely to perceive that they have sufficient discussion time** to share diverse information and perspectives **in the early stages** of engagement **when the number of team members increases**.	Marginally supported
Team Diversity	3a	The **effectiveness** of multidisciplinary GHG assurance teams will **increase when the level of diversity in the team increases**.	Conditionally Supported (Only significant in the absence of team processes)
	3b	Multidisciplinary GHG assurance teams will be **more likely to perceive that they have sufficient elaboration** on different information and perspectives **when the level of diversity in the team increases**.	Supported

3.6 Sensitivity Analyses

3.6.1 Client Size

As previously mentioned in section 3.4.4.2, client size is a control variable and is measured by the number of facilities the client has. The number of facilities is measured categorically: single facility (small), two to five facilities (medium), and more than five facilities (large). Table 3.2 reveals that most of the clients recalled in this study had more than five facilities (62 percent). In the main analysis, the dummy variable "*SIZE"* was included to control for the client size effect. This variable is coded 1 if the client has more than five facilities and 0 if the client has less than five facilities, Another way to treat this variable is as a continuous variable by assuming equal intervals between 1 (small), 2 (medium), and 3 (large). This alternative measure was included in the regression as a sensitivity analysis.

The regression results for GHG assurance team effectiveness after substituting a dummy variable for client size with a continuous variable remain mostly unchanged to those in the main analysis (un-tabulated). The only differences are that in Model 1, the "*DIRECT*" coefficient, which is a control variable, becomes significant ($t = 1.80$, $p = 0.400$, one-tailed) and the "*SIZE*" coefficient becomes highly significant ($t = 3.61$, $p = 0.001$, one-tailed) compared with the insignificant coefficients in the main analysis. Therefore, the results stay largely consistent with the main analysis and provide the same support for the hypotheses as previously reported.

3.6.2 Alternative Team Diversity Variable

In the main analyses, the level of educational diversity in MDGHGTs (*DIVERSITY*) is measured using Blau's (1977) index of heterogeneity. Because the majority of team members have accounting, engineering/science, or mixed backgrounds (accounting and engineering/science degrees), the Blau's index value is calculated based on these three categories. Given that having mixed backgrounds also reflects the diversity in MDGHGTs' educational backgrounds, *DIVERSITY* could be alternatively treated as a dichotomous variable in which teams with members who have mixed backgrounds are coded as "1" and teams with no members with mixed backgrounds are coded as "0" (hereafter, *MIXBG*). To test whether the results are sensitive to the measure of diversity, all models were rerun and *DIVERSITY* was replaced with *MIXBG.*

The regression results show that the results for H1a and H1b are unchanged (un-tabulated), with *ELABORATE* remaining significant both in the presence and absence of *DISCUSS* (t =2.18, p = 0.018 and t = 3.00, p = 0.003, respectively, one-tailed). *DISCUSS* remain insignificant (significant) in the presence (absence) of *ELABORATE* (t = 7.63, p = 0.000 and t = 0.79, p = 0.216, respectively, one-tailed). The results for H2a remain the same as for the main analyses and for other control variables (un-tabulated). However, the results for H2b, H3a, and H3b change after replacing *DIVERSITY* with *MIXBG.*

The marginally significant relationship between *TEAMSIZE* and *DISCUSS* becomes significant (H2b: t = -1.87, p = 0.035, one-tailed). Further, while *DIVERSITY* is significantly associated with team effectiveness only in the absence of team processes, *MIXBG* is significantly associated with team effectiveness both in the presence and absence of team processes (t = 4.61, p = 0.000, and t = 4.95, p = 0.000, respectively, one-tailed). Therefore, H3a is unconditionally supported. More importantly, the significant relationship between *DIVERSITY* and *ELABORATE* becomes insignificant after replacing *DIVERSITY* with *MIXBG* (t = -0.14, p = 0.447, one-tailed). These findings indicate that H3a and H3b are sensitive to the measure of diversity. In particular, whether or not team members have mixed backgrounds is significantly and positively associated with team effectiveness but is not significantly associated with the perceived sufficiency of elaboration on different information and perspectives.

3.6.3 Demographic Variables

Two demographic variables, number of years involved in conducting GHG assurance *(GHGYEAR)* and number of training hours on GHG assurance *(TRAINING),* were included in the main analyses to control for the differences in respondents' GHG assurance experience. As explained in Section 3.4.4.2, two other potential demographic variables were excluded due to multicollinearity concerns: number of GHG assurance engagements undertaken (*GHGNUM)* and number of GHG assurance engagements undertaken as a team leader *(LEADNUM)*. To see whether the results are sensitive to the inclusion of any of these demographic variables, all demographic variables were included in all models one at a time, starting with *GHGYEAR,* then *TRAINING,* then *GHGNUM,* and *LEADNUM.*

After re-running the analyses, the results examining the determinants of MDGHGT effectiveness remain unchanged from the main analysis except when only *GHGYEAR* was included in the model (un-tabulated). When only controlling for the number of years involved in conducting GHG assurance, the association between *DISCUSS* and team effectiveness changes from insignificant to highly significant ($t = 3.52$, $p = 0.001$, one-tailed) in the presence of *ELABORATE,* and *DIRECT* becomes marginally significant ($t = 1.91$, $p = 0.063$, two-tailed). However, after including other demographic variables in the model, *DISCUSS* and *DIRECT* are no longer significant, and *ELABORATE* remains highly significant with the inclusion of all demographic variables (un-tabulated). Thus, the results remain largely consistent and provide the same support as in the main analyses for H1a and H1b.

In sum, the results for H2 and H3 remain largely the same as in the main analyses. However, two noticeable changes occur in H2b and H3b. For H2b (un-tabulated), the direct relationship between *TEAMSIZE* and *DISCUSS* changes from marginally significant to significant after including all four demographic variables in the main model ($t = -1.69$, $p = 0.049$, one-tailed). For H3b, the relationship between *DIVERSITY* and *ELABORATE* becomes marginally significant after including *GHGNUM* and *LEADNUM* in the model ($t = 1.67$, $p = 0.051$, and $t = 1.63$, $p = 0.056$, respectively, one-tailed). These findings suggest that the relationships between team composition and team processes are sensitive to the inclusion of additional GHG assurance experience variables*: GHGNUM* and *LEADNUM*. However, because these two variables are highly correlated with *GHGYEAR* and *TRAINING*, these changes may have been caused by multicollinearity among these variables.

3.6.4 Insignificant Variables

Given the small sample size obtained in this study, a major concern is that the results could be affected by the lack of degrees of freedom. To test whether the results are sensitive to the change in the degrees of freedom, six control variables that are not significantly associated with the dependent variables in all models are dropped from the analyses: *DIRECT*, *TASK*, *SIZE*, *COMPLEX*, *PUBLIC*, and *IMPORTANCE*. The analyses for Hypotheses 1 to 3 were rerun without these six variables. The regression results show that the results for H1a and H1b remain the same (un-tabulated), with *ELABORATE* remaining highly significant both in the presence and absence of *DISCUSS* (t = 2.76, p = 0.004 and t = 3.39, p = 0.001, respectively, one-tailed). *DISCUSS* remains insignificant (significant) in the presence (absence) of *ELABORATE* (t = 1.43, p = 0.103 and t = 3.67, p = 0.001, respectively, one-tailed). The results for H2a, H3a, and H3b also remain as per the main analysis and the demographic variables (un-tabulated), except for H2b. For H2b, the marginally significant relationship between *TEAMSIZE* and *DISCUSS* becomes highly significant (t = -3.65, p = 0.001, one-tailed). Overall, the results are robust to the change in the degrees of freedom.

3.7 Secondary Analysis

As mentioned earlier, although the main research aim of this study is to explore factors that are related to MDGHGTs' effectiveness, the research framework suggests a potential mediating relationship between team composition, team processes, and team effectiveness. The results found for H1 and H3 suggest that mediation may be present between *DIVERSITY*, *ELABORATE*, and team effectiveness. Therefore, a secondary analysis was conducted to test for this mediating relationship by adopting the three-variable path diagram developed by Baron and Kenny (1986), which is the most frequently reported procedure for such analyses (Warner 2013).

The model shown in Figure 3.3 involves three causal paths feeding directly and indirectly into the outcome variable (*TEAMEF*): the direct effect of the independent variable (*DIVERSITY*, *path c*), the direct effect of the mediator (*ELABORATE*, *path b*), and the effect of the independent variable on the mediator (*DIVERSITY* to *ELABORATE*, *path a*). To test whether the presumed mediator variable actually serves as a mediator, all four of the following conditions must be met: (1) the independent variable must be significantly associated with the outcome variable (*path c*); (2) the independent variable must be significantly associated with the presumed mediator (*path a*); (3) the presumed mediator must be significantly associated with the outcome variable (*path b*); and (4) the association between the independent variable and the outcome variable is no longer significant (full mediation) or less significant (partial mediation) after controlling for the presumed mediator (Baron and Kenny 1986, p. 1176).

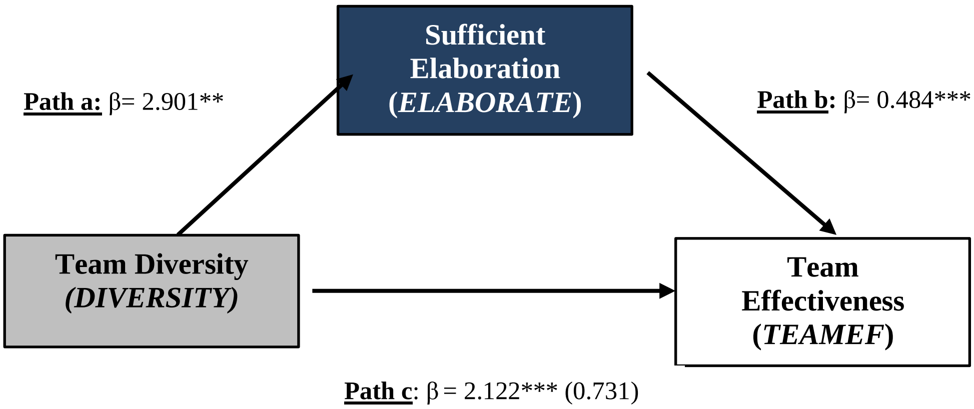


FIGURE 3.3 Mediating Relationship of Team Diversity, Sufficient Discussion in the Early Stages of Engagement, and Team Effectiveness.

Previously in the hypotheses testing, significant associations have been established between *ELABORATE* and *TEAMEF* (H1b: see Table 3.8, Model 1) and *DIVERSITY* and *ELABORATE* (H3b: see Table 3.11, Model 3). The results for H3a also show that in the absence of *ELABORATE*, *DIVERSITY* is highly significantly associated with *TEAMEF* (H3a: see Table 3.9, Model 1c), while in the presence of *ELABORATE*, *DIVERSITY* is no longer significant (see Table 3.8, Model 1). Since all four conditions are met, these results indicate that the sufficiency of elaboration on different information and perspectives fully mediates the relationship between the level of educational diversity in the team and the effectiveness of MDGHGTs[16].

[16] Although Baron and Kenny's (1986) Causal-Steps approach is the best-known procedure, Fritz and MacKinnon (2007) suggest that it has relatively lower statistical power than the Sobel test (Sobel 1982) and the bootstrapping approach (Preacher and Hayes 2004). However, when violations of the normality assumption occur and when the sample size is not very large, such as in the present study, the bootstrapping approach is more suitable than the Sobel test (Preacher and Hayes 2008). Thus, to formally

3.8 Additional Analyses

3.8.1 Regulatory Effect

Since respondents were asked to identify whether they undertook GHG assurance engagement on a regulatory or voluntary basis, additional analyses were conducted to test for a regulatory effect. Table 3.13 shows that just over half of the full sample assured their emissions due to regulatory requirements (53 percent). When compared to regulatory assurance observations, voluntary assurance observations are larger and are more likely to entail limited assurance engagements. The frequencies of client and GHG assurance engagement characteristics in Table 3.13 reveal that the top two industries for regulatory assurance companies are production (45 percent) and mining (32 percent), while the top two industries for voluntary assurance companies are mining (30 percent) and utilities (22 percent). To test for the regulatory effect, the dummy variable *REG* was included in all models. This variable was coded into 1 if the client undertook the GHG assurance engagement on a regulatory basis, and 0 if they undertook the GHG assurance engagement on a voluntary basis.

After controlling for the regulatory effect, the results remain largely the same for the variables of interest in H1, with only a few changes in the control variables (un-tabulated). These changes include *SIZE* (which becomes significant), *IMPORTANCE, GHGYEAR* (which becomes marginally significant), and *FAMILIAR* (which becomes less significant). The results for H2 and H3 also remain unchanged (un-tabulated) except for the control variables: *SIZE* and *TRAINING*, which become more significant when testing the regulatory effect on the direct relationship between team composition variables and team effectiveness (H2a and H3a).

---

test the significance of the indirect effect in the mediation model, Preacher and Hayes' (2004) bootstrapping approach was also adopted. Based on 1,000 bootstrapping samples, the results suggest that no significant mediation relationship occurs between *DIVERSITY*, *ELABORATION*, and *TEAMEF*. Because the confidence interval contains zero (indirect effect = 1.404, confidence interval at 95 percent: -0.678 to 3.486, un-tabulated), the indirect effect is not significant (Preacher and Hayes 2004). This result is not surprising given the small sample size. Repeating 58 observations 1,000 times could result in more sampling errors, which is the most important limitation of the bootstrapping approach (Haukoos and Lewis 2005). Because the bootstrapping approach assumes that the sample represents the variety and range of possible values in the population from which it was sampled, a small sample size may add another level of sampling error and result in invalid statistical estimations (Haukoos and Lewis 2005, p. 364). Therefore, the minimum sample size of 150 to 200 is recommended for testing mediation models (Warner 2013).

TABLE 3.13 Frequencies of Client and GHG Assurance Engagement Characteristics Relating to Regulatory and Voluntary GHG Assurance

1. Client's reason for undertaking GHG assurance	Full Sample		More Effective Teams		Less Effective Teams	
	Number of Responses	**Percent of 58 Responses**	**Number of Responses**	**Percent of 29 Responses**	**Number of Responses**	**Percent of 29 Responses**
Regulatory	31	53%	19	66%	12	41%
Voluntary	27	47	10	35	17	59
Total	58	100	29	100	29	100

2. Client's number of facilities	Full Sample		Regulatory		Voluntary	
	Number of Responses	**Percent of 58 Responses**	**Number of Responses**	**Percent of 31 Responses**	**Number of Responses**	**Percent of 27 Responses**
Single Facility	6	10%	3	10%	3	12%
2–5 Facilities	16	28	14	45	2	7
5+ Facilities	36	62	14	45	22	81
Total	58	100	31	100	27	100

3. Client's industry sector	Full Sample		Regulatory		Voluntary	
	Number of Responses	**Percent of 58 Responses**	**Number of Responses**	**Percent of 31 Responses**	**Number of Responses**	**Percent of 27 Responses**
Mining	18	31%	10	32%	8	30%
Production	17	29	14	45	3	11
Utilities	9	16	3	10	6	22
Transport	5	9	2	6	3	11
Property	3	5	2	6	1	4
Finance	2	3	0	0	2	7
Services	2	3	0	0	2	7
Government	1	2	0	0	1	4
Not answered	1	2	0	0	1	4
Total	58	100	31	100	27	100

4. Client's type of company	Full Sample		Regulatory		Voluntary	
	Number of Responses	**Percent of 58 Responses**	**Number of Responses**	**Percent of 31 Responses**	**Number of Responses**	**Percent of 27 Responses**
Public	45	78%	23	74%	22	81%
Private	13	22	8	26	5	19
Total	58	100	31	100	29	100

5. GHG assurance engagement type	Full Sample		Regulatory		Voluntary	
	Number of Responses	**Percent of 58 Responses**	**Number of Responses**	**Percent of 31 Responses**	**Number of Responses**	**Percent of 27 Responses**
Limited	33	57%	11	35%	22	81%
Reasonable	25	43	20	65	5	19
Total	58	100	31	100	27	100

Therefore, the additional analysis accounting for the regulatory effect provides the same support for the hypotheses as the main analyses in this study. The results suggest that whether the client undertakes GHG assurance engagement on a regulatory or voluntary basis does not affect the factors affecting MDGHGT effectiveness.

3.8.2 Factors Associated with MDGHGT Composition

According to the proposed framework in Figure 3.1, three components can directly influence the MDGHGT composition: client characteristics and risks, client–assurer relationship, and task characteristics. ISAE 3410 highlights the relationship between team composition and client risks. In addition, ISAE 3410 suggests that different areas of expertise are required to deal with quantifying and reporting emissions, particularly when the engagement is relatively complex (IFAC 2012a). As noted in ISAE 3410, the complexity of GHG assurance engagements varies depending on client characteristics (e.g., industry and number of facilities), scope of emissions (for details, please see footnote 9 in section 3.2.3.1), and GHG quantification methods. Thus, the specific client characteristics may influence team composition by decreasing or increasing the number of members and the combination of accountants and non-accountants in the team.

The nature of the task should also be considered when staffing a team. Previous literature finds that *task interdependence* and *task type* are significantly associated with team effectiveness (e.g., Stewart and Barrick 2000). ISAE 3410 recognises the need for scientific and engineering expertise, particularly when a client's GHG emissions profile involves significant Scope 1 emissions that result in a high proportion of direct measurements used to quantify the emissions (IFAC 2012a, para. A19). In such complex tasks, team members may need to rely more on the expertise of non-accountant practitioners on the team. Therefore, the degree of task interdependence may be high in this case.

According to ISAE 3410, GHG assurance practitioners will work on engagements providing one of two levels of assurance: reasonable and limited (IFAC 2012a). Given the higher level of assurance obtained in a reasonable assurance engagement compared to a limited assurance engagement, the amount of work and the complexity in the reasonable assurance engagement may affect staffing decisions. For example, more team members or more diverse members may be required to perform substantive tests in reasonable assurance engagements compared to limited assurance engagements.

While the relationships between team effectiveness and the three components of task characteristic, client characteristics and risks, and client–assurer relationship are well established, the empirical evidence on the relationships between these components and team composition is very limited. Therefore, additional analysis was conducted to test the effect of task characteristics, client characteristics and risks, and client–assurer relationship on MDGHGT composition.

Two additional models were adopted to explore the determinants underlying MDGHGT composition (i.e., *TEAMSIZE* and *DIVERSITY*). Task characteristics variables (i.e., *DIRECT* and *TASK*) were the variables of interest in this model, while other environmental factors were treated as control variables. Thus, the following regression models were used:

TEAMSIZE = f(***DIRECT***, ***TASK***, *SIZE*, *COMPLEX*, *PUBLIC*, *AVGIC*, *FAMILIAR*, *IMPORTANCE*, *GHGYEAR*, *TRAINING*) (4)

DIVERSITY = f(***DIRECT***, ***TASK***, *SIZE*, *COMPLEX*, *PUBLIC*, *AVGIC*, *FAMILIAR*, *IMPORTANCE*, *GHGYEAR*, *TRAINING*) (5)

Model 4 in Table 3.14 shows a negative relationship between *DIRECT* and *TEAMSIZE*; however, this is not statistically significant ($t = -1.65$, $p = 0.106$, two-tailed). The relationship between *TASK* and *TEAMSIZE*, however, is positive but not statistically significant ($t = 0.91$, $p = 0.370$, two-tailed).

TABLE 3.14 Effect of Task Characteristics and Environmental Factors on Team Size and Team Diversity

	TEAMSIZE		***DIVERSITY***	
	Model 4		Model 5	
Intercept	6.050	(6.53)***	0.652	(3.51)***
DIRECT	**-1.119**	**(-1.65)**	**0.008**	**(0.08)**
TASK	**0.609**	**(0.91)**	**-0.209**	**(-0.60)**
SIZE	-0.100	(-0.23)	-0.072	(-1.10)
COMPLEX	-0.043	(-0.478)	-0.197	(-1.91)*
PUBLIC	0.887	(1.45)	0.075	(1.77)*
AVGIC	2.938	(1.62)	0.004	(0.11)
FAMILIAR	-0.102	(-1.28)	-0.009	(-0.43)
IMPORTANCE	-0.362	(-1.51)	-0.160	(-1.18)
RESPONDENT [a]	YES		YES	
INDUSTRY [a]	YES		YES	
Number of observations[b]	57		57	
Regression R^2 (%)	23.29		10.92	

Notes:
*, **, *** Significant at $p < 0.1$, $p < 0.05$, and $p < 0.01$, respectively, two-tailed. Definitions of variables used in the regression are presented in Appendix 2.
[a] The standard errors clustered by respondent and industry sector are used to compute the *t*-statistics. For each variable, the regression coefficient is reported, followed by the *t*-statistics in parentheses.
[b] A missing value is found for *DIRECT*; therefore, one observation is deleted.

Table 3.14 also presents the factors associated with the level of educational diversity in the team. Model 5 reveals that *DIRECT* is positively and *TASK* is negatively associated with *DIVERSITY*. However, these relationships are not statistically significant (all p > 0.50, two-tailed). In addition to the team composition variables, two control variables, *COMPLEX* and *PUBLIC*, are found to be marginally significantly associated with *DIVERSITY* (t = -1.91, p = 0.063 and t = 1.77, p = 0.084, respectively, two-tailed). The level of educational diversity in the team is negatively related to the complexity of the client's GHG emissions profile (*COMPLEX*). That is, when more complexity is present in the client's emissions profile, the members in the team are less diverse. This may be because when the emissions profile becomes more complex, more non-accountant practitioners with scientific knowledge (i.e., engineers and scientists) are required (IFAC 2012a). Once the proportion of non-accountant practitioners becomes higher than the proportion accountant practitioners, the level of diversity becomes lower[17]. *DIVERSITY* is also positively related to whether or not the client is a public company (*PUBLIC)*. The results suggest that the level of educational diversity in the team will increase when the client is a public company.

3.8.3 Evaluation of GHG Assurance Engagement Teams

For both assurance engagements, respondents were asked to answer additional questions related to the evidence gathering and evaluation stages of assurance engagement and five open-ended questions: (1) List the factors that you feel contributed to the GHG assurance team working well together; (2) List the factors that you feel inhibited the GHG assurance team's ability to work well together; (3) Was there any other mechanism by which different information and perspectives from different team members was shared and integrated at any stage of the engagement? If so, please specify; (4) In your opinion, what factors, if they had been present, would have made the team work together better? These may be some of the factors listed above or others; and (5) Please comment on any issues not properly covered in the above questions or anything else you wish to raise regarding the team for this GHG assurance engagement. This narrative data was coded independently by the author and one research assistant, who is a PhD student with a financial audit background. A sophisticated coding procedure, similar to the procedure used by Gibbins and Trotman (2002), was used. To avoid any potential influence that other responses in the questionnaire may have had, all coding was done from a photocopied excerpt. The inter-rater coding agreement was 91.60 percent, and the *kappa* coefficient was 0.84, which represents a high level of reliability. All differences in the coding were discussed and resolved. Additional analyses were conducted to explore other factors that could enhance the effectiveness of MDGHGTs.

[17] An additional test was conducted to see if the complexity of the client's emissions profile is associated with the proportion of non-accountant practitioners in the team. The result confirms that the complexity of the client's emissions profile is positively associated with the proportion of non-accountant practitioners in the team ($t = 2.260$, $p = 0.029$, two-tailed). That is, the proportion of non-accountant practitioners increased when the complexity of the client's emissions profile increased.

3.8.4 Factors that Contributed to and Inhibited the Effectiveness of MDGHGTs

Table 3.15 presents the results for the responses to three open-ended questions: (1) List the factors that you feel contributed to the GHG assurance team working well together; (2) List the factors that you feel inhibited the GHG assurance team's ability to work well together; (3) In your opinion, what factors, if they had been present, would have made the team work together better? The answers for questions (1) and (3) were combined to form a comprehensive set of factors that were perceived to enhance MDGHGTs' effectiveness.

The responses for the three questions were coded into 17 categories, which are presented in Table 3.15. Factors that respondents infrequently identified that did not fall into the 17 categories were presented under the category "Other". Because responses to the three questions contained positive and negative versions of the same factors, they were analysed by counting either the positive or the negative versions of each factor identified by respondents in each case only once (e.g., Gibbins and Trotman 2002).

In all, respondents identified 217 non-repeated factors (average of 3.74 per respondent per case), with 137 factors considered to contribute to the team working well together and 80 factors considered to inhibit the team working well together (averages of 2.91 and 1.50 factors per respondent per case, respectively). Table 3.15 also reports the number of contributing and inhibiting factors identified by more effective and less effective teams.

Analysis of the 58 responses (full sample) indicated seven major factors (i.e., those with the highest frequency) that were perceived to contribute to the MDGHGT working well together (i.e., those accounting for 54 percent of the 137 contributing factors). These factors used positive wording, including 1) sufficient knowledge of and communication with the client; 2) team members with sufficient technical skills and experience; 3) involvement of all team members; 4) clearly defined roles and responsibilities of accountant/non-accountant practitioners in the team; 5) effective team communication; 6) sufficient planning; and 7) familiarity and good working relationship with other team members. Further, seven major factors were indicated that were perceived to inhibit the MDGHGT's ability to work well together (i.e., those accounting for 60 percent of the

80 inhibiting factors). These factors used negative wording, included 1) unclear roles and responsibilities of accountant/non-accountant practitioners in the team; 2) lack of time to prepare and work on site; 3) team members with lack of competence and experience; 4) lack of client cooperation, preparation, and data quality; 5) lack of understanding in the engagement or government requirements; 6) lack of understanding of other team members' knowledge and expertise; and 7) high complexity and difficulties of the subject matter and assurance processes. All seven major contributing and inhibiting factors are indicated in Table 3.15 in bold numbers.

The main differences in the "contributed factors" and "inhibited factors" identified between the more effective and less effective teams appear in the shaded areas (see Table 3.15). The more effective teams identified four contributed factors more frequently than the less effective teams: clearly defined roles and responsibility (7 vs. 2); strong knowledge, skills, and engagement of the team leader (6 vs. 1); familiarity with other team members (6 vs. 3); and mixture and integration of diverse skills expertise (5 vs. 2). With regard to the inhibiting factors, the less effective teams identified three factors more frequently than the more effective teams: lack of client cooperation, preparation, and data quality (6 vs. 1); team members with lack of competence and experience (5 vs. 2); and lack of knowledge, skills, and engagement of the team leader (3 vs. 0). However, no significant difference is found when comparing the proportion of each factor identified between the more effective teams and the less effective teams ($p > 0.10$).

The majority of these factors reflect some difficulties MDGHGTs face specifically related to the differences between accountant and non-accountant practitioners on the team. Because these practitioners are different in their knowledge and skill-sets, they may not understand each other (O'Dwyer 2011) or may not have sufficient technical knowledge to undertake GHG assurance engagements, especially when the subject matter and assurance processes are highly complex. Further, the roles and responsibilities of accountant and non-accountant practitioners on the team are not clearly defined. However, the findings also reveal ways to enhance MDGHGTs' effectiveness. In addition to addressing the difficulties mentioned earlier, MDGHGTs' effectiveness can be improved by having sufficient knowledge of and communication with the client; sufficient client cooperation and preparation; clear assurance guidance and methodology; sufficient planning and time to prepare and work on-site; and team leaders with strong knowledge and engagement.

TABLE 3.15 Factors that Contributed to/Inhibited MDGHGTs' Ability to Work Effectively Together

Contributed/Inhibited factors	Number of Factors Identified by Respondents					
	Full Sample (Percent of 58 responses)		**More Effective Teams (Percent of 29 responses)**		**Less Effective Teams (Percent of 29 responses)**	
	Contributed	**Inhibited**	**Contributed**	**Inhibited**	**Contributed**	**Inhibited**
1. Sufficient/lack of knowledge of and communication with the client	**16 (27.6%)**	4 (6.9%)	7 (24.1%)	2 (6.9%)	9 (31.0%)	2 (6.9%)
2. Team members with sufficient technical skills and experience/lack of competence and experience	**12 (20.7%)**	**7 (12.1%)**	6 (20.7%)	2 (6.9%)	6 (20.7%)	5 (17.2%)
3. Involvement/lack of involvement of all team members	**10 (17.2%)**	4 (6.9%)	5 (17.2%)	2 (6.9%)	5 (17.2%)	2 (6.9%)
4. Clearly defined roles and responsibilities/unclear roles and responsibilities of accountant/non-accountant practitioners	**9 (15.5%)**	**9 (15.5%)**	7 (24.1%)	5 (17.2%)	2 (6.9%)	4 (13.8%)
5. Effective team communication/lack of team communications	**9 (15.5%)**	4 (6.9%)	4 (13.8%)	2 (6.9%)	5 (17.2%)	2 (6.9%)
6. Sufficient planning/lack of planning	**9 (15.5%)**	2 (3.4%)	5 (17.2%)	0 (0.0%)	4 (13.8%)	2 (6.9%)
7. Familiarity/unfamiliarity with other team members	**9 (15.5%)**	1 (1.7%)	6 (20.7%)	1 (3.4%)	3 (10.3%)	0 (0.0%)
8. Strong/lack of knowledge, skills, and engagement of the team leader	7 (12.1%)	3 (5.2%)	6 (20.7%)	0 (0.0%)	1 (3.4%)	3 (10.3%)
9. Good mixture and integration of diverse expertise/lack of integration of diverse views	7 (12.1%)	1 (1.7%)	5 (17.2%)	0 (0.0%)	2 (6.9%)	1 (3.4%)
10. Shared/not shared common objectives, commitment, language, and core understanding of methodology (e.g., materiality)	6 (10.3%)	1 (1.7%)	3 (10.3%)	0 (0.0%)	3 (10.3%)	1 (3.4%)
11. Sufficient/lack of collaboration and support from staff inside and outside the team	6 (10.3%)	0 (0.0%)	3 (10.3%)	0 (0.0%)	3 (10.3%)	0 (0.0%)

TABLE 3.15 (Continued). Factors that Contributed to/Inhibited MDGHGTs' Ability to Work Effectively Together

Contributed/Inhibited factors	Number of Factors Identified by Respondents					
	Full Sample (Percent of 58 responses)		More Effective Teams (Percent of 29 responses)		Less Effective Teams (Percent of 29 responses)	
	Contributed	Inhibited	Contributed	Inhibited	Contributed	Inhibited
12. Clear guidance and methodology/lack of understanding in the engagement or government requirements	5 (8.6%)	**6 (10.3%)**	3 (10.3%)	2 (6.9%)	2 (6.9%)	4 (13.8%)
13. Understanding/lack of understanding of other team members' knowledge and expertise	5 (8.6%)	**5 (8.6%)**	3 (10.3%)	3 (10.3%)	2 (6.9%)	2 (6.9%)
14. Sufficient/lack of time to prepare and work on site	4 (6.9%)	**9 (15.5%)**	3 (10.3%)	6 (20.7%)	1 (3.4%)	3 (10.3%)
15. Willing to share knowledge and learn from each other/lack of knowledge sharing and integration between team members	4 (6.9%)	2 (3.4%)	3 (10.3%)	1 (3.4%)	1 (3.4%)	1 (3.4%)
16. Low/high complexity and difficulties of the subject matter and assurance processes	2 (3.4%)	**5 (8.6%)**	2 (6.9%)	2 (6.9%)	0 (0.0%)	3 (10.3%)
17. Sufficient/lack of client cooperation, preparation and data quality	1 (1.7%)	**7 (12.1%)**	1 (3.4%)	1 (3.4%)	0 (0.0%)	6 (20.7%)
18. Other (various)	16 (27.6%)	10 (17.2%)	7 (24.1%)	6 (20.7%)	9 (31.0%)	4 (13.8%)
Total	**137**	**80**	**79**	**35**	**58**	**45**

Note: The factors are presented in the order of factors that participants mainly suggested as contributing to the GHG assurance teams' ability to work effectively together. The percentages in the parentheses are calculated based on 58 responses for the full sample and 29 responses for the sub-samples. The seven major factors identified are indicated in bold numbers. The main differences between more effective and less effective teams are indicated in shaded areas. However, these differences are not statistically significant.

3.8.5 Team Processes Related to the Evidence Gathering and Evaluation Stages of the Assurance Engagement

Table 3.16 provides descriptive data on the extent of various team processes used in the evidence gathering and evaluation stages of GHG assurance engagement for the full sample and the two sub-samples. With regard to assurance procedures used to gather evidence, detailed substantive testing is mostly used (52.81 percent) followed by analytical procedures (26.43 percent) and tests of controls (21.09 percent). No significant differences are found in the extents to which more effective and less effective teams used these procedures (all $p > 0.58$). However, more effective teams had a clearer separation between the information search and the information processing stages ($z = -2.228$, $p = 0.026$, two-tailed) and more discussion on the information collected before final evaluations and decisions were made ($z = -3.762$, $p = 0.000$, two-tailed) compared to less effective teams.

TABLE 3.16 Team Processes Related to the Evidence Gathering and Evaluation Stages of GHG Assurance Engagement

The Evidence Gathering and Evaluation Processes	Full Sample (n = 58)		More Effective Teams (n = 29)		Less Effective Teams (n = 29)		Prob. Of Mean Diff.
	Mean	Median	Mean	Median	Mean	Median	
Percentage of tests of controls used	21.09	17.50	21.48	20.00	20.69	10.00	0.750
Percentage of substantive tests used	52.81	52.50	52.00	50.00	53.62	60.00	0.745
Percentage of analytical procedures used	26.43	20.00	26.48	20.00	26.38	20.00	0.583
The extent of a clear separation between the information search/ collection stage and the information processing/ decision-making stage	4.77	5.00	5.22	5.00	4.31	5.00	0.026**
The extent of discussion on the information collected before final evaluations and decisions were made	6.41	7.00	7.53	7.00	5.29	5.00	0.000***

*, **, *** Probability of difference is significant at the 0.10, 0.05, and 0.01 levels (2-tailed).
[a] Probability of difference using a paired-samples t-test (Wilcoxon Signed Ranks test) for difference of means.

In addition to team discussion, elaboration, evidence gathering, and evaluation processes discussed in the earlier sections, this study explores whether any other mechanism supports information sharing and integration within MDGHGTs. To explore this issue, respondents were asked, "Was there any other mechanism by which different information and perspectives from different team members was shared and integrated at any stage of the engagement? If so, please specify". The responses for this question were coded into nine categories, which are presented in Table 3.17.

TABLE 3.17 Other Mechanisms by which Different Information and Perspectives from Different Team Members were Shared and Integrated at Any Stage of the Engagement

Other mechanisms	Responses No. (Percent of 58 Responses)	Responses No. (Percent of 29 Responses)	
	Full Sample	More[a] Effective Teams	Less Effective Teams
1. No mechanism	28 (48.3%)	10** (34.5%)	18 (62.1%)
2. Meeting and team discussions	17 (29.3%)	12** (41.4%)	5 (17.2%)
3. Ongoing sharing of perspectives with others	5 (8.6%)	4 (13.8%)	1 (3.4%)
4. Internal calls and catch-ups	5 (8.6%)	3 (10.3%)	2 (6.9%)
5. Documents and issue logs	4 (6.9%)	2 (6.9%)	2 (6.9%)
6. Review process	3 (5.2%)	2 (6.9%)	1 (3.4%)
7. Training sessions	1 (1.7%)	1 (3.4%)	0 (0.0%)
8. Team leaders as a facilitator	1 (1.7%)	1 (3.4%)	0 (0.0%)
9. Past experience with sustainability issues	1 (1.7%)	1 (3.4%)	0 (0.0%)
Total	**65**	**36**	**29**

** Significant at $p < 0.05$

[a] The proportion of more effective and less effective teams are compared using a test of column proportion (z-test).

Of the 58 responses, 48 percent reported that no mechanism was in place to assist in sharing and integrating different perspectives within the team, while 29 percent reported that they did so through meetings and team discussions. Other mechanisms are also been reported, including ongoing sharing of perspectives with others; internal calls and catch-ups; documents and logs of issues; review process; training sessions; team leaders as a facilitator; and past experience with sustainability issues. The comparison between more effective and less effective teams shows that the proportion with no mechanisms reported in the less effective teams is significantly higher than in the more effective teams ($p < 0.05$), while the proportion of meeting and team discussions reported in the more effective teams is significantly higher than in the less effective teams ($p < 0.05$). These findings indicate that mechanisms to share and integrate different knowledge are important to the effectiveness of MDGHGTs, particularly meetings and team discussions. However, such mechanisms are often not established in practice.

3.9 Discussion and Limitations

While financial audit teams comprising members with accounting backgrounds have been studied extensively in the audit literature, very little research has been conducted on assurance teams comprising practitioners from different disciplines. Because newly emerging assurance services, such as GHG assurance, require diverse knowledge and skill-sets (i.e., environmental science and accounting/financial auditing), international assurance standard ISAE 3410 recognises the need to include practitioners from other disciplines into the assurance team to perform assurance engagements other than financial information audits. The importance of MDTs in the assurance context is accentuated specifically because GHG assurance is becoming a major service provided by leading assurance firms and a high-quality GHG assurance function is required to add the necessary credibility to successfully implement emissions reporting and/or trading schemes (PwC 2007; KPMG 2008, 2013; Simnett et al. 2009a). However, because the multidisciplinary nature of GHG assurance teams is new to assurance firms and the audit literature, how MDGHGTs should be operationalised and how to enhance the effectiveness of MDGHGTs remains unknown.

This study is the first to examine the factors contributing to MDGHGTs' effectiveness. Utilising team effectiveness frameworks suggested by previous research in psychology, this study examines opportunities to optimise the quality of the newly emerging GHG assurance services through various team-level factors. Given that little is known about how MDGHGTs are being operationalised in practice, a retrospective field research instrument is used to help identify factors perceived as affecting engagement performance. This is achieved by gauging GHG assurance team members' assessments of issues related to the client and GHG engagement characteristics, the client–assurer relationship, GHG assurance team composition, and team processes. As such, this study adds to the psychology literature and the limited literature relating to GHG assurance by deepening the understanding of the factors underlying the success of MDGHGTs.

Most team effectiveness frameworks (Guzzo and Dickson 1996; Cohen and Bailey 1997; Ilgen et al. 2005; Mathieu et al. 2008) suggest that team processes play a significant role in determining team effectiveness. Consistent with these frameworks, this study provides empirical evidence that team process factors are highly significantly associated with MDGHGTs' effectiveness. The results show that when team members perceive they have sufficient time for discussion in the early stages of engagement and have sufficient elaboration on diverse perspectives, the teams work more effectively together. The results are consistent with Larson et al.'s (1994) findings that providing sufficient discussion time in the start-up phase improves team effectiveness because it increases the chance that unique information will be shared and considered. The highly significant relationship between information elaboration and team effectiveness also supports van Knippenberg et al.'s (2004) proposition that information elaboration is an important process that drives the positive effects of diversity on team effectiveness. However, the significant relationship between sufficient discussion time and MDGHGTs' effectiveness is not significant in the presence of sufficient information elaboration, which indicates that discussion of different information and perspectives is an important component of information elaboration; this is in line with van Knippenberg et al.'s (2004) definition of information elaboration.

In addition to team process variables, the previous literature suggests that team composition variables (e.g., team size and diversity) are potential determinants of MDGHGTs' effectiveness. Although no direct relationship is found between team size and MDGHGTs' effectiveness, this study provides some evidence that team size negatively affects the perceived sufficiency of discussion time to share diverse information and perspectives. These findings suggest that larger MDGHGTs face more coordination and communication difficulties during the team discussion (Smith et al. 1994; LePine et al. 2008) than do smaller MDGHGTs. Moreover, different mindsets and professional language used among practitioners from different disciplines (van Someren et al. 1998) can slow down the discussion process (van Knippenberg et al. 2004), which then puts larger teams under more time pressure (Paulus et al. 2012).

With regard to team diversity, the results show a positive, significant relationship between each MDGHGTs' level of educational diversity measured using Blau's (1977) index of heterogeneity weighted by each team member's involvement level and team effectiveness. The findings also reveal that MDGHGT members are more likely to perceive that they have sufficient elaboration on different information and perspectives when the team becomes more diverse. The results are consistent with the view that cognitive diversity among MDGHGT members leads to a lack of understanding and disagreements on the task, which then forces the team members to thoroughly elaborate (exchange, discuss, and integrate) all task-relevant information. By thoroughly elaborating different information and perspectives, the benefits outweigh the negative effect of diversity (i.e., coordination and communication difficulties). These findings are consistent with the positive relationship between task conflict and MDT performance found by Jehn (1995) and Jehn et al. (1999). The results also support the superior role of information elaboration as a "primary process underlying the positive effects of diversity on group performance" as suggested by van Knippenberg et al.'s (2004, p.1012) categorisation-elaboration model (CEM).

Control variables, including environmental and demographic variables, also play significant roles in explaining GHG assurance team effectiveness. In terms of environmental variables, the quality of the client's internal control increases MDGHGTs' effectiveness, whereas familiarity with the client negatively affects MDGHGTs' effectiveness. In terms of demographic variables, MDGHGTs are more likely to perceive that they have enough discussion time to share different information and perspectives if they have more training on GHG assurance but are less likely to perceive so as they gain more GHG assurance experience over time. These findings indicate that, in addition to team-level factors, control risks arising from the client's report preparers, the client's systems to capture and record GHG data, and the assurer–client relationship also contribute to the effectiveness of MDGHGTs, while individual characteristics, including training and working experience, only effect team processes.

Additional analyses were conducted to explore the factors associated with MDGHGT composition. No significant relationship is found between the team composition variables (i.e., team size and diversity) and the task characteristics variables (i.e., task interdependence and task type). However, the complexity of the client's emissions profile is negatively significant with the team's level of educational diversity, and public companies are found to be positively and marginally significant. These findings indicate that teams become less diverse when the task becomes more complex. This could be explained by the fact that, in the context of GHG assurance, more complex tasks usually require more non-accountant practitioners with scientific knowledge and skill-sets (IFAC 2012a). Once the proportion of non-accountant practitioners increases, the level of educational diversity in the teams decreases.

This study also reveals additional factors that MDGHGTs think contribute to and inhibit the effectiveness of MDGHGTs. MDGHGT members suggest that the following important factors help improve their team effectiveness: clearly defined roles and responsibility; strong knowledge, skills and engagement of the team leader; familiarity with other team members; and mixture and integration of diverse skills expertise. On the other hand, the following factors inhibit MDGHGTs' effectiveness: lack of client cooperation; lack of preparation and data quality; lack of competence and experience of team members; and lack of knowledge, skills, and engagement of the team leader. Since MDT members are not only expected to bring diverse knowledge and perspective to bear on the task but also to utilise this information (van Knippenberg et al. 2004), it is crucial to the success of MDGHGTs to have a mechanism facilitating information sharing and integrating. Surprisingly, the results show that the majority of MDGHGTs do not have any such mechanisms, while a minority share and integrate their perspectives during meetings, team discussions, internal calls and catch-ups, documents, the review process, and training sessions.

The results of this study must be considered in light of several limitations. First, this study examines perceptions of team effectiveness, which are open to interpretation and are not externally validated. The prior team literature on psychology and auditing suggests a wide range of outcomes that are more objective than the one used in this study and that may be appropriate for assessing MDGHGTs' effectiveness, such as quantity and quality of ideas/risks/hypotheses generated (Osborn 1953; Trotman et al. 2009; Chen et al. 2014). To improve external validity, future research may consider using these objective outcome measures to assess MDGHGTs' effectiveness.

Second, the MDGHGTs were in the early stages of development when the data was collected. Consequently, this study is unable to examine over time stable process constructs (e.g., shared mental models) or other time-related variables (e.g., team tenure). Because team tasks and team processes can change over time (McGrath 1991), future research could examine the effect of these variables on MDGHGTs' effectiveness.

Third, this study does not find the expected relationship between task characteristic variables and team effectiveness using the proportion of direct measurement used by clients as a proxy for task interdependence and the type of assurance engagement as a proxy for task type. Future research could measure task interdependence using self-assessment measures (e.g., Stewart and Barrick 2000; Van der Vegt et al. 2000). Further, other characteristics and aspects of GHG assurance tasks could be explored following McGrath's (1984) Group Task Circumplex (e.g., planning, decision making, negotiating tasks) or Hackman and Oldham's (1975) task characteristics (i.e., task variety, task identity, task significance, autonomy, and feedback).

Fourth, the use of retrospective recall and experiential questionnaires could increase bias in the responses compared with other research methods, including archival and experimental research. For example, respondents may answer questions based on what they think happened in the engagement rather than what actually happened. However, at this stage of knowledge and development associated with this type of assurance service, this technique was identified as an appropriate mechanism to explore the issues outlined in this dissertation. To minimise potential recall biases, the approaches used in Gibbins and Trotman (2002) were adopted, including asking a broad range of questions about factors that could affect respondents' perceptions of team effectiveness. Although retrospective recall provides a larger amount of rich data, which allows a wider understanding of what happens in practice, future research is required to gain a complete understanding of MDGHGTs' effectiveness, specifically those employing laboratory and archival studies.

Finally, the study is limited by the small number of GHG assurance team members in Australia. The population of assurance practitioners is relatively small at the time of the study. However, this is an assurance service with significant growth prospects (GRI 2013; KPMG 2013). As we develop a better understanding of this assurance service, future research will be able to address the limitations raised in this study.

CHAPTER 4

STUDY TWO: EFFECTS OF DIFFERENT TEAM FORMATS ON THE PERFORMANCE OF MULTIDISCIPLINARY GREENHOUSE GAS ASSURANCE ENGAGEMENT TEAMS

4.1 Introduction

Assurance standards (IFAC 2012a, para. A42) and prior research (Huggins et al. 2011) highlights the need for multidisciplinary teams (MDTs) to undertake greenhouse gas (GHG) assurance engagements. Unlike a financial audit, the subject matter being assured in a GHG assurance engagement involves non-financial data, specifically, emissions data. The quantification of GHG emissions relies heavily on scientific estimation and uncertainties (Green and Li 2009; Simnett et al. 2009a). Because the knowledge and expertise required by GHG assurance practitioners goes beyond the traditional roles of accountants, assurance on GHG emissions are currently undertaken by practitioners such as accountants, engineers, and environmental scientists with different disciplinary backgrounds (Huggins et al. 2011; Nugent 2008). ISAE 3410 suggests the use of experts from various disciplines to deal with quantifying and reporting emissions, particularly when the engagement is relatively complex (IFAC 2012a).

ISAE 3410 requires multidisciplinary greenhouse gas assurance teams (MDGHGTs) to discuss during the planning stage the susceptibility of the entity to material misstatements in GHG statements due to fraud or error (IFAC 2012a, para. 29). Although it can be expected that interactions between educationally diverse team members could result in better decision making, empirical research on work team diversity has found mixed evidence of the advantages and disadvantages of MDTs in terms of team performance (van Knippenberg and Schippers 2007). On one hand, diverse members bring greater knowledge and skill-sets to the task that enhance team creativity and decision making (Guzzo and Dickson 1996; Williams and O'Reilly 1998). On the other hand, individuals with diverse educational backgrounds may have different frames of reference, professional language and problem-solving styles that impede the optimal sharing and integration of diverse ideas and information (van Knippenberg and Schippers 2007). Tension arises from different mindsets held by team members with different backgrounds and is documented in the sustainability assurance setting, in which accountant and non-accountant assurers work together on an engagement (O'Dwyer 2011).

To gain a better understanding of how to improve MDGHGTs' effectiveness, Study One employs a retrospective recall methodology in which GHG assurance practitioners provided their perceptions relating to factors they believed enhanced GHG team effectiveness. Study One finds that the team process is the most important factor underlying team effectiveness. The term "team process" is defined as a mechanism that integrates different knowledge and expertise possessed by team members and coordinates the effort to resolve task demands (Kozlowski and Ilgen 2006). Such a mechanism varies depending upon the team format used (Kerr and Tindale 2004). Therefore, different team formats have been examined in research on group performance and decision making in psychology (see Kerr and Tindale 2004 for reviews) and auditing (see Rich et al. 1997b and Nelson and Tan 2005 for reviews) in an attempt to improve team performance. While the specific guidance for GHG assurances provided in ISAE 3410 requires accountant and non-accountant practitioners to interact with each other through discussions to assess the risks of material misstatements, research has not yet examined what team format works well for these MDTs to deliver effective assessments of risks and thereby high-quality assurance.

The present study employs an experiment to explore ways of improving MDGHGTs' performance through three different types of team format: nominal, interacting, and review teams. In particular, the outcomes of these teams are compared for two risk assessment tasks: risk generation and risk selection. The research framework in Figure 4.1 illustrates the focus of this study. Of the six components of the MDGHGT effectiveness framework proposed in Study One (see Section 3.2.2), this study focuses on only three components: team composition, team processes, and team outcomes. The team composition of MDGHGTs is expected to lead to cognitive diversity between team members, while the team processes underlying different team formats are expected to have different effects on their team performances. To test these expectations, this study examines three main issues: (1) is there cognitive diversity in MDGHGTs working together on risk generation and selection tasks; (2) how do different team formats affect the ability of MDGHGTs to generate risks; and (3) how do different team formats affect the utilisation of diverse information and perspectives by MDGHGTs.

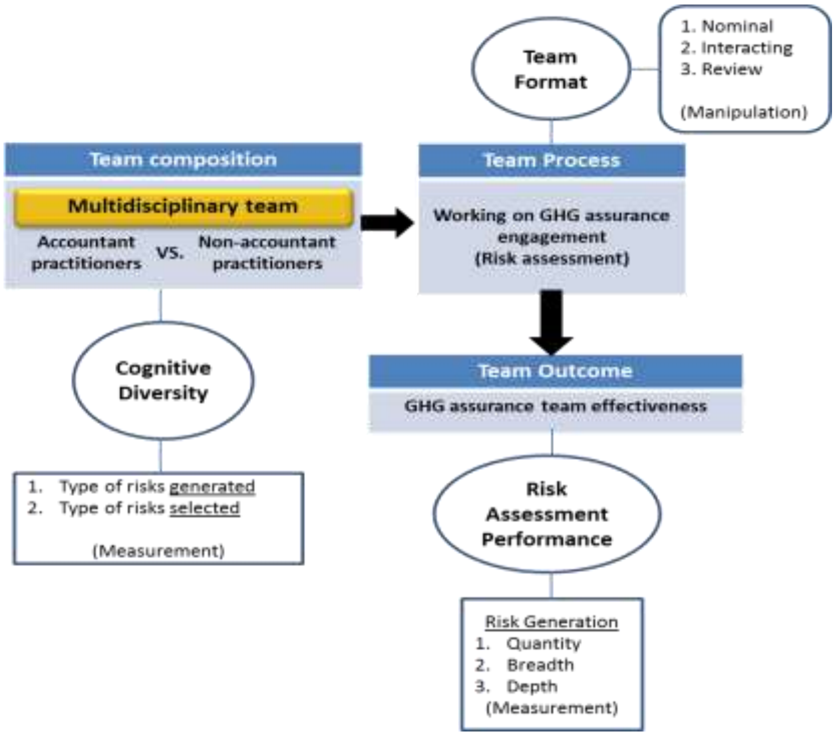


FIGURE 4.1 Research Framework

The participants in this study are 66 GHG assurance practitioners from the Big Four audit firms in Australia. All participants were randomly assigned to 36 two-person MDGHGTs comprising an accountant practitioner (participant with accounting/financial audit background) and a non-accountant practitioner (participant with no accounting/financial audit background, typically with engineering and/or science background).[18] A GHG assurance case scenario developed in conjunction with GHG assurance experts from a Big Four audit firm is used to examine the MDGHGTs' risk assessment performance. The case was developed to include a range of embedded GHG risk elements, thereby providing the opportunity for a broad range of risks to be generated by participants.

---

[18] Output from nominal team members was used as the input for the review teams to reduce the number of participants needed.

To address whether cognitive diversity is present in MDGHGTs working together on risk generation and selection tasks, the types of risks[19] generated and selected are compared between accountant and non-accountant practitioners. Cognitive diversity refers to the differences in the team members' knowledge bases, perspectives, attitudes, values, and beliefs (e.g., Kilduff et al. 2000 and Milliken et al. 2003). Cognitively diverse members are expected to bring a broad set of knowledge and perspective to a given task, which explains why team compositions incorporating differences in educational backgrounds have been increasingly adopted by organisations facing complex tasks (van Knippenberg and Schippers 2007). Therefore, to the extent that cognitive diversity exists between accountant and non-accountant practitioners, the practitioners are likely to complement each other when working together on GHG assurance engagements and thereby support the suggestions to use multidisciplinary teams in ISAE 3410.

To address how different team formats affect the MDGHGTs' ability to generate risks, the performances of three different team formats are compared: nominal, interacting, and review teams. In the nominal team, accountant and non-accountant practitioners complete the tasks independently and then the generated and selected risks are combined by the researcher to form their team performance. This form of team process has normally been treated as a baseline for evaluating team performance (Diehl and Stroebe 1987, 1991). In the interacting team, team members communicate with each other through discussion to generate and select risks, which is in line with the ISAE 3410requirements. In the review team, an accountant practitioner reviews and adds to the risks generated by an individual non-accountant practitioner. Literature on group decision-making (e.g., Osborn 1953) suggests that increasing the quantity of risks generated will also increase the chance that more quality risks are generated, including primary risks. Given the importance of quantity and the difficulty of obtaining an unambiguous measure of the quality of the risks, risk generation performance in this study is measured in terms of the quantity of risks generated.

[19] Two types of risks are compared: (1) risks associated with the measurement of the subject matter, i.e., risks resulting from the measurement of the GHGs emitted into the atmosphere; and (2) risks associated with comparing the subject matter with suitable criteria, i.e., risks resulting from noncompliance with National Greenhouse and Energy Reporting (NGER) reporting criteria and accounting/audit criteria.

To address how different team formats affect MDGHGTs' utilisation of diverse information and perspectives, the breadth and depth of risks generated by nominal, interacting, and review teams are compared. The breadth is the range of issues covered, while the depth reflects the extent to which the issues have been completely examined. Both dimensions are important because insufficient breadth or depth can reduce the quality of ideas generated (Dahlin et al. 2005). For example, teams may generate a broad range of issues but not explore any of those adequately in depth, or they may explore one particular issue in great depth but disregard other important issues. The breadth of risk area coverage is measured by the number of risk categories generated. The depth of risks generated is measured by the number of risk subcategories generated within a risk category.

The remainder of this chapter is organised as follows. Section 4.2 presents the relevant literature and hypotheses development, and Section 4.3 suggests a number of research questions. Section 4.4 describes the research method used in this study; Section 4.5 reports descriptive statistics and tests of the hypotheses; Section 4.6 provides the results of sensitivity analyses; Section 4.7 reports additional analyses; and Section 4.8 summarises and discusses the implications and limitations of this study.

4.2 Relevant Literature and Hypotheses Development

The focus of this study is comparing the performance of different forms of assurance teams comprising practitioners with diverse educational backgrounds (e.g., accounting, engineering, and science). Four hypotheses are developed to address the research questions. Hypothesis 1 examines the cognitive diversity between accountant and non-accountant practitioners by testing the differences in the types of risks they generate and select. Hypothesis 2 examines the effect of the interactions between accountant and non-accountant practitioners on the number of risks generated and the information utilisation, specifically whether the teams employed a breadth or depth approach to generating risks. Hypothesis 3 investigates the effect of the review process on the quantity, breadth, and depth of risks generated. Hypothesis 4 focuses on the difference in the types of risk selected between accountant and non-accountant practitioners.

4.2.1 Risk Assessment Procedures: Risk Generation and Selection

International Standard on Auditing (ISA) 315 defined risk assessment procedures as "the audit procedures performed to obtain an understanding of the entity and its environment, including the entity's internal control, to identify and assess the risks of material misstatement, whether due to fraud or error, at the financial statement and assertion levels" (IFAC 2012b, para. 4). Risk assessment has become increasingly important as recent auditing standards, including ISAE 3410 assurance engagements on GHG statements, have moved towards a risk-based approach requiring auditors to perform such procedures to assess the entity's risk of material misstatement (IFAC 2012a).

Risk assessment procedures involve two important activities: risk generation and risk selection/evaluation. Practitioners are required to generate or identify potential risks of material misstatements (i.e., risk generation) once they gather sufficient knowledge about the clients' business and look for unexpected changes in account balances or ratios (IFAC 2012a, 2012b). However, in order to allocate limited resources to the most important audit areas, not all risks listed by practitioners will be addressed. Thus, practitioners need to exercise judgment and make prioritising decisions to focus attention on risks that are more significant (i.e., risk selection); as ISA 315 states, "the auditor shall determine whether any of the risks identified are, in the auditor's judgment, a significant risk"[20] (IFAC 2012b, para. 27). These diagnostic tasks enable auditors to better understand the nature of their clients' business processes, policies, and control environments, which then forms a basis for the design of substantive tests and resource allocations. Therefore, the failure to identify and select significant risks or to

[20] "Significant risk" has been defined as "an identified and assessed risk of material misstatement that, in the auditor's judgment, requires special audit consideration" (IFAC 2012b, para. 4). To decide which risks are significant risks, auditors should consider whether the risk is related to fraud; recent significant economic, accounting, or other developments; the complexity of transactions; significant transactions with related parties; or the degree of measurement uncertainty. Further, auditors should consider whether the risk involves significant transactions outside the normal course of the entity's business (IFAC 2012b, para. 28). In the GHG assurance context, auditors should also consider the likelihood of non-compliance with the provisions of laws and regulations directly affecting the content of the GHG statement, the omission of a potentially significant emission, the nature of quantification methods, the degree of complexity in determining the organisational boundary, whether Scope 3 emissions are included in the GHG statement, whether the entity makes significant estimates, and the data on which estimates are based (IFAC 2012a, para. 34).

effectively discuss or communicate the information, particularly in the planning stage, can lead to ineffective audit results (Low 2004; Fukukawa and Mock 2011).

4.2.2 Types of Risks Generated by Accountant and Non-Accountant Practitioners: Hypothesis 1

MDTs are often formed because members with diverse educational backgrounds are expected to bring different knowledge and perspectives to problems and decisions encompassing various disciplines (van Asselt 2000). By combining these various cognitive perspectives, MDTs are expected to improve the quality of judgment and decision making (van Knippenberg and Schippers 2007) as well as creativity (Bantel and Jackson 1989). Based on the psychology literature, the differences in team members' knowledge bases and perspectives (Milliken et al. 2003) as well as attitudes, values, and beliefs (Kilduff et al. 2000) are referred to as cognitive diversity. These differences could result from different educational backgrounds[21] because the curriculum of study not only indicates one's personality and cognitive style (Holland 1973) but also shapes the way that person thinks or believes (Dahlin et al. 2005). Therefore, educational diversity is the most salient and important source of creative thinking and reasoning (Nijstad and Paulus 2003).

In the GHG assurance setting, the unique knowledge and skill-sets possessed by accountant and non-accountant practitioners are perceived as necessary to GHG assurance engagements (IFAC 2012a). Accountant practitioners who provide assurance on financial statements are usually well-trained in financial accounting and audit methodologies (Huggins et al. 2011). Because accounting seems to be the only discipline that provides assurance training to members (Gray 2000), accountant practitioners are well recognised for their audit competency. However, non-accountant practitioners claim a competitive advantage in GHG assurance engagements because of

[21] An individual's cognitive style could also be influenced by his or her functional background, such as work experience (Milliken et al. 2003). However, because differences in functional backgrounds could prompt an in-group/out-group identification, which is referred to as social categorisation diversity, the differences in educational background are suggested as a "purer indicator" of cognitive diversity (Dahlin et al. 2005, p. 1108).

their specific expertise and knowledge of the subject matter (Corporate Register 2008; Huggins et al. 2011). The unique and complementary skill-sets that practitioners from accounting and engineering/science backgrounds bring to the GHG assurance engagement and the complexities of the subject matter in this setting demonstrate the benefit of adopting MDTs.

The differences outlined above suggest that when assessing the risk of material misstatement in the GHG setting, accountant and non-accountant practitioners are expected to identify different types of risks. ISAE 3410 requires GHG assurers to evaluate the appropriateness of the entity's quantification methods and the completeness of emissions sources, which requires scientific knowledge, a certain understanding of industrial processes, and the assessment of the consistency of reported emissions with the applicable criteria (IFAC 2012a). Therefore, the risks of material misstatements could be categorised into two types. The first type comprises the risks associated with *the measurement of the subject matter*, such as risks resulting from the measurement of GHGs that are emitted into the atmosphere (e.g., inaccurate, insufficient, and incomplete metering) and the identification of emissions sources. The second type comprises the risks associated with *comparing the subject matter with suitable criteria*, such as the National Greenhouse and Energy Reporting (NGER)[22] scheme for reporting criteria and accounting/audit criteria (e.g., whether the facility boundary and operational control has been determined in accordance with the NGER legislation, whether the methods used to calculate GHG emissions are correctly applied or in line with NGER requirements, and whether activity data is recorded in the correct reporting period).

Given that accountant practitioners are familiar with comparing financial data with audit/accounting criteria, they are more likely to emphasise generating risks associated with *comparing the subject matter with suitable criteria* than non-accountant practitioners. On the other hand, accountant practitioners possess less scientific knowledge and skill-sets and have less of an understanding of quantification methods and the industrial operation/process than non-accountant practitioners. Therefore, they

[22] The Australian Government's NGER system took effect in September 2007. The NGER Act requires entities that meet or exceed one or more of three thresholds (carbon dioxide equivalent, energy production, and energy consumption) to register and report their GHG emissions on a yearly basis (Australian Government ComLaw 2009).

are likely to put less emphasis on generating risks associated with *the measurement of the subject matter*. This leads to the following hypothesis:

Hypothesis 1*: Compared to non-accountant practitioners, accountant practitioners generate a higher proportion of risks associated with **comparing the subject matter with suitable criteria** and a lower proportion of risks associated with **the measurement of the subject matter.***

4.2.3 The Effect of Team Interaction on the Quantity, Breadth, and Depth of Risks Generated by MDGHGTs: Hypothesis 2

Accountant and non-accountant practitioners possess unique knowledge-based and complementary skill-sets that are necessary for undertaking GHG assurance engagements (Huggins et al. 2011). To increase the effectiveness of MDGHGTs, ISAE 3410 requires accountant and non-accountant practitioners to be involved in planning and discussions related to assessing the entity's potential material misstatements (IFAC 2012a, para. 27).

Early studies in psychology (Osborn 1957; Paulus et al. 1993) suggest that exchanging different ideas through group discussions can improve the quantity and quality of ideas generated because it helps stimulate and integrate ideas. Research has focused on how to increase the quantity of ideas generated because it will increase the chance that more quality ideas are generated, or as Osborn (1953) states, "quantity breeds quality". However, empirical research in psychology has consistently found that when various team formats are examined, interacting (or brainstorming) groups generate a lower quantity and quality of ideas than nominal groups (e.g., Hill 1982; Diehl and Stroebe 1987; Mullen et al. 1991; Argote and Kane 2003; Dennis et al. 1999; Rietzschel et al. 2006). Although process gains (i.e., improved performance) would logically be expected when team members work together because additional perspectives and ideas are enabled (van Knippenberg and Schippers 2007), these studies conclude that interacting groups fail to generate more ideas than nominal groups because more process losses (i.e., reduced performance) occur than process gains (e.g., Diehl and Stroebe 1987; Dennis and Valacich 1993). Process losses occur due to production blocking (Diehl and Stroebe 1991), free riding (Strobe and Diehl 1994; Chen et al. 2014), and evaluation apprehension[23] (Camacho and Paulus 1995), which all arise from group interaction. Therefore, group interaction seems to inhibit the idea-generation process.

An extensive amount of audit literature examines the risk generation performance of nominal and/or interacting teams, specifically in a fraud brainstorming setting (Carpenter 2007; Hoffman and Zimbelman 2009; Lynch et al. 2009; Trotman et al. 2009; Carpenter et al. 2011; Chen et al. 2014). While other studies examine alternative forms of interacting brainstorming (Trotman et al. 2009; Hoffman and Zimbelman 2009), Carpenter (2007) is the first to compare face-to-face brainstorming and nominal teams. Carpenter's results show that brainstorming teams generate fewer risks than nominal teams. Lynch et al. (2009) examine electronic brainstorming and find that electronic brainstorming teams generate more relevant fraud risks compared to face-to-face brainstorming teams. However, no difference is found between electronic and nominal brainstorming teams. While Lynch et al. (2009) use undergraduate student subjects, Chen et al. (2014) compare the performance of nominal and interacting electronic brainstorming in hierarchical audit teams. They find that nominal teams generate more fraud risk factors and fraud hypotheses than interacting teams. These results suggest that social loafing by less-experienced auditors explains the differences between nominal and interacting teams in this setting. Moving away from hierarchical audit teams, Carpenter et al. (2011) compare fraud identification performance of interacting brainstorming and nominal teams of internal auditors. Consistent with studies on hierarchical audit teams, this study finds that interacting brainstorming teams identify a lower number of fraud risks than nominal teams. The findings from these studies indicate that nominal teams are superior to interacting teams in risk generation tasks.

[23] Production blocking occurs because group members have to take turns verbalising their ideas and therefore have to listen to others' ideas while thinking, which could interfere with their own thoughts. Free riding occurs because group members rely on others to complete the task for various reasons, such as the perception that their inputs are unidentifiable or dispensable. Evaluation apprehension occurs because group members are afraid of being evaluated by other group members and therefore withhold their ideas.

There are several plausible explanations as to when and why interaction does not work for idea generation groups. Trotman (1985) and Libby et al. (1987) suggest that interacting groups will outperform nominal groups if enough variation is present in group members' performance and if they are able to recognise the differences in their expertise. Further, Trotman (1985) and Libby et al. (1987) suggest that the group task has to be complex enough to detect differences in expertise. Another plausible explanation is the low cognitive diversity among participants in previous research (Nijstad and De Dreu 2002). Since the majority of brainstorming studies in psychology (e.g., Hill 1982; Diehl and Stroebe 1987; Dennis and Valacich 1993) use students with similar educational backgrounds, the students are unlikely to bring sufficiently different knowledge and perspectives to the task and are therefore less likely to generate different ideas. Similarly, brainstorming studies in auditing comprise auditors with similar backgrounds (e.g., Carpenter 2007; Chen et al. 2014)

Stroebe and Diehl (1994) present one of the very few studies to test the effect of cognitive diversity on group idea generation. They find that interacting groups with heterogeneous members (in terms of dominant associations regarding environmental concerns) generate almost the same number of ideas as nominal groups. They suggest that a broad range of ideas and perspectives shared within cognitively diverse groups stimulates the creativity of members in the team, which in turn outweighs the productivity losses usually observed in interacting groups, such as production blocking.

However, the literature on work team diversity suggests that the relationship between cognitive diversity (i.e., diverse educational backgrounds) and team performance is more complicated than expected. On one hand, cognitive diversity in teams is predicted to bring different opinions and perspectives to the decision-making task along with a broader range of task-relevant knowledge, skills, and abilities (e.g., van Knippenberg and Schippers 2007; Jackson and Joshi 2011). Previous studies in psychology (Jehn 1995; Jehn et al. 1999) find that the cognitive diversity among MDT members leads to disagreements on task-related issues, which forces teams to engage in the thorough exchange, clarification, and reconciliation of different knowledge and perspectives, thus increasing team effectiveness (Jehn 1995; Jehn et al. 1999; van Knippenberg et al. 2004). On the other hand, if too much difference is present in the group members' educational backgrounds, they will have difficulty following each other's reasoning because of the different frames of reference, interpretation, and professional language used (e.g., van Someren et al. 1998). These difficulties have been evident across MDT studies that find that team members with different professional backgrounds often use different language or terminology only understood by people in the same profession or that have different meanings in other fields (Carlile 2004; Sheehan et al. 2007). Therefore, while some degree of diversity in the knowledge or expertise within a group is likely to result in process gains, too much diversity can have negative effects (Nijstad and Paulus 2003; Paulus 2008).

Support for tensions between cognitively diverse accountant and non-accountant practitioners has been observed when they work together on sustainability assurance engagements (O'Dwyer 2011). O'Dwyer (2011) suggests that these tensions emerge because assurers from different disciplines have different ways of interpreting, judging, and approaching data, specifically qualitative data. These tensions may also arise in the GHG assurance setting, in which practitioners from two distinctive disciplines (i.e., accounting and engineering/science) work together on a risk assessment task. Therefore, although their complementary knowledge and skill-sets are potential sources of process gains, process losses may arise from the distinctly different mindsets (e.g., different interpretation, concepts, and professional languages) of accountant and non-accountant practitioners. These differences can cause a lack of understanding or misunderstandings among members (van Someren et al. 1998; van Asselt 2000), which in turn affects the idea stimulation process or the ability to build on others' ideas. Further, interacting MDTs need to communicate and reconcile different perspectives, which requires time. On balance, the difficulties encountered due to differing mindsets means that these potential process losses are expected to outweigh the process gains. Thus, when all teams consist of cognitively diverse accountant and non-accountant practitioners, the following hypothesis can be formulated:

Hypothesis 2a: *Interacting teams will generate a lower* ***quantity*** *of risks of potential material misstatements than nominal teams.*

Although it could be difficult for multidisciplinary interacting teams to outperform nominal teams in terms of the quantity of ideas generated, the benefits of interaction between cognitively diverse members are likely to be captured in the quality measures of performance. Accordingly, previous research suggests a range of different dimensions of quality that could be used to further examine the performance of idea-generating groups. These include uniqueness (Parnes and Meadow 1959); diversity (Paulus and Yang 2000); originality and feasibility (Rietzschel et al. 2006); and breadth and depth of generated ideas (Nijstad et al. 2002; Dahlin et al. 2005; Smith 2008; Kohn and Smith 2011). Among these quality measures, the breadth and depth of ideas have been widely used in the social psychology literature to examine the idea-generation performance of cognitively or educationally diverse teams (e.g., Stroebe and Diehl 1994; Nijstad et al. 2002; Dahlin et al. 2005, Kohn and Smith 2011). The breadth and depth of ideas are also used to measure the quality of the audit procedures identified in the audit judgment and decision-making literature (Asare et al. 2000; Green and Trotman 2003). Increasing the breadth of ideas allows many possible alternatives to be analysed, while increasing the depth of ideas allows important issues to be focused upon and thereby explored more completely (Dahlin et al. 2005).

MDTs have advantages over teams composed of members with similar educational backgrounds in terms of the breadth of ideas generated because of their chance of being stimulated by a broader range of knowledge, skills, and perspectives (Stroebe and Diehl 1994; Nijstad 2002; Dahlin et al. 2005). Stroebe and Diehl (1994) find that interacting groups with heterogeneous members (in terms of dominant associations regarding environmental concerns) generate almost the same number of categories (breadth) of ideas as nominal groups. They suggest that the broad range of ideas and perspectives shared within the cognitively diverse groups triggers group members to explore different categories of ideas, which thus outweighs the productivity losses usually observed with interacting groups.

While groups exposed to semantically heterogeneous ideas generated more categories of ideas (more breadth), Nijstad et al. (2002) find that groups exposed to semantically homogeneous ideas generated more new ideas within the same category (more depth).They conclude that diverse cognitive stimulation increases the breadth of ideas generated, while homogeneous cognitive stimulation increases the depth of ideas generated. The social psychology literature on information use in MDTs shows that teams that are more educationally diverse use a wider range of information and analyse ideas in greater depth compared to teams that are less diverse (Dahlin et al. 2005). This literature suggests that the broader set of knowledge and frameworks possessed by MDTs allows them to analyse familiar information in terms of both breadth and depth while leaving more time for them to deeply process unfamiliar information. However, these benefits of diversity only hold up to a certain point. The literature shows that once the level of educational diversity in the team becomes too high, the breadth and depth of information use decreases. Therefore, too much diversity in education makes it difficult for team members to understand each other and thus inhibits information sharing, exploring, and integrating (West 2002; Dahlin et al. 2005).

Unlike the cognitively diverse teams in previous studies (i.e., Strobe and Diehl 1994; Nijstad et al. 2002), MDGHGTs are much more diverse in their educational backgrounds. Given the relatively high level of diversity, accountant and non-accountant practitioners in MDGHGTs are therefore likely to have difficulty understanding each other's reasoning (as evidenced by O'Dwyer 2011) and thus may not be able to fully build on shared ideas or deeply explore unfamiliar issues. Therefore, the process gains from interactions between MDGHGT members will not necessarily cancel out process losses and take the performance of the interacting teams to the level of nominal teams both in terms of the number of risk categories generated (breadth) and the number of risks generated within a category (depth). Moreover, it could be difficult for interacting teams to explore issues in depth because various issues compete for attention when members are diverse in their educational backgrounds (Dahlin et al. 2005). However, when accountant and non-accountant practitioners generate risks alone, they are likely to spend time focusing on categories of risks that they are familiar with or are knowledgeable about because they do not have a chance to be stimulated by different perspectives. This is suggested to result in nominal teams exploring particular issues in more depth than interacting teams. Accordingly, when all teams consist of accountant and non-accountant practitioners, the following hypotheses are tested:

Hypothesis 2b: *Interacting teams will generate fewer categories of risks (**breadth**) than nominal teams.*

Hypothesis 2c: *Interacting teams will generate fewer risks within categories (**depth**) than nominal teams.*

4.2.4 The Effect of the Review Process on the Quantity, Breadth, and Depth of Risks Generated by MDGHGTs: Hypothesis 3

The review process is another team format that has been widely studied in the audit judgment and decision-making literature (see Rich et al. 1997a, 1997b for reviews). By having a more senior auditor evaluating the work of a more junior auditor, the review process has consistently been found to improve audit judgments (Trotman and Yetton 1985; Trotman 1985). Trotman and Yetton (1985) examine the effect of the review process on the consensus of internal control evaluations and find that the review process significantly improves the level of consensus. However, they point out that similar improvements could also be obtained using nominal or interacting groups. They suggest that the addition of a second opinion, regardless of its form, seems to improve audit effectiveness. However, in a more complex task that allows an easier differentiation of expertise, Trotman (1985) finds that review teams significantly outperform nominal teams because the review process reduces the systematic bias and variance in individual judgments. Consistent with this view, Ismail and Trotman (1995) examine the effectiveness of the review process on a hypotheses-generation task and find that the review process increases the number of plausible hypotheses generated regardless of the group members' experience (i.e., senior or manager) or the group interaction (i.e., with or without discussion). They suggest that auditors may benefit from a larger pool of information or may be stimulated by other auditors' ideas during the review process.

Recent studies on the audit review process examine the effect of alternative review formats on *reviewers* and reviewees' performances (Brazel et al. 2004; Agoglia et al. 2009; Payne et al. 2010). Focusing on the performance of audit workpaper preparers (reviewees), Brazel et al. (2004) examine the effect of face-to-face and electronic reviews on preparers' effectiveness and efficiency. This study finds that preparers anticipating a face-to-face review prepare more effective workpapers, make higher quality judgments, feel more accountable, and are less efficient compared to preparers anticipating an electronic review. This study also compares the two review formats with a control group (no review). While a significant difference is found between the face-to-face and control groups, no difference is found between the electronic review and control groups. This study suggests that the nature of face-to-face review provides advantages over electronic review because it allows for a real-time response, the presence of a reviewer, and more effective communication. Agoglia et al. (2009) extend Brazel et al.'s (2004) work by investigating the effects of face-to-face and electronic reviews on the quality of reviewers' judgments. This study finds that reviewers in the electronic review condition make lower quality judgments than reviewers in the face-to-face condition. The quality of the preparers' workpapers is also found to be lower when preparers anticipate an electronic review as opposed to a face-to-face review. The mediation analysis reveals that reviewers in the electronic condition have difficulties recognising and mitigating lower-quality workpapers, which then leads to the lower quality of their going concern judgments. Payne et al. (2010) compare the effects of adding a discussion after preparing written review notes (as opposed to adding no discussion). This study finds that a face-to-face discussion of written review notes increases audit effectiveness because preparers examine the audit evidence more thoroughly compared to when they anticipate written review comments. The findings from these studies suggest that the review format affects the audit teams' effectiveness.

Although the review process seems promising as a source of process gains, it is typically used in hierarchical teams comprised of practitioners with similar educational backgrounds. It is unknown whether the review process will also provide the same advantages when it is applied to educationally diverse teams. While accountant and non-accountant practitioners are required to work together as a team in the GHG assurance setting, it is common practice for a partner in the financial audit practice to act as the signing partner for the firm. Therefore, in practice, accountant practitioners review the work of non-accountant practitioners. Given the differences in their educational backgrounds, accountant practitioners are likely to be stimulated by the different knowledge and perspectives non-accountant practitioners had on the task while reviewing the ideas of their non-accountant team members (e.g., Ismail and Trotman

1995; Paulus and Yang 2000). Unlike review teams, individual accountant and non-accountant practitioners in nominal teams do not have the opportunity to see each other's ideas. While cognitive stimulation does occur for the review teams, the review teams are more likely to generate a larger number of risks than nominal teams because no cognitive stimulation is expected for members in nominal teams. Interacting GHG assurance teams are also less likely to outperform review teams in terms of the quantity of risks generated. This is because although cognitive stimulation is expected for interacting teams, there are process losses due to interaction as discussed earlier. On the other hand, since no discussion is allowed for review teams, they have an advantage in that they have an opportunity as reviewers to build on the reviewees' ideas while not having to spend time discussing the ideas. Thus, when all teams consist of accountant and non-accountant practitioners, the following hypotheses can be proposed:

Hypothesis 3a: *Review teams will generate a higher* ***quantity*** *of risks of potential material misstatements than interacting teams.*

Hypothesis 3b: *Review teams will generate a higher* ***quantity*** *of risks of potential material misstatements than nominal teams*

Diverse ideas are found to increase categories of ideas generated by stimulating individuals to think differently and come up with ideas that they may not think of before interacting (Nijstad et al. 2002), particularly when they see the ideas of others for the first time (Kohn et al. 2011). Based on these findings, when accountant practitioners review the risks generated by non-accountant practitioners, they are likely to be triggered to explore new categories or form new combinations of risks (i.e., generate risks with more coverage or breadth), which then gives review teams an advantage over nominal teams. As discussed earlier, nominal teams do not have a chance to be exposed to diverse cognitive stimulation, therefore, they are less likely to generate risks with more breadth than review teams. While interacting teams are likely to be exposed to such stimulation, they then need more time to exchange, discuss, and integrate those different perspectives. As such, when all teams consist of accountant and non-accountant practitioners, review teams are likely to generate more categories of risks (breadth) than interacting teams and nominal teams. This leads to the following hypotheses:

Hypothesis 3c: *Review teams will generate more categories of risks (**breadth**) than interacting teams.*

Hypothesis 3d: *Review teams will generate more categories of risks (**breadth**) than nominal teams.*

Dahlin et al. (2005) show that MDTs tend to drill deep down into a category when they have enough time to deeply process unfamiliar information after they process familiar information. In comparison with interacting teams, review teams are more likely to generate risks with a greater depth because the reviewers in review teams do not need to spend time exchanging, discussing, and integrating different information and perspectives with the reviewees. Thus, they are more likely to have time to explore unfamiliar issues in depth than interacting teams. However, in comparison with nominal teams, review teams are more likely to generate risks with less depth because being exposed to diverse stimulation makes it difficult for them to focus on specific issues. Unlike review teams, nominal teams are not stimulated by diverse perspectives. As such, they are likely to go deeper into a few risk categories they are familiar with and thus generate more risks within their domain of knowledge. Therefore, when all teams consist of accountant and non-accountant practitioners, the above expectations can be tested using the following hypotheses:

Hypothesis 3e: *Review teams will generate more risks within categories (**depth**) than interacting teams.*

Hypothesis 3f: *Review teams will generate fewer risks within categories (**depth**) than nominal teams.*

4.2.5 Types of Risks Selected by Accountant and Non-Accountant Practitioners: Hypothesis 4

The majority of brainstorming studies focus heavily on idea generation rather than idea selection (West 2002; Paulus 2008). From a practical point of view, not all ideas can be implemented; therefore, these ideas should be evaluated and only some should be selected for further implementation (Reiter-Palmon et al. 2012). Idea selection may be completed by groups or individuals that generate the ideas or by other groups or individuals not involved in the idea-generation process (Paulus 2008). Groups are expected to be better at selecting ideas than individuals because they have a larger pool of knowledge and perspectives, which in turn helps screen out inappropriate alternatives (Laughlin and Hollingshead 1995; Paulus 2008). However, research in psychology (e.g., Faure 2004; Rietzschel et al. 2006; Rietzschel et al. 2010) and auditing (e.g., Bedard and Biggs 1991; Hirst and Koonce 1996; Asare and Wright 1997; Green and Trotman 2003; Moreno et al. 2007; Luippold and Kida 2012; Pike et al. 2013) has shown that neither individuals nor groups perform well at idea selection and evaluation. Previous psychology literature suggests that people select ideas based on what they think is more important to them rather than what is actually important, which explains why people usually fail to select the best ideas (Reitzschel et al. 2010).

Cognitive diversity among MDT members is evident in sustainability assurance, in which accountant and non-accountant practitioners have different concepts of materiality (O'Dwyer 2011). Therefore, what accountant practitioners think is significant may not be significant for non-accountant practitioners and vice versa. This cognitive diversity is also likely to be found in the GHG assurance context because accountant and non-accountant practitioners on GHG assurance teams are highly diverse in their educational backgrounds (i.e., accounting vs. environmental science). Based on the findings discussed above, accountant and non-accountant practitioners in the GHG assurance setting are likely to select different types of risks. Given their educational backgrounds, an accountant practitioner may perceive risks associated with *comparing the subject matter with suitable criteria* to be more important than a non-accountant practitioner would. On the other hand, accountant practitioners may perceive risks associated with *the measurement of the subject matter* to be less important than non-accountant practitioners would. Thus:

Hypothesis 4: *Compared to non-accountant practitioners, accountant practitioners select a higher proportion of risks associated with* ***comparing the subject matter with suitable criteria*** *and a lower proportion of risks associated with* ***the measurement of the subject matter.***

A summary of all the proposed hypotheses is provided in Table 4.1.

TABLE 4.1 Summary of Hypotheses

Stage	Hypothesis	Expectation
Risk Generation: Types of Risks	1	Compared to non-accountant practitioners, accountant practitioners **generate** a higher proportion of risks associated with **comparing the subject matter with suitable criteria** and a lower proportion of risks associated with **the measurement of the subject matter.**
Risk Generation: Interaction	2a	Interacting teams will **generate** a lower **quantity** of risks of potential material misstatements than nominal teams.
	2b	Interacting teams will **generate** fewer categories of risks (**breadth**) than nominal teams.
	2c	Interacting teams will **generate** fewer risks within categories (**depth**) than nominal teams.
Risk Generation: Review process	3a	Review teams will **generate** a higher **quantity** of risks of potential material misstatements than interacting teams.
	3b	Review teams will **generate** a higher **quantity** of risks of potential material misstatements than nominal teams.
	3c	Review teams will **generate** more categories of risks (**breadth**) than interacting teams.
	3d	Review teams will **generate** more categories of risks (**breadth**) than nominal teams.
	3e	Review teams will **generate** more risks within categories (**depth**) than interacting teams.
	3f	Review teams will **generate** fewer risks within categories (**depth**) than nominal teams.
Risk Selection: Types of risks	4	Compared to non-accountant practitioners, accountant practitioners **select** a higher proportion of risks associated with **comparing the subject matter with suitable criteria** and a lower proportion of risks associated with **the measurement of the subject matter.**

4.3 Research Questions

4.3.1 The Effect of Team Format on the Breadth and Depth of Risks Selected: Research Question 1

Social psychology literature suggests that team formats not only affect the quantity of ideas but also the breadth and depth of ideas generated (Nijstad et al. 2002; Dahlin et al. 2005; Smith 2008; Kohn and Smith 2011). In particular, the greater breadth and depth of information used by MDTs represent the main advantages MDTs have over teams comprising members with similar educational backgrounds (Dahlin et al. 2005). While the effects of different team formats on the breadth and depth of risks has been extensively studied in idea generation literature, it is not well established in idea selection literature. As MDGHGTs do not engage only in idea generation but also in idea selection (Paulus 2008), how the breadth and depth of risks selected by nominal, interacting, and review teams vary is an empirical question. Understanding these effects will give more insight into the quality aspect of the risks selected by educationally diverse teams. Consequently, the following research questions are considered:

Research Question 1a: Does the type of team format affect the ***breadth of risks selected*** *by teams comprising accountant and non-accountant practitioners?*

Research Question 1b: Does the type of team format affect the ***depth of risks selected*** *by teams comprising accountant and non-accountant practitioners?*

4.3.2 Relationship between Elaboration of Task-Relevant Information and the Performance of Interacting MDGHGTs: Research Question 2

Van Knippenberg et al. (2004) propose a “categorisation-elaboration model (CEM)” positing that the “elaboration of task-relevant information is the primary process underlying the positive effects of diversity on performance” (p. 1012). Elaboration is defined as “the exchange of information and perspectives, individual-level processing of information and perspectives, the process of feeding back the results of this individual-level processing into the group, and discussion and integration of its implications" (van Knippenberg et al. 2004, p. 1011). Thus, the elaboration of task-relevant information has been suggested as a moderator of the effects of diversity on MDT performance. A few studies test the role of elaboration on decision-making quality, including van Ginkel and van Knippenberg (2008). They find that when group members have the same representation emphasising information elaboration and realise this similarity, they elaborate more and make higher quality decisions. A study by van Ginkel et al. (2009) also finds that information elaboration positively affects group decision making when groups have the opportunity to discuss their task, goals, and how to reach the goals beforehand. Although previous studies show the benefit of information elaboration on decision-making tasks, the research has yet to establish a link between elaboration and idea generation or idea selection in MDTs. Consequently, the following research questions are considered to explore these relationships:

Research Question 2a: *Is the level of elaboration of task-relevant information correlated with the quantity, breadth, and depth of risks* ***generated*** *by interacting teams of accountant and non-accountant practitioners?*

Research Question 2b*: Is the level of elaboration of task-relevant information correlated with the breadth and depth of risks* ***selected*** *by interacting teams of accountant and non-accountant practitioners?*

4.3.3 Relationship between Cross-Understanding and the Performance of Interacting MDGHGTs: Research Question 3

Huber and Lewis (2010) propose another potential factor, "cross-understanding", as a means of explaining inconsistent findings relating to performance in the group literature. Cross-understanding refers to "the extent to which group members have an accurate understanding of one another's mental models" (Huber and Lewis 2010, p. 7). This study suggests that group members will be more likely to predict other members' behaviours by understanding "what others know, believe, are sensitive to, and prefer"

(Huber and Lewis 2010, p. 9). Consequently, the members will be able to select their responses more effectively. The high level of cross-understanding is posited to enhance communication, elaboration effectiveness, and collaborative behaviours by encouraging members to share, discuss, and integrate their diverse knowledge and perspectives while increasing the quantity and quality of task-relevant information discussed in the group. Therefore, a high level of cross-understanding between members can potentially cancel out the negative effects of cognitive diversity found in MDTs, which could improve the performance of MDGHGTs. These relationships are explored in this study by considering the following research questions:

Research Question 3a: Is the level of cross-understanding correlated with the quantity, breadth, and depth of risks ***generated*** *by interacting teams of accountant and non-accountant practitioners?*

Research Question 3b: Is the level of cross-understanding correlated with the breadth and depth of risks ***selected*** *by interacting teams of accountant and non-accountant practitioners?*

4.3.4 Perceived Ability and Knowledge of Accountant and Non-Accountant Practitioners in MDGHGTs: Research Question 4

O'Dwyer (2011) conducts a case study on sustainability assurance practice by interviewing accountant and non-accountant (i.e., no financial audit background) assurers involved in sustainability assurance engagements. The interview results show that accountant practitioners think non-accountant practitioners have insufficient knowledge of audit criteria and procedures, while non-accountant assurers think accountant assurers do not have enough knowledge of the subject matter to work together on a task.

Assurance and subject matter experts' perception of their team members' ability and knowledge to perform the risk assessment task in the GHG assurance setting is explored in this study through the following research questions:

Research Question 4a: Do accountant and non-accountant practitioners perceive themselves to be different from the other team member in terms of their ***ability to identify risks****?*

Research Question 4b: Do accountant and non-accountant practitioners perceive themselves to be different from the other team member in terms of their ***knowledge of the audit criteria and process****?*

Research Question 4c: Do accountant and non-accountant practitioners perceive themselves to be different from the other team member in terms of their ***knowledge of the subject matter****?*

4.4 Research Methods

4.4.1 Participants

Sixty-six participants from the Big Four audit firms in Australia with GHG assurance engagement experience participated in this experiment. Contact partners and managers from dedicated assurance groups performing GHG assurance at each firm were asked to nominate potential GHG assurance group members to participate in this study. All the participants voluntarily participated, and the confidentiality of their information was ensured.[24] All participants received a $50 gift voucher for participating in the experiment.

Nine experimental sessions involving 60 participants were conducted at the Big Four audit firms' offices in Sydney and Melbourne. In addition, experimental sessions were conducted remotely via teleconferencing or videoconferencing for six participants located in Melbourne, Brisbane, and Perth.[25]

Prior to each experimental session, a list of participants was provided by the firm, including details of their educational specialisations, professional experience, and

[24] The study was approved and strictly followed the ethical requirements of the Human Research Ethics Advisory Panel at the University of New South Wales.

[25] These participants were assigned to work individually in either the nominal or the review treatment because it was not possible for them to be on an interacting team.

expertise categorisation. Participants were then classified as either an “accountant” or a “non-accountant” practitioner. If they had an accounting degree and/or audit experience, they were classified as an “accountant”.[26] Six participants were dropped from the analysis,[27] resulting in 60 participants for the experiment: 28 from Firm A, 10 from Firm B, 12 from Firm C, and 10 from Firm D. To test the effect of different team formats on the performance of MDGHGTs, an accountant practitioner and a non-accountant practitioner were grouped together in a team of two (a dyad).[28] In all, the study included 12 dyads for nominal teams, 12 dyads for interacting teams, and 12 dyads for review teams (Table 4.2).

Accountant practitioners were randomly allocated to either do the task individually (nominal) or collaboratively with a non-accountant practitioner (interacting) or to review the work of an individual non-accountant practitioner (review). Non-accountant practitioners were randomly allocated to either do the task individually (nominal) or collaboratively with a non-accountant practitioner (interacting). The work of each individual non-accountant practitioner in the nominal teams was used as the input for the review team. In other words, 12 non-accountant practitioners in the nominal teams were used to form the review teams. In this way, 60 participants effectively became 72 team members.

---

[26] Five participants had both accounting and engineering/environmental science backgrounds (i.e., they had double degrees or a master’s degree in another area). These participants were categorised according to their professional experience. Three participants with financial audit experience were classified as “accountants”, while two participants without any financial audit experience were classified as “non-accountants”.

[27] Three non-accountant practitioners who could not be matched with any accountant practitioners (i.e., leftover participants) and three participants who were misclassified as an accountant or a non-accountant practitioner were dropped from the analysis. For the three participants who were misclassified, one reviewer was excluded because of the misclassification of the reviewee’s educational background (which resulted in an accountant reviewing another accountant’s work); and two participants from the same interacting team were excluded because both were non-accountant practitioners. Misclassification occurred because the classifications of the participants provided by the firms differed across the firms. To ensure a consistent background for the teams created for the experiment, misclassifications were determined by rechecking the participants’ backgrounds (as provided by the firm) against each participants’ answers to demographic questions in the post-experimental questionnaire.

[28] One interacting team comprised two members with accounting backgrounds and audit experience. However, this team was not dropped from the analysis because one member had one year of experience in GHG assurance, while the other member had no experience in GHG assurance. As such, the team member with GHG experience was classified as a non-accountant practitioner, while the one without GHG experience was classified as an accountant practitioner. Further, three participants in three review teams had economics or commerce backgrounds with no financial audit experience. Given that they had completed accounting courses, they were classified as accountant practitioners. Sensitivity analysis was conducted by excluding these participants and yielded similar results.

TABLE 4.2 Participant Distribution

	Treatment			
	Nominal Team	**Interacting Team**	**Review Team**	**Totals**
Teams	**12**	**12**	**12**	**36**
Participants	**24**	**24**	**24**	**72***
Accountant practitioners	12	12	12	36
Non-accountant practitioners	12*	12	12*	36

*The nominal and review teams shared the same non-accountant practitioners. In all, 60 participants took part in the experiment.

The participants' demographic data, based on 72 team members, is reported in Table 4.3. In terms of the audit firm membership (Panel A), 31 team members were from Firm A, 12 were from Firm B, 17 were from Firm C, and 12 were from Firm D. The chi-square tests of independence show no significant differences in the audit firm membership across the three experimental conditions ($\chi^2 = 1.775$, $p > 0.99$).

Panel B of Table 4.3 shows that the team members consisted of auditors/consultants (16); senior auditors/consultants (21); audit/sustainability managers and directors (31); and audit/sustainability partners (4). The team members' current positions did not differ significantly ($p > 0.10$) across the three conditions. Panel C of Table 4.3 shows that, on average, team members had 2.69 years[29] of financial audit experience and 2.42 years of GHG assurance experience.[30] They had 4.75[31] days training on GHG assurance on

[29] No significant difference is found in the average financial audit experience across all teams (all $p > 0.10$).

[30] No significant difference is found in the average GHG assurance experience between the practitioners in the nominal versus interacting treatments ($z = -0.151$, $p = 0.908$) and practitioners in the interacting versus review treatments ($z = -1.433$, $p = 0.152$). However, practitioners in the review treatment had more GHG assurance experience than practitioners in the nominal treatment (2.94 vs. 2.00, $z = -2.708$, $p = 0.009$). Additional analyses showed no correlation between GHG assurance experience and all dependent variables (all $p > 0.10$). After controlling for GHG assurance experience, the statistical significance levels for all hypotheses between the review and nominal teams remained the same.

[31] One participant in the review treatment had significantly more training in GHG assurance than the rest of the participants (50 days of training). After excluding this participant, the average number of days of GHG assurance training dropped to 4.18 days. No significant difference is found in the average training days between practitioners in the nominal versus interacting treatments ($z = -0.854$, $p = 0.393$) and practitioners in the nominal versus review treatments ($z = -1.304$, $p = 0.192$). However, practitioners in the review treatment had more training than practitioners in the interacting treatment (7.10 vs. 2.75, $z = 1.967$, $p = 0.049$). After excluding the participant in the review team who had extensive training, a marginally significant difference was found between the review and interacting treatments in terms of

average. For the interacting teams, the average familiarity between members is 5.00 (un-tabulated) as measured on a seven-point scale (1 = not familiar at all; 7 = very familiar).

TABLE 4.3 Participants' Demographics by Experimental Conditions

Panel A. Firm

Firm	Number of Team Members Nominal Team (*n* = 24)	Interacting Team (*n* = 24)	Review Team (*n* = 24)	Total (*N* = 72)*
Firm A	10	12	9	31
Firm B	3	4	5	12
Firm C	7	4	6	17
Firm D	4	4	4	12

Panel B. Current Position

Current Position	Number of Team Members Nominal Team (*n* = 24)	Interacting Team (*n* = 24)	Review Team (*n* = 24)	Total (*N* = 72)*
Auditor/Consultant	5	6	5	16
Senior Auditor/Consultant	8	7	6	21
Audit/Sustainability Manager/Director	10	10	11	31
Audit/Sustainability Partner	1	1	2	4

Panel C. Experience

Experience	Team Members' Experience Nominal Team (*n* = 24)	Interacting Team (*n* = 24)	Review Team (*n* = 24)	Totals (*N* = 72)*
Average (range) years of financial audit experience	3.27 (25–0)	2.54 (17–0)	2.25 (13–0)	2.69 (25–0)
Average (range) years of GHG assurance experience	2.00 (8–0)	2.33 (12–0)	2.94 (10–0)	2.42 (12–0)
Average (range) days of training on GHG assurance	4.40 (30–0)	2.75 (10–0)	7.10 (50–0)	4.75 (50–0)

* Although only 60 participants took part in the experiment, the nominal and review teams shared the same 12 non-accountant practitioners, bringing the notional total to 72.

GHG assurance training days (z = -1.689, p = 0.091). Additional analyses showed no correlation between training days and all dependent variables (all p > 0.10) and the statistical significance levels for all hypotheses remain the same after excluding this participant.

4.4.2 Task Development

4.4.2.1 Task Context and Elements

The case was co-developed with the assistance of an expert from the Climate Change and Sustainability group from one of the Big Four firms. This expert, who is a Chartered Accountant and holds a Masters of Environmental Science, worked in the audit division and the global services group for the firm on developing and rolling out sustainability and GHG assurance training. The case was also reviewed by a non-accountant expert from the same Big Four firm with an engineering/science background to ensure that the content and numbers in the case were accurate from a scientific point of view.

The case materials contain planning documents to help participants understand the entity, including the entity's industry, operations, and sources of emissions. The entity's emissions data for the current and prior years as well as the assumptions and ratio analysis, which were prepared by the engagement team, are also provided. The case materials are provided in Appendix 3.

The task used in this study involves the completion of three stages of risk identification and the related planning process for an unexpected variation in reported GHG emissions. First, participants identified the potential risks of material misstatements that could occur for the entity. Second, they were asked to focus on the risks they considered most important by selecting the four most significant risks. Third, they were asked to identify appropriate procedures to address the selected risks.

In developing the case, several criteria were applied to ensure that the case was sufficiently nuanced to allow the detection of differences for team members' background expertise (i.e., accountant or non-accountant) and to test the effect of the team formats (i.e., nominal, interacting, and review) on the MDGHGTs' performance. The first criterion is that the case had to be complex enough to detect differences in expertise (Trotman 1985). The complex nature of the steelmaking processes chosen for the case allows the use of different types of raw material, energy, and quantification methodologies as well as the inclusion of multiple facility locations. In addition, the case materials designate that the process emissions from steel production are calculated using the “Carbon Mass Balance” method, which requires the measurement and analysis of the carbon content of all inputs and outputs. Therefore, the high level of complexity embedded in this case is expected to allow an examination of the different types of risks generated and selected by participants from different disciplines.

The second criterion is that the case had to involve a reasonable level assurance. According to ISAE 3410, a reasonable level assurance requires practitioners to identify and assess the risks of material misstatements at both the GHG statement level and the assertion level, while a limited level assurance requires them to identify and assess the risks of material misstatements only at the GHG statement level (IFAC 2012a). Therefore, identifying risks at the assertion level allows more specific risks to be identified and increases the need for different expertise.

The third criterion is that the case had to cover various risk categories from a range of issues mentioned in ISAE 3410. These include risks at the GHG statement level (e.g., risks of fraud, inconsistent quantification methods and reporting policies, errors in unit conversion when consolidating information from facilities) and at the assertion level (e.g., inaccurate quantification of emissions, incomplete emissions recorded, emissions recorded in the incorrect reporting period). Because one of the task requirements is for participants to select the top four risks from their list of risks (refer to section 4.4.4.2 for specific task requirements), it was necessary to ensure that enough potential risk issues were present in the case to be identified and selected. The nature of the entity’s business and the steelmaking processes designed in this case allows at least 11 risk categories to be identified. The details of the potential risk categories will be discussed in the next section.

The last criterion is that the case had to allow participants to observe their team members’ working behaviour. The spreadsheets included in the case contain audit working papers regarding the entity’s emissions data. These working papers provide details of the calculations of total emissions for the current and prior years, the assumptions used in the emissions calculations, and the ratio analyses for the current and prior years. Including these spreadsheets in the case not only provides an understanding of the entity's emissions calculation processes but also allows tests to be undertaken to determine whether accountant practitioners focus more on assessing key numerical indicators and quantitative data or emphasise data accuracy, as found in the sustainability assurance setting (e.g., O'Dwyer 2011).

4.4.2.2 Risk Issues

Various risk issues were embedded in the case materials. These risk issues were derived from the current GHG reporting and assurance standards, including ISAE 3410 (IFAC 2012a), the Greenhouse Gas Protocol (WBCSD and WRI 2004, 2005), and other government publications (e.g., Defra 2009). These publications suggest five steps that are required when assessing reported GHG emissions: determining the entity's organisational boundary, identifying emissions sources, collecting data, calculating emissions, and reporting emissions data. Figure 4.2 illustrates the five steps and the related risks embedded in the case.

Step 1: Determining the Entity's Organisational Boundary

This step involves identifying the parts of the business/operations that are owned or controlled by the entity and thereby should be included in the entity's GHG statements. ISAE 3410 requires GHG assurance practitioners to evaluate the appropriateness of the entity's organisational boundary determination and to ensure compliance with applicable criteria (IFAC 2012a). Therefore, a risk category related to *inaccurate and incomplete boundary determination* could be identified. This category of risks is embedded in the case by adding the following information: a recycle company (RecycleCo.) is located on the land owned by the entity and has the authority to manage the site; on-site contractors use the entity's natural gas for welding and forklifting; and an off-site rolling mill.

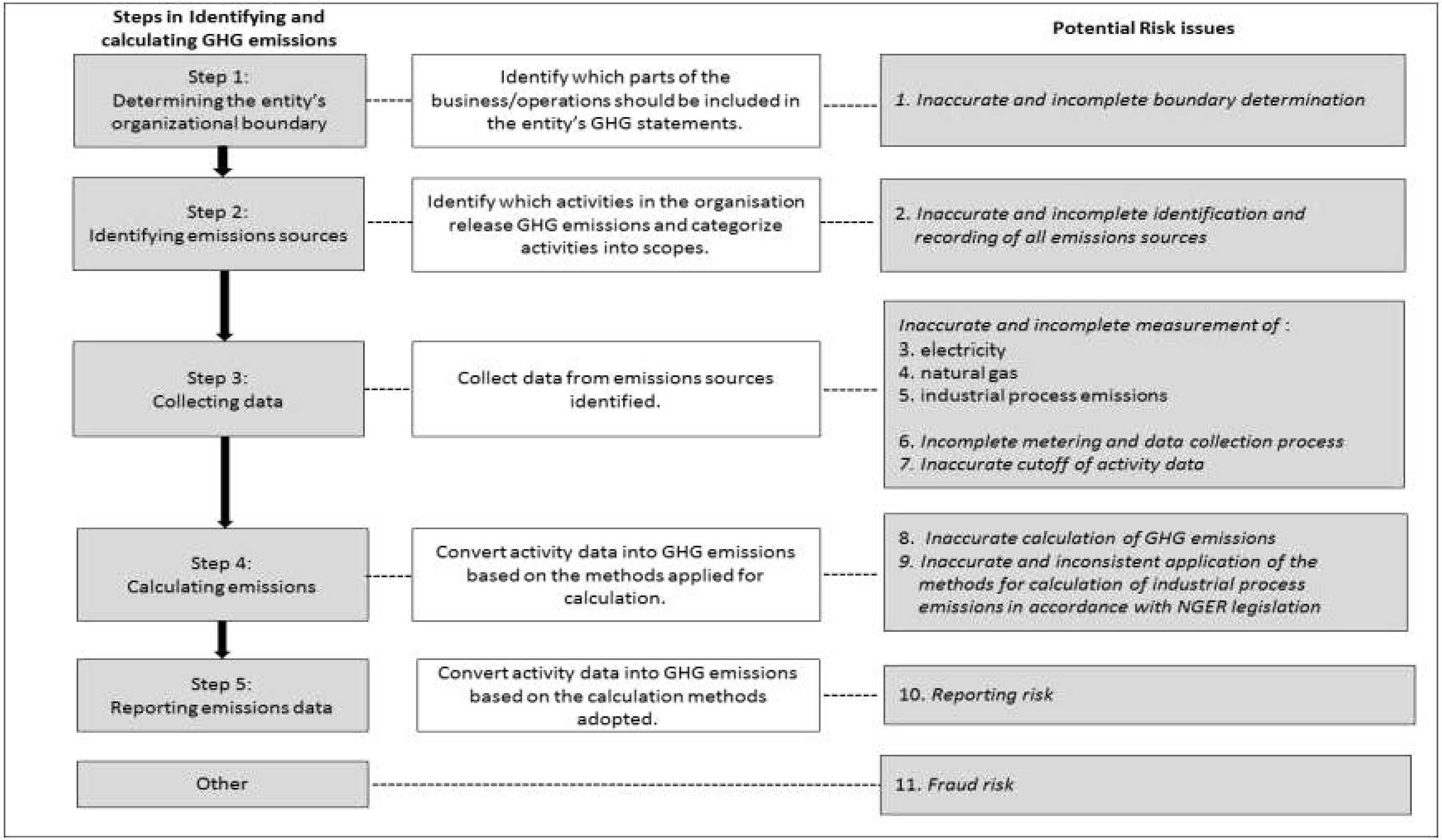


FIGURE 4.2 The Five Steps of Identifying and Calculating GHG Emissions and Potential Risk Issues

Step 2: Identifying Emissions Sources

This step involves identifying and categorising the GHG sources within the entity's organisational boundary. Emissions sources must be categorised into Scope 1[32] (direct emissions), Scope 2 (indirect emissions), and Scope 3 (other indirect emissions) because GHG calculation methods vary based on scopes of emissions. ISAE 3410 requires practitioners to obtain an understanding of the sources and completeness of emissions and recognise the risks associated with the *inaccurate and incomplete identification and recording of all emissions sources*. This category of risk is embedded in the case because the entity in the case has multiple sources of emissions, including Scope 1 and Scope 2 emissions.

Step 3: Collecting Data

This step involves collecting activity data from emissions sources identified in Step 2. The data required to calculate GHG emissions depends on the sources of emissions and the quantification methods adopted. The major source of emissions for the entity in the case is from Scope 2 emissions (i.e., electricity and natural gas). For Scope 2 emissions, activity data[33] is required to calculate the GHG emissions. With regard to Scope 1 emissions, data from steel production is required to calculate the GHG emissions. Fuel samples are analysed for carbon, energy, ash, or moisture content are used to determine carbon content factors. Therefore, carbon content and quantity (measured in tonnes) for all inputs, outputs, and wastes must be measured.

According to the various emissions sources and calculation methods used, five risk categories related to the measurement of emissions could be identified in the case: inaccurate and incomplete measurements of (1) *electricity*, (2) *natural gas*, and (3)

[32] The Greenhouse Gas Protocol defines Scope 1 emissions as direct GHG emissions from sources owned or controlled by the company, such as emissions from combustion in boilers and furnaces and emissions from chemical production. Scope 2 emissions are defined as indirect GHG emissions from the generation of electricity/gas purchased or transferred to the company and then consumed by the company. Scope 3 emissions are defined as other indirect emissions from a company's upstream and downstream activities and emissions associated with outsourced or contract manufacturing, leases, or franchises not included in Scope 1 and 2 (WBCSD and WRI 2004).

[33] Scope 2 activity data are converted into GHG emissions by applying designated emissions factors to activity data (activity data x emissions factor = GHG emissions). Examples of activity data collected are electricity and gas usage (e.g., total kilowatt hours used) from invoices, receipts, or meters.

industrial process emissions; (4) *incomplete metering and data collection process*; and (5) *inaccurate cut-off of activity data* (recording activity data in the wrong reporting period).

Step 4: Calculating Emissions

The NGER technical guidelines suggest a framework for selecting one of the four emissions calculation methods (Australian DCCEE 2010). Method 1 is the approach used most often to calculate GHG emissions; it applies designated emission factors to activity data (Defra 2009). Instead of applying published emission factors to activity data, Method 2 requires entities to undertake additional measurements, including the analysis of carbon content for various sources (e.g., carbon, energy, ash, or moisture content) to gain accurate estimates of emissions for particular facilities. Because the entity in the case adopted Method 1 for Scope 2 emissions and Method 2 for Scope 1 emissions, two categories of risks are embedded in the case: (1) *inaccurate calculation of GHG emissions*, and (2) *inaccurate and inconsistent application of the methods for calculation of industrial process emissions in accordance with NGER legislation*.

Step 5: Reporting Emissions Data

To report the entity's total GHG emissions, GHG data from different facilities must be gathered and summarised, particularly if the facilities are located in different countries and business units. Therefore, it is important to have a documented mechanism of collecting information and preparing reports and a control system to ensure the sufficient review and approval of reports and a consistent basis of preparation across different sites. These mechanisms therefore relate to *reporting risk*, and this risk is embedded in the case by including multiple facilities and locations.

Fraud Risk

In addition to the five steps involved in assessing reported GHG emissions, *fraud risk* is another risk category that could be identified in the case. ISAE 3410 suggests that misstatements in GHG statements can arise when incentives to under or overstate emissions are present. The incentives may result from the entity's climate change strategy or may be in connection with emissions trading markets. At the time of this study, the Emissions-Intensive Trade Exposed (EITE) scheme was being implemented in Australia to compensate industries affected by the carbon price.[34] Businesses can have an incentive to understate EITE activities to receive the same levels of assistance.[35] Given that the entity in this case is a steel-making company, which is a highly emissions-intensive activity (Australian CER 2013), this entity is eligible for EITE assistance; therefore, *fraud risk* related to the EITE assistance is embedded in the case.

4.4.3 Research Design

The experiment employed a 3x1 design. The independent variable is the team format which was manipulated at three levels: (1) nominal team, in which the ideas generated and decisions made by both the individual accountant and the individual non-accountant practitioners were combined without any interaction between participants, (2) interacting team, in which the individuals were asked to discuss their ideas and make decisions as a team comprising one accountant practitioner and one non-accountant practitioner, and (3) review team, in which the accountant practitioner was asked to review and add to the ideas generated by the non-accountant practitioner and to make a team decision without discussing the ideas. The employed design allowed the work of the individual non-accountant practitioner in the nominal teams to be reviewed by an accountant practitioner in the review team.

4.4.4 Procedures

The experimental sessions took place at the Sydney and/or Melbourne offices of the Big Four firms. Meeting rooms were provided in which participants could use a computer with internet access. Because the tasks were computerised, participants were told in

[34] The carbon pricing mechanism is an emissions trading scheme that puts a price on carbon pollution. In Australia, carbon pricing was introduced by the Clean Energy Act of 2011 and related legislation and was applied to Australia's biggest carbon emitters. However, the carbon pricing mechanism in Australia has now been revoked effective 1 July 2014 (Australian CER 2014). Companies conducting emissions-intensive activities (e.g., the steel-making industry) may receive a free allocation of carbon permits to compensate for their carbon liability. However, greater levels of assistance are provided for high emission-intensive activities than for moderate ones.

[35] The EITE assistance rates will decrease by a carbon productivity contribution of 1.3% per annum to encourage industry to reduce emissions (Deloitte 2011).

advance to bring their own computers to the room. The rooms were organised so participants in the nominal and review teams were in a separate room from the interacting teams. Participants in the nominal and review teams all worked on the task individually, so they could be seated in the same room.

At least one researcher (or research assistant) was in each room[36], and they started the experimental sessions at the same time, except for the review teams. Because the participants in this condition were asked to review other individuals' work from Stage 1 of the task, different arrival times were arranged for the reviewers. They were asked to arrive 20 minutes later than the participants in other conditions[37]. The 20-minute delay periods allowed individual non-accountants in the review teams to finish their first-stage task before their work was made available electronically to the accountant reviewer participants in the review treatment.

Before beginning the tasks, the participants were instructed to log on to the web-based instrument and read through the experimental study overview, definitions of some technical terms used in this study[38], and the instructions before they started reading the paper-based case materials. Paper-based case materials were used because they were easier for participants to refer to while working on the computer-based tasks. Appendix 3 provides the paper-based case materials, and Appendix 4 provides screen shots of each stage of the computerised experiment. The participants in each treatment were then instructed to begin reading the paper-based case materials and were given 10 minutes to read them.

All participants were given the same case and instructions for each treatment. They were informed that they were involved in the planning phase of a GHG assurance

---

[36] A small group of research assistants was recruited for each experimental session so that the author or at least one research assistant was in each room for every session of the experiment.

[37] The twenty-minute delay period for the reviewer treatment was operationalised in the first two experimental sessions only. In both sessions, review teams were seated in a separate room from the nominal teams. However, for the later sessions, there was no need to delay the participants in the review treatment as there was output from several non-accountant individuals left over to be reviewed from the previous sessions. Therefore, the individuals' output was ready for the reviewers to review from the beginning of the session.

[38] The definitions of three technical terms used in the instruments were provided to all participants including "assurance expert", "subject matter expert", and "risks of potential material misstatements".

engagement of SteelCo., which is a steel-making company. They were asked to assume that they were a member of a two-person team comprising one assurance expert and one subject matter expert,[39] which had been asked to provide input into the GHG risks assessment for SteelCo. They were then asked to generate as many risks of potential material misstatements as they could in the time allowed (20 minutes) and to rank their top four risks.[40]

Table 4.4 shows the four stages included in the experiment. As shown in the table, the participants were given 60 minutes to complete the experiment. The tasks in all stages were completed individually, except for participants in the interacting condition who completed Stages 1, 2, and 3 interactively through team discussion.

TABLE 4.4 Research Design and Experimental Procedures

	Experimental Treatment			
Procedures	**Nominal**	**Interacting**	**Review**	**Time**
Introduction				2 minutes
Read instruction pages	Individual	Individual	Individual	3 minutes
Read case material	Individual	Individual	Individual	10 minutes
Stage 1 Risk generation	Individual	Team	Individual	20 minutes
Stage 2 Risk selection	Individual	Team	Individual	3 minutes
Stage 3 Plan generation	Individual	Team	Individual	12 minutes
Stage 4 Post-experimental and demographic questions	Individual	Individual	Individual	10 minutes
Total				**60 minutes**

[39] In the instruction, the term "assurance expert" was used for "accountant" practitioner and "subject matter expert" was used for "non-accountant" practitioner. However, the definitions of "assurance expert" and "subject matter expert", which were provided to participants, were similar to the definitions of "accountant" and "non-accountant" practitioners used in this study. The results from the post-experimental question confirm that participants in interacting and review teams who had a chance to either interact with their team members or to review their team member's work perceived that their team member lived up to their expectation of an expert in the GHG assurance setting (Mean score=6.04, on a seven-point scale where 1=did not live up to expectations at all and 7=fully lived up to expectations).

[40] They were also asked to identify procedures to address each of their top four risks. Unfortunately, a number of participants did not complete the procedures for all top four risks. Further, a wide range of top four risks were selected, thus the procedures are not directly comparable. Thus, the procedures provided by participants are not reported in this study.

4.4.4.1 Stage 1: Risk Generation

Participants were given 20 minutes to complete the risk generation stage and were allowed to refer to the case during this stage. The following instructions were given to participants in each treatment:

Participants in treatment 1 (the nominal team) were asked to record a list of risks of potential material misstatements, rank their top four risks in order of significance, and identify the appropriate procedures for each of the top four risks. They were told that it was important to work independently and that the other team member was not present at the time.

Participants in treatment 2 (the interacting team) were told that they and another team member had been asked to discuss and record a list of risks of potential material misstatements, rank the top four risks in order of significance, and identify the appropriate procedures for each of the top four risks. They were told that it was important that they work as a team.

Participants in treatment 3 (the review team) were told that the other member of their team had already listed a set of risks of potential material misstatements for the entity in the case. They were asked to review their team members' list and create a joint list of risks of potential material misstatements by adding to their team members' list of risks as considered appropriate, rank their teams' top four risks in order of significance, and identify the appropriate procedures for each of the top four risks. They were told that it was important to work independently and that the other team member was not present at the time.

All participants were told to document as many risks as they could[41] in the time provided and not to leave the recording until the end of the period. The task instructions

[41] Initially, a maximum of 15 text boxes was provided to the participants to enter the risks they identified. However, results from the first experimental session showed that one interacting team was able to generate 15 risks in 20 minutes, while the other interacting teams and individuals were able to generate nine risks on average. Therefore, the number of text boxes provided in this task was adjusted up to 20 for the subsequent experimental sessions to avoid the ceiling effect that could limit the number of risks generated by the instrument.

also required participants to enter the risks of potential material misstatements in the boxes provided on their computer screen and to be as specific as possible. For the interacting teams, either the accountant or the non-accountant practitioner was randomly assigned to record the team answers. Participants in all treatments were also told that some participants' answers (without identifying any individual) would be selected for review by other members of their organisations.[42] Once the time elapsed (as shown in Table 4.4), the program automatically prompted a notification for participants to move on to the next task.

4.4.4.2 Stage 2: Risk Selection

Stage 2 was the risk selection stage. Participants were given three minutes to complete this stage, which asked them to rank[43] their top four risks of potential material misstatements in order of significance. The participants in the nominal treatment were asked to individually select their top four risks, while participants in the interacting treatment were asked to select their top four risks as a team. In the review treatment, accountant practitioners (reviewers) were asked to individually select their top four risks from the set of risks they added in Stage 1 (risk generation).[44] To facilitate this stage, the computer program presented the participants with the list of risks generated by themselves or by their team in Stage 1 and required them to enter the number "1", "2", "3", or "4" in the boxes provided next to each generated risk to identify the top four risks (1 = most significant; 4 = least significant).

4.4.4.3 Stage 3: Plan Generation

Stage 3 of the experiment involved identifying the appropriate audit procedures to address each of the top four risks selected in Stage 2. The task prompts required

---

[42] The answers provided by non-accountant practitioners in the nominal treatment were reviewed by accountant practitioners in the review treatment. The answers provided by some participants in the interacting teams were reviewed by the experts from Big Four firms.

[43] Using a rank ordering procedure in the selection task allows the comparison of the top four risks selected by the nominal, interacting, and review teams. The four risks that received the highest ranking within each team will be considered, i.e., the risks with 1, 2, 3, and 4 rankings for the interacting and review teams and the risks with 1 and 2 rankings for each member of the nominal teams.

[44] In Stage 4, reviewers in the review treatment were also asked to re-rank their top four risks from a pooled list of risks, which allows them to select risks generated by both themselves and non-accountant practitioners in their team (reviewees). This set of risks was used to test the sensitivity of results for RQ1.

participants to enter the procedures for each of their top four risks in the boxes provided and requested them to be as specific as possible. Participants were given 12 minutes to complete this stage[45].

4.4.4.4 Stage 4: Post-Experimental and Demographic Questions

Stage 4 involved answering post-experimental and demographic questions and was carried out by all individuals in all treatments. They were given 10 minutes to complete this stage. Because some questions were not applicable to every individual, the post-experimental questions were tailored for each treatment (See Appendix 4).

Participants in treatment 1 (nominal team*)* were asked to rate themselves on a seven-point scale (low/medium/high) in terms of their ability to identify risks and procedures and their knowledge of the subject matter in environmental reports and the audit criteria and processes. Participants in treatment 2 (interacting team) and the reviewers[46] in treatment 3 (review team) rated both themselves and their team members. Participants in treatment 2 and the reviewers in treatment 3 were asked to rate the overall performance of their team member on a seven-point scale ranging from one (poor performance) to seven (excellent performance) with a midpoint of four (moderate performance). The participants were also asked to rate the extent to which their team member lived up to their expectations for an expert in their team members' area on a seven-point scale ranging from one (did not live up to expectations at all) to seven (fully lived up to expectations) with a midpoint of four (moderately lived up to expectations).

The set of questions relating to team member behaviours was also asked to address research questions 2 to 4. To address research question 2, the participants in treatment 2 (interacting team) were asked about the other team member's behaviour regarding the elaboration of task-relevant information; this question follows the three-item self-reported measure used by Homan et al. (2007; 2008). The anchors on the seven-point scale were one (completely disagree) and seven (completely agree), with a midpoint of four (neither agree nor disagree). The level of cross-understanding between team

[45] Please see footnote 40

[46] This only happened for reviewers in the review team because the other team member (reviewees) did not have the opportunity to see the reviewers' work.

members in the interacting condition was assessed to address research question 3. The participants in treatment 2 were asked to rate the extent of the other team member's behaviours based on the 10-item self-reported measure suggested by Huber and Lewis (2010). The anchors on the seven-point scale were one (almost never) and seven (almost always) with a midpoint of four (sometimes). To address research question 4, the participants in all treatments were asked to assess their own (and their team member's, for treatments 2 and 3) ability to identify risks, knowledge of the subject matter, and knowledge of the audit criteria and process, all on a seven-point scale (1 = extremely low; 7 = extremely high).

All the participants were then asked to answer the following demographic questions: position in the firm, tertiary education background, industry specialisation, years of financial audit experience, years of conducting environmental/GHG/sustainability assurance engagements, number of environmental/GHG/sustainability assurance engagements undertaken, number of financial/GHG/sustainability assurance engagements for clients in manufacturing, and training hours or days of assurance for GHG emissions. Participants in treatment 2 (interacting team) were also asked to rate their familiarity with their assigned team member for the experiment on a seven-point scale before performing the team task (1 = not familiar at all, 4 = moderately familiar, and 7 = very familiar).

4.4.5 Dependent Variables

Several dependent variables of interest were present in this study. For the risk generation task, the types of risks generated by accountant and non-accountant practitioners in the nominal teams were used to test H1. Further, the quantity, breadth, and depth of risks generated by the nominal, interacting, and review teams were used to test H2 and H3. For the risk selection task, the types of risks selected by the accountant and non-accountant practitioners in the nominal teams were used to test H4. These dependent variables are discussed in detail in the following sections.

4.4.5.1 Types of Risks Generated

To analyse the types, quantity, breadth, and depth of risks generated and the types of risks selected, a coding system for risks generated by participants was established. The first step was to prepare a preliminary list of risks generated. The author identified 28 categories and 74 subcategories of risks. This preliminary list of risks was sent to the expert involved in the case material design to review whether the categories and subcategories of risks were appropriate (i.e., discrete and valid given the case materials). Because of this process, the preliminary list of risks was reclassified into 11 categories and 57 subcategories (see Table 4.5). Risks that did not fall into any of the 11 categories and that were infrequently generated were coded into an "other" category. Consequently, there were 12 risk categories in total. These risks include the risk issues discussed in Section 4.4.2.2 and a number of other risks listed by participants.

As described in Section 4.2.2, the risks of potential material misstatements in this study could be classified into two main types: (1) risks associated with *the measurement of the subject matter* and (2) risks associated with *comparing the subject matter with suitable criteria*. Risks associated with *the measurement of the subject matter* were defined as risks resulting from the measurement of Scope 1 emissions that are emitted into the atmosphere and the identification of emissions sources. Among the 12 categories and 57 subcategories, three categories and 17 subcategories (marked with an asterisk in Table 4.5) were classified as risks associated with measurements of the subject matter: (1) inaccurate and incomplete identification and recording of emissions sources; (2) inaccurate and incomplete measurements for industrial process emissions; and (3) inaccurate, insufficient, and incomplete metering and data collection process.

Risks associated with *comparing the subject matter with suitable criteria* were defined as risks resulting from noncompliance with NGER reporting criteria and accounting/audit criteria. Among the 12 categories and 57 subcategories, 10 categories and 40 subcategories were classified as risks associated with comparing the subject matter with suitable criteria: (1) inaccurate and incomplete boundary established for determining emissions; (2) inaccurate and incomplete measurement of electricity;[47] (3) inaccurate and incomplete measurement of natural gas; (4) incorrect cut-off; (5) inaccurate calculation of GHG emissions; (6) inaccurate and inconsistent application of methods of calculation of industrial process emissions; and (7) others (i.e., non-compliance with the EITE activity definitions and requirements, inappropriate estimation and disclosure of uncertainty, inappropriate presentation of criteria and methods disclosure in the GHG report, and incorrect prior year figures).

The other two categories, i.e., reporting risks and fraud risks, could not be categorised into either of the two main types. Thus, they were not included in the analysis of the types of risks generated.

All risks generated by the accountant and non-accountant practitioners in the nominal teams were categorised into (1) risks associated with *the measurement of the subject matter* and (2) risks associated with *comparing the subject matter with suitable criteria*. The number of risk subcategories for each type of risk was counted to determine differences in the types of risks generated between accountant and non-accountant practitioners.

4.4.5.2 Quantity of Risks Generated

The risks in the 12 categories and 57 subcategories generated by participants were classified as “total risks” and serve as the measure of the quantity of risks generated. The quantity of risks was assessed by counting the number of risk subcategories listed by each individual and those listed by the teams. The number of total risks for the nominal and review teams is the sum of the number of risk subcategories generated by each individual in the team after deleting redundancies. The number of total risks for the interacting teams is the number of risk subcategories generated as a team after deleting redundancies.

[47] Measurements of electricity and natural gas were considered as risks associated with comparing subject matter with suitable criteria (and not as the measurement of subject matter) because electricity and natural gas usage are mainly collected from invoices or receipts from third-party service providers. Thus, risks may arise from transcription errors, incomplete lists of invoices, and improper allocations.

TABLE 4.5 Categories and Subcategories of Risks

	Risk Categories and Subcategories
1	**Inaccurate and incomplete boundary established for determining emissions**
	1.1 The facility boundary and operational control has not been determined in accordance with the NGER legislation
	1.2 Inaccurate and incomplete boundary determination of operational control of all facilities in the overall group, small facilities, head office, data centres, and stockpile management (different locations)
	1.3 Inaccurate and incomplete boundary determination of operational control of contractors
	1.4 Inaccurate and incomplete boundary determination of operational control over off-site rolling mill
	1.5 Inaccurate and incomplete boundary determination of operational control over RecycleCo.
	1.6 Inaccurate and incomplete boundary determination of operational control over export or site transportation activities
***2**	**Inaccurate and incomplete identification and recording of emissions sources**
	2.1 Inaccurate and incomplete identification and recording of all sources of GHGs (i.e., different processes, sites, and corporate offices)
	2.2 Inaccurate and incomplete identification and recording of emissions from waste/slag and chemical reactions
	2.3 Inaccurate and incomplete identification and recording of emissions from on-site contractors (e.g., supply of fuel data, LNG, diesel)
	2.4 Inaccurate and incomplete identification and recording of emissions from the NSW rolling mill (located off-site)
	2.5 Inaccurate and incomplete identification and recording of emissions from the casting and refining process
	2.6 Inaccurate and incomplete identification and recording of emissions from transportation activities or fuel use for vehicles used on and between sites (diesel, petrol)
	2.7 Inaccurate and incomplete identification and recording of emissions from oils, greases, non-combusted emissions
	2.8 Inaccurate and incomplete identification and recording of emissions from stationary combustion sources, electrodes, and energy production
	2.9 Inaccurate and incomplete identification and recording of emissions from nitrogen, coke, and acetylene
	2.10 Inaccurate and incomplete identification and recording of emissions from plant emergencies, start-up, and shutdown
3	**Inaccurate and incomplete measurement of electricity usage**
	3.1 Inaccurate and incomplete measurement of electricity usage
	3.2 Inaccurate and incomplete measurement of electricity used by backup generators or power plant
	3.3 Inaccurate and incomplete measurement of electricity used by electrodes, electric arc furnace, and casting process (e.g., improper allocations)
	3.4 Inaccurate and incomplete reporting on electricity consumption (e.g., transcription error, incomplete list of invoices, or inaccurate invoice by third-party service providers)

TABLE 4.5 (Continued) Categories and Subcategories of Risks

	Risk Categories and Subcategories
4	**Inaccurate and incomplete measurement of natural gas usage**
	4.1 Inaccurate and incomplete measurement of natural gas usage
	4.2 Inaccurate and incomplete measurement of natural gas used by GlassCo.
	4.3 Inaccurate and incomplete measurement of natural gas used by on-site contractors for welding and forklift activities
	4.4 Inaccurate and incomplete measurement of natural gas used by RecycleCo.
	4.5 Inaccurate and incomplete measurement of natural gas used by rolling mills (e.g., lost in transport due to NSW rolling mill being located off-site)
	4.6 Inaccurate and incomplete measurement of natural gas used to monitor the chemical composition of the melt
	4.7 Inaccurate and incomplete report on natural gas consumption (e.g., transcription error, incomplete list of invoices, or inaccurate invoice by third-party service providers)
***5**	**Inaccurate and incomplete measurement for industrial process emissions**
	5.1 Inaccurate and incomplete measurement of industrial process emissions
	5.2 Inaccurate and incomplete measurement of the input data (e.g., volume and carbon content) used in the carbon mass balance calculation
	5.3 Inaccurate and incomplete measurement of output data (e.g., volume and carbon content) used in the carbon mass balance calculation
	5.4 Inaccurate and incomplete measurement of emissions from chemical reaction and waste by-product
***6**	**Inaccurate, insufficient, and incomplete metering and data collection process**
	6.1 Inaccurate and insufficient meter reading, maintenance and calibration-energy consumption usage
	6.2 Inaccurate and insufficient meter reading, maintenance, and calibration (other than energy consumption usage meters)
	6.3 Inaccurate and incomplete meters being selected for measurement at each site and across sites (i.e., duplication of measurement or not captured)
7	**Inaccurate cut-off**
8	**Inaccurate calculation of GHG emissions**
	8.1 The methods applied for the calculation of greenhouse gas emissions, energy produced, and energy consumed are not correctly applied or not in accordance with the NGER legislation
	8.2 Inaccurate invoice by third-party service providers due to estimate amounts or inaccurate capture
	8.3 Inaccurate emission factors applied
	8.4 Inaccurate emissions factors applied across different facilities and states
	8.5 Inappropriate and inconsistent underlying assumptions applied when calculating emissions
	8.6 Inaccurate calculation (not mathematically accurate or wrong formulas in spreadsheets)
	8.7 Inaccurate and inconsistent conversion of energy sources or unit conversion applied
	8.8 Inaccurate data entry, data processing, or transcription from source documentation
	8.9 Inaccurate figures, weighting, or calibration of input and output data
	8.10 The uncertainty of method 1 and 2 inputs are not calculated and not validated in accordance with the NGER legislation

TABLE 4.5 (Continued) Categories and Subcategories of Risks

	Risk Categories and Subcategories
9	**Inaccurate and inconsistent application of the methods for calculation of industrial process emissions in accordance with NGER legislation**
	9.1 Inaccurate adjustments to the chemical composition measured and recorded due to variances in the materials or inputs used (i.e., uncertainty around the EAF)
	9.2 Method 2: Unreasonable and inappropriately calculated Method 2 emission factors in accordance with the NGER Act
	9.3 Method 2: Analysis of fuels for carbon, energy, ash, or moisture content is inaccurate, incomplete, or not in accordance with Australian or equivalent International Standards
	9.4 Method 2 is not correctly or consistently applied to calculate greenhouse gas emissions, energy produced, and energy consumed or not in accordance with the NGER legislation
	9.5 Not complying with the technical sampling requirements for calculating the process emissions via the carbon mass balance (representative and unbiased sampling)
10	**Reporting risk**
	10.1 Inadequate, poorly controlled, or poorly documented mechanism for collecting data, quantifying emissions, and preparing GHG statements
	10.2 Lack of appropriate review and approval of reports, assumptions, and calculations
	10.3 Incomplete and inconsistent basis of preparation, data collection, and measures from and across different sites
	10.4 Lack of staff competence in collecting data, quantifying emissions, and preparing GHG statements
	10.5Failure to report audited energy consumption and emissions to the regulator or any other public disclosures
11	**Fraud risk**
	11.1 Incentive to overstate or understate emissions due to fraud or management bias
	11.2 Incentive to overstate emissions due to EITE assistance or activity
	11.3 Incentive to understate emissions due to the benefit of less emissions when set against free carbon permits provided by EITE
12	**Other**

Notes:

- This table presents the list of risk categories and subcategories generated by participants in all treatments
- Items with an asterisk (*) are the three categories and 17 subcategories classified as risks associated with the measurement of the subject matter
- Other categories and subcategories (other than categories 10 and 11) are classified as risks associated with comparing the subject matter with the suitable criteria

4.4.5.3 Breadth and Depth of Risks Generated

Consistent with the prior literature (e.g., Nijstad et al. 2002; Dahlin et al. 2005; Kohn and Smith 2011), measures of the *breadth* and *depth* of risks were used to examine the qualitative aspects of MDGHGT performance.[48] The *breadth* of risk area coverage was calculated by determining the number of risk categories generated. The *depth* of risk refers to how many risk subcategories were generated within a category and is calculated by dividing the number of total risks by the number of risk categories generated.

4.4.5.4 Types of Risks Selected

Consistent with Rietzschel et al. (2006), participants were asked in Stage 2 to select the top four risks from the subcategories they generated in Stage 1 based on the significance of the risks. Because the rankings were completed independently for the nominal teams, the top four subcategories for the team was derived by including those ranked 1 and 2 from each team member's list. For the interacting teams, the top four subcategories with 1, 2, 3, and 4 rankings were included. For the review teams, the top four subcategories were derived by including items ranked 1, 2, 3, and 4 from the reviewer's list.[49]

To address Hypothesis 4, which focuses on the difference in the types of risks selected between accountant and non-accountant practitioners, the 'top four risks' selected by each accountant and non-accountant practitioner in the nominal teams were included. Similar to the types of risks generated in 4.4.5.1, the 'top four risks' selected could be classified into two types: (1) risks associated with *the measurement of the subject matter* and (2) risks associated with *comparing the subject matter with suitable criteria*. The top four subcategories selected by the accountant (accountant-selected risks) and non-

---

[48] The quality of risks generated is not measured because no clear measure of quality was available, i.e., the aim was to identify all identifiable risks and understand how MDGHGTs utilise their diverse knowledge and perspectives; therefore, the quantity, breadth, and depth of risks generated are the key relevant measures.

[49] As in later stages, reviewers were asked to re-rank their top four risks from a pooled list of risks (i.e., select risks generated by themselves and non-accountant practitioners in their team). The top four risks selected from the pooled list with 1, 2, 3, and 4 rankings were used to test the sensitivity of results for RQ1.

accountant (non-accountant-selected risks) practitioners in each of the two risk types were compared.

4.4.6 Coding

A coding scheme for the risks generated by participants was initially developed based on the five steps of identifying and calculating GHG emissions and the potential risk issues presented in Figure 4.2 and by pre-coding the risks from all participants. This process resulted in 28 risk categories and 74 subcategories. This preliminary list of risks was sent to the Big Four expert involved in designing the case material to review and assess whether the categories and subcategories were appropriate. The preliminary list of risks was reclassified by the expert into 12 categories and 57 subcategories.[50]

Because the experimental instruments were computerised, the participants had already typed up all the responses. A number was randomly allocated to all responses so the coders were not able to identify which treatment the responses belonged to. The coding of risks generated by participants was performed by the author and one research assistant blind to the hypotheses. This research assistant is a PhD student who has auditing experience and is knowledgeable in GHG assurance. The two coders initially discussed the coding instructions to clarify the procedures and then individually coded 10 responses based on the coding scheme that had been agreed on with the expert. After the codes were compared, both coders discussed any disagreements until they reached consensus. They then coded the rest of the responses. The coders discussed and resolved all differences in their coding. The final agreed-upon coding was used in the analysis. The inter-rater reliability was measured using Cohen's (1960) kappa. The kappa coefficient was 0.79 for the categories and subcategories of risks generated, which represents a high reliability between coders.

[50] This coding scheme was refined through several rounds of discussion with the expert. After deleting redundancies, merging several categories, and reclassifying some categories into subcategories, the number reduced from 28 to 12 categories (74 to 57 subcategories).

4.5 Results

4.5.1 Types of Risks Generated: Hypothesis 1

Hypotheses 1 examines the difference in the **types** of risks **generated** between the individual accountant and non-accountant practitioners. Accountant practitioners are expected to generate a higher proportion of risks associated with *comparing the subject matter with suitable criteria*, and a lower proportion of risks associated with *the measurement of the subject matter* than accountant practitioners. To address these hypotheses, risk categories generated by the 12 accountant and 12 non-accountant practitioners in the nominal treatment are categorised into two types: (1) risks associated with *the measurement of the subject matter* and (2) risks associated with *comparing the subject matter with suitable criteria*.[51] Table 4.6 provides descriptive statistics with a list of the types and number of risks generated by the accountant and non-accountant practitioners in the nominal treatment. The accountant (non-accountant)-generated risks are calculated by summing the number of subcategories generated by the 12 accountant (non-accountant) practitioners in each risk category.

As reported in Table 4.6, 29 (46 percent) of the 134 risks generated by the accountant practitioners are associated with the *measurement of the subject matter* compared with 38 (57 percent) of the 119 risks generated by non-accountant practitioners. At the same time, 105 (56 percent) of the 134 risks generated by the accountant practitioners are associated with *comparing the subject matter with suitable criteria* compared with 81 (44 percent) of the 119 risks generated by non-accountant practitioners. These proportions are significantly different ($\chi^2 = 3.428$, $p = 0.032$, one-tailed) and are in the expected direction. The results indicate that differences are present in the types of risks generated by practitioners from different educational backgrounds, with accountant practitioners generating more risks associated with *comparing the subject matter with suitable criteria*, and fewer risks associated with *the measurement of the subject matter* than non-accountant practitioners. Thus, Hypothesis 1 is supported.

[51] Two categories generated by the individuals in the nominal treatment, i.e., reporting risks and fraud risks, could not be categorised into the two main types of risks and are not reported in Table 4.6.

TABLE 4.6 Types of Risks of Potential Material Misstatements Generated by Individual Accountant and Non-Accountant Practitioners in the Nominal Teams

Descriptive Statistics – Frequency and Percentage

Types of Risk Categories	**Number of Total Risks Generated (Percentage)** Total	Accountant-Generated Risks (*n* = 12)	Non-Accountant-Generated Risks (*n* = 12)
Risks associated with the measurement of the subject matter:			
1. Inaccurate and incomplete identification and recording of emissions sources	28 (100%)	14 (50%)	14 (50%)
2. Inaccurate and incomplete measurement for industrial process emissions	27 (100%)	12 (44%)	15 (56%)
3. Inaccurate, insufficient, and incomplete metering and data collection process	12 (100%)	3 (25%)	9 (75%)
Total	**67 (100%)**	**29 (43%)**	**38 (57%)**
Risks associated with comparing the subject matter with suitable criteria:			
1. Inaccurate calculation of GHG emissions	49 (100%)	27 (55%)	22 (45%)
2. Inaccurate and incomplete measurement of natural gas (Scope 2 emissions/indirect measurement)	39 (100%)	25 (64%)	14 (36%)
3. Incorrect and inconsistent application of methods to calculate industrial process emissions in accordance with NGER legislation	39 (100%)	24 (62%)	15 (38%)
4. Incorrect boundary established for determining emissions	38 (100%)	16 (42%)	22 (58%)
5. Inaccurate and incomplete measurement of electricity (Scope 2 emissions/indirect measurement)	12 (100%)	7 (58%)	5 (42%)
6. Incorrect cut-off (emissions have been recorded in the wrong reporting period)	5 (100%)	2 (40%)	3 (60%)
7. Other	4 (100%)	4 (100%)	0 (0%)
Total	**186 (100%)**	**105 (56%)**	**81 (44%)**
Total Number of Risks Generated	**253**	**134**	**119**

Notes:

- This table presents the descriptive statistics and analysis for the types of risks generated by individual accountant and non-accountant practitioners in the nominal teams and includes risks associated with *the measurement of the subject matter* and risks associated with *comparing the subject matter with suitable criteria*. This table provides a list of risk categories and the number of 'total risks', 'accountant-generated risks', and 'non-accountant-generated risks' within each category.
- Risks associated with *the measurement of the subject matter* are defined as risks resulting from the measurement of Scope 1 emissions that are emitted into the atmosphere and the identification of emissions sources.
- Risks associated with *comparing the subject matter with suitable criteria* are defined as risks resulting from non-compliance with NGER reporting criteria and accounting/audit criteria.

4.5.2 Quantity, Breadth, and Depth of Risks Generated: Hypotheses 2 and 3

Table 4.7 provides a list of 12 risk categories generated in stage 1 (i.e., risk generation) by the nominal, interacting, and review teams and describes the risks generated by these teams. The numbers without parentheses represent the number of total risks generated within each risk category and are calculated by summing the number of subcategories[52] generated within each category by the teams in each treatment. The numbers with parentheses represent the *breadth of risks* generated, which is the number of risks generated in each risk category. The mean numbers of *total risks* and the *breadth of risks* generated by teams in each treatment are presented in the final row of the table.

The descriptive information reported in Table 4.7 reveals that the main differences in risk generation between the three team formats are between the mean number of total risks generated by the interacting team and by the other two teams (interacting = 11.75, nominal = 18.67, and review = 18.25). However, the mean number of categories *(*breadth of risks*)* generated by the three team formats is quite similar (interacting = 7.67, nominal = 8.75 and review = 8.92). The differences in the quantity of risks generated by the nominal versus the interacting teams are addressed in Hypothesis 2a, while the differences between the nominal and interacting teams versus the review teams are addressed in Hypotheses 3a and 3b.

4.5.2.1 The Effect of Team Interaction on Risk Generation: Hypothesis 2

Hypothesis 2a examines the effect of interactions between accountant and non-accountant practitioners on the **quantity** of risks **generated**, i.e., whether the interacting teams generate a lower number of total risks than the nominal teams. The descriptive statistics for risks generated by the 36 teams are provided in Panel A of Table 4.8. Overall, the mean number of total risks generated by the interacting teams is lower than for the nominal teams (11.75 vs. 18.67). The independent-samples comparison in Panel B of Table 4.8 confirms that teams in the interacting treatment generate a significantly lower number of total risks compared to those in the nominal treatment ($t = -4.507$, $p = 0.000$, one-tailed), providing support for Hypothesis 2a.

[52] Across all categories, 57 subcategories were generated. Details for the 57 subcategories generated are provided in Table 4.5.

TABLE 4.7 Number of Teams Generating Each Category and Number of Total Risks Generated by Teams in the Risk Generation Stage (Stage 1)

	Risk Categories	Number of Total Risks Generated within Each Category (Number Categories Generated: Breadth of Risks)			
		Nominal Team (*n* = 12)	**Interacting Team (*n* = 12)**	**Review Team (*n* = 12)**	**Total Teams (*N* = 36)**
1	Inaccurate boundary established for determining emissions	28 (12)	22 (12)	29 (12)	79 (36)
2	Inaccurate and incomplete identification and recording of emissions sources	27 (11)	14 (8)	27 (11)	68 (30)
3	Inaccurate and incomplete measurement of electricity usage	12 (8)	7 (6)	11 (9)	30 (23)
4	Inaccurate and incomplete measurement of natural gas usage	28 (12)	21 (12)	22 (12)	71 (36)
5	Inaccurate and incomplete measurement of industrial process emissions	20 (11)	10 (7)	17 (8)	47 (26)
6	Inaccurate, insufficient, and incomplete metering and data collection process	11 (7)	11 (9)	16 (9)	38 (25)
7	Inaccurate cut-off	5 (5)	5 (5)	4 (4)	14 (14)
8	Inaccurate calculation of GHG emissions	43 (12)	19 (9)	39 (11)	101 (32)
9	Inaccurate and inconsistent application of methods of calculating industrial process emissions in accordance with NGER legislation	31 (12)	19 (12)	28 (12)	78 (36)
10	Reporting risk	12 (8)	6 (5)	17 (11)	35 (24)
11	Fraud risk	2 (2)	1 (1)	3 (2)	6 (5)
12	Other	5 (5)	6 (6)	6 (6)	17 (17)
	Total	**224 (105)**	**141 (92)**	**219 (107)**	**584 (304)**
	Mean Number of *Total Risks* (Mean Number of Risk Categories) generated per team - as per Table 4.8	**18.67 (8.75)**	**11.75 (7.67)**	**18.25 (8.92)**	

Notes:

- This table presents the risk categories generated and the number of total risks generated within each category by the nominal, interacting, and review teams. The number shown in parentheses represents the number of categories generated by the three teams (breadth of risks). The total risks are the sum of the number of subcategories generated by the nominal, interacting, and review teams.
- Nominal: team in which no communication occurred between two team members.
- Interacting: team in which verbal communication occurred between two team members.
- Review: team in which no verbal communication occurred between two team members, but the accountant team member has a chance to see the risks generated by the non-accountant team member before adding to the list of risks.

Hypothesis 2b predicts that the interacting teams will **generate** risks across fewer risk categories (i.e., less **breadth** of categories**)** than the nominal teams. The *breadth of risks* generated is measured by the number of risk categories generated by teams. As shown in Panel A of Table 4.8, the interacting teams generated 7.67 categories, while the nominal teams generated 8.75 categories on average. The analysis of the *breadth of risks* in Panel B shows that the number of categories generated by the interacting teams is only marginally significantly different from the nominal teams ($t = -1.622$, $p = 0.060$, one-tailed), thus Hypothesis 2b is marginally supported.

Hypothesis 2c predicts that the interacting teams will **generate** fewer risks in each risk category (i.e., less **depth**) than the nominal teams. The *depth of risks* generated is calculated by dividing the number of total risks by the number of categories generated by teams. Panel A of Table 4.8 reports that the interacting teams generated fewer risks within the same category than the nominal teams (1.53 vs. 2.16). The comparison between the interacting and nominal teams (Panel B of Table 4.8) shows that this difference is highly significant ($t = -4.520$, $p = 0.000$, one-tailed), thus providing support for Hypothesis 2c.

In summary, Hypotheses 2a and 2c are supported. Interacting teams generate fewer risks (H2a) and have less depth within categories (H2c) than nominal teams. However, only a marginal difference is found between the interacting and nominal teams for the breadth of risks generated (H2b), thus Hypothesis 2b is marginally supported.

TABLE 4.8 Comparison of Mean Number of Risks of Potential Material Misstatements <u>Generated</u> by Teams in the Risk Generation Stage (Stage 1)

Panel A. Descriptive Statistics: Mean

	Mean Number of Risks Generated by Teams		
Treatment	**Total Risks**	**Breadth of Risks**	**Depth of Risks**
Nominal (***n*** = 12)	18.67	8.75	2.16
Interacting (***n*** = 12)	11.75	7.67	1.53
Review (***n*** = 12)	18.25	8.92	2.04

Panel B. Independent-Samples Comparison (*t*-Test)

	t-statistics and p-values*		
Treatment	**Total Risks**	**Breadth of Risks**	**Depth of Risks**
Interacting vs. Nominal	-4.507, p = 0.000	-1.622, p = 0.060	-4.520, p = 0.000
Review vs. Interacting	4.776, p = 0.000	2.175, p = 0.021	5.218, p = 0.000

Panel C. Paired-Samples Comparison (*t*-Test)

	t-statistics and p-values*		
Treatment	**Total Risks**	**Breadth of Risks**	**Depth of Risks**
Review vs. Nominal	-0.439, p = 0.335	0.304, p = 0.384	-1.043, p = 0.160

Notes:
- This table presents the descriptive statistics (Panel A) and analysis for the quantity, breadth, and depth of risks generated by the nominal, interacting, and review teams (Panels B and C).
- Nominal: team in which no communication occurred between two team members.
- Interacting: team in which verbal communication occurred between two team members.
- Review: team in which no verbal communication occurred between two team members, but an accountant team member has a chance to see the risks generated by a non-accountant team member before adding to the list of risks.
- *Total risks* are the number of subcategories generated by each team out of 57 subcategories.
- *Breadth of risks* is the number of categories generated by each team out of 12 categories.
- *Depth of risks* is calculated by dividing the *total risks* by the *breadth* of risks.
- An independent-samples t-test is used to determine the statistical differences between the interacting and nominal teams and the review and interacting teams, while a paired-samples t-test in Panel B is used to determine statistical differences between review and nominal teams.

*All reported p-values are one-tailed.

4.5.2.2 The Effect of the Review Process on Risk Generation: Hypothesis 3

Hypotheses 3a and 3b examine the effect of the review process between the accountant and non-accountant practitioners on the **quantity** of risks **generated**, that is, whether the review teams generate a larger number of total risks than the interacting teams (H3a) and the nominal teams (H3b). It should be noted that the review and nominal teams have the same person as the non-accountant. Therefore, a paired-samples t-test is used to determine the statistical differences between the review and nominal teams. The descriptive statistics of risks generated by the teams in all treatments are provided in Panel A of Table 4.8. The mean number of total risks generated by the review teams is higher than for the interacting teams (18.25 vs. 11.75) but lower than for the nominal teams (18.25 vs. 18.67). As shown in Panel B of Table 4.8, the differences in the number of total risks generated between the review and interacting teams is highly significant ($t = 4.776$, $p = 0.000$, one-tailed), thus providing support for Hypothesis 3a. However, the results reported in Panel C show no significant difference between the review and nominal teams ($t = -0.439$, $p = 0.335$, one-tailed), thus Hypothesis 3b is not supported. The findings indicate that although the review process enhances MDGHGTs' risk generation performance compared to interacting teams, the review process does not outperform nominal teams.

Hypotheses 3c and 3d predict that review teams will **generate** risks with more **breadth** than the interacting (H3c) and nominal (H3d) teams. As shown in Panel A of Table 4.8, the review teams generated a larger number of risk categories (*breadth of risks*) than the interacting teams (8.92 vs. 7.67) and nominal teams (8.92 vs. 8.75). Panel B of Table 4.8 reports that the difference in the breadth of risks generated between the review and interacting teams is statistically significant ($t = 2.175$, $p = 0.021$, one-tailed), thus providing support for Hypothesis 3c. However, as shown in Panel C of Table 4.8, no significant difference is found for the breadth of risks generated between the review and nominal teams ($t = 0.304$, $p = 0.384$, one-tailed). Thus, Hypothesis 3d is not supported.

Hypotheses 3e and 3f predict that the review teams will **generate** risks with more category **depth** than the interacting teams (H3e) but with less category depth than the nominal teams (H3f). The significant difference reported in Panel B of Table 4.8 indicates that the review teams generated significantly more risks within the same category than the interacting teams (2.04 vs. 1.53, $t = 5.218$, $p = 0.000$). Thus, Hypothesis 3e is supported. Panel C of Table 4.8 shows no significant difference between the review and nominal teams in terms of the depth of risks generated (2.04 vs. 2.16, $t = -1.043$, $p = 0.160$). Thus, Hypotheses 3f is not supported.

In summary, all the hypotheses related to the comparison of the quantity, breadth, and depth of risks generated between the review and interacting teams are supported. The review teams generated more risks (H3a) with greater breadth (H3c) and greater depth (H3e) than the interacting teams. However, contrary to expectations, the hypotheses related to the comparison between the review and nominal teams are not supported. No significant differences are found for the quantity (H3b), breadth (H3d), and depth of risks generated (H3f) between the two treatments.

4.5.3 Types of Risks Selected by Accountant and Non-Accountant Practitioners: Hypothesis 4

Table 4.9 lists the 12 risk categories **selected** in the risk selection stage (Stage 2) by the nominal, interacting, and review teams and the number of the 'top four risks' selected within each category. The top four risks are calculated by summing the number of subcategories selected by the teams in each category for each treatment.

To compare the types of risks selected between the three conditions, the four risks that received the highest ranking within each group are determined as follows. For the interacting teams, the top four subcategories with rankings of 1, 2, 3, and 4 are included. For the nominal teams, the top four subcategories with rankings of 1 and 2 are included for each team member. For the review teams, the top four subcategories are derived by including those ranked 1, 2, 3, and 4 from the reviewers' lists.

The descriptive information reported in Table 4.9 indicates that differences exist for risk category 3, in which interacting teams tend to select more risks related to the inaccurate and incomplete measurement of electricity usage, which is the major source of emissions for the entity in the case, than the nominal and review teams (6 vs. 3 and 2, respectively). However, these differences are not statistically significant. The differences in the types of risks selected between the accountant and non-accountant practitioners in the nominal teams are further examined in Hypotheses 4.

TABLE 4.9 Number of Risks Selected by Teams in the Risk Selection Stage (Stage 2)

	Risk Categories	Number of *Top Four Risks* Selected in Each Category			
		Nominal Team (*n* = 12)	**Interacting Team (*n* = 12)**	**Review Team (*n* = 12)**	**Total Teams (*N* = 36)**
1	Inaccurate boundary established for determining emissions	9	9	12	30
2	Inaccurate and incomplete identification and recording of emissions sources	6	5	7	18
3	Inaccurate and incomplete measurement of electricity usage	3	6	2	11
4	Inaccurate and incomplete measurement of natural gas usage	10	7	7	24
5	Inaccurate and incomplete measurement of industrial process emissions	5	3	2	10
6	Inaccurate, insufficient, and incomplete metering and data collection process	2	4	3	9
7	Inaccurate cut-off	0	1	0	1
8	Inaccurate calculation of GHG emissions	6	4	8	18
9	Inaccurately and inconsistently applied methods of calculating industrial process emissions in accordance with NGER legislation	5	7	7	19
10	Reporting risk	0	1	0	1
11	Fraud risk	1	0	0	1
12	Non-compliance with EITE activity definitions and requirements	1	1	0	2
	Total	**48**	**48**	**48**	**144**

Notes:
- This table presents the list of risk categories selected and the number of top four risks selected within each category. The top four risks were calculated by summing the number of subcategories selected as the top four risks by the nominal, interacting, and review teams in each category (each team could select more than one risk from the same category).
- For the interacting teams, the top four subcategories with rankings of 1, 2, 3, and 4 were assessed. For the nominal teams, the top four subcategories with rankings of 1 and 2 (from each team member's list) were assessed. For the review teams, the top four subcategories the reviewers selected with rankings of 1, 2, 3, and 4 were included.
- Nominal: team in which no communication occurred between two team members.
- Interacting: team in which verbal communication occurred between two team members.
- Review: team in which no verbal communication occurred between two team members, but an accountant team member has a chance to see the risks generated by a non-accountant team member before adding to the list of risks.

Hypothesis 4 examines the difference in the **types** of risks **selected** in the risk selection stage (Stage 2) between individual accountant and non-accountant practitioners in the nominal treatment. Specifically, this hypothesis addresses whether the accountant practitioners select more risks associated with *comparing the subject matter* with suitable criteria, fewer risks associated with the *measurement of the subject matter* than non-accountant practitioners. In Stage 2, the participants were asked to select the four most significant risks; thus, the risks selected by individuals in the nominal teams with rankings of 1 to 4 were included. To address this hypothesis, risk categories selected by the 12 accountant and 12 non-accountant practitioners in the nominal treatment were categorised into two types: (1) risks associated with *the measurement of the subject matter* and (2) risks associated with *comparing the subject matter with suitable criteria*.[53] Table 4.10 provides descriptive statistics, with a list of the number and types of the top four risks selected by the accountant and non-accountant practitioners. The accountant-selected (non-accountant-selected) risks are calculated by summing the number of subcategories from within the same category that were selected as the top four risks by the 12 accountant (non-accountant) practitioners.

As shown in Table 4.10, 10 (40 percent) of the 46[54] risks selected by the accountant practitioners were associated with the *measurement of the subject matter* compared with 15 (60 percent) of the 45 risks selected by non-accountant practitioners. At the same time, 36 (61 percent) of the 46 risks selected by accountant practitioners were associated with *comparing the subject matter with suitable criteria* compared with 30 (39 percent) of the 45 risks selected by non-accountant practitioners. While the results are in the expected direction, these proportions are not significantly different (χ^2 = 1.535, p = 0.108, one-tailed). The results indicate that accountant and non-accountant practitioners do not select different types of risks. Thus, Hypothesis 4 is not supported.

[53] Three categories generated by the individuals in the nominal treatment were not included in the analysis because one category (i.e., other) was not selected by the nominal teams and two categories (i.e., reporting risks and fraud risks) were selected but could not be categorised into the two types of risks. Thus, these three risks are not reported in Table 4.10.

[54] As mentioned in the previous footnote, two categories were selected but could not be categorised into the two types of risks. Consequently, the proportions were calculated based on 46 risks for accountant practitioners and 45 for non-accountant practitioners rather than 48 (i.e., four risks were selected by 12 accountant and non-accountant practitioners).

TABLE 4.10 Types of Risks of Potential Material Misstatements Selected by Individual Accountant and Non-Accountant Practitioners in the Nominal Teams

Descriptive Statistics: Frequency and Percentage

Types of Risks (Categories)	**Number of Top Four Risks Selected (Percentage)** **Total Selected Risks (*N* = 24)**	**Accountant-Selected Risks (*n* = 12)**	**Non-Accountant-Selected Risks (*n* = 12)**
Risks associated with the measurement of the subject matter:			
1. Inaccurate and incomplete measurement of industrial process emissions	10 (100%)	7 (70%)	3 (30%)
2. Inaccurate and incomplete identification and recording of emissions sources	9 (100%)	2 (22%)	7 (78%)
3. Inaccurate, insufficient, and incomplete metering and data collection process	6 (100%)	1 (17%)	5 (83%)
Total	**25 (100%)**	**10 (40%)**	**15 (60%)**
Risks associated with comparing the subject matter with suitable criteria:			
1. Inaccurate and incomplete measurement of natural gas (Scope 2 emissions/indirect measurement)	19 (100%)	12 (63%)	7 (37%)
2. Incorrect boundary established for determining emissions	15 (100%)	6 (40%)	9 (60%)
3. Inaccurate calculation of GHG emissions	13 (100%)	6 (46%)	7 (54%)
4. Inaccurate and incomplete measurement of electricity (Scope 2 emissions/indirect measurement)	7 (100%)	4 (57%)	3 (43%)
5. Incorrect cut-off (emissions recorded in the wrong reporting period)	1 (100%)	1 (100%)	0 (0%)
Total	**66 (100%)**	**36 (61%)**	**30 (39%)**
Total Number of Risks Selected	**91**	**46**	**45**

Notes:

- This table presents the descriptive statistics and analysis for the types of the top four risks selected by individual accountant and non-accountant practitioners in the nominal teams. Panel A provides the list of risk categories selected and the number of the top four risks selected within each category. The top four risks were calculated by summing the number of subcategories selected as the top four risks by 12 accountant (accountant-selected risks) and 12 non-accountant (non-accountant-selected risks), which fell into the same category (each practitioner could select more than one risk from the same category).
- Risks associated with the measurement of the subject matter are defined as risks resulting from the measurement of the Scope 1 emissions that are emitted into the atmosphere and the identification of emissions sources.
- Risks associated with comparing the subject matter with suitable criteria are defined as risks resulting from non-compliance with NGER reporting criteria and accounting/audit criteria.

TABLE 4.11 Summary of Results

Stage	Hypothesis/ Research Questions	Expectation/Questions	Outcome
Risk Generation: Types of risks	H1	Compared to non-accountant practitioners, accountant practitioners **generate** a higher proportion of risks associated with **comparing the subject matter with suitable criteria** and a lower proportion of risks associated with **the measurement of the subject matter.**	Supported
Risk Generation: Interaction	H2a	Interacting teams generate a **lower quantity** of risks than nominal teams	Supported
	H2b	Interacting teams generate **fewer** categories of risk (**breadth**) than nominal teams	Marginally supported
	H2c	Interacting teams generate **fewer** risks within categories (**depth**) than nominal teams	Supported
Risk Generation: Review process	H3a	Review teams generate a **higher quantity** of risks than interacting teams	Supported
	H3b	Review teams generate a **higher quantity** of risks than nominal teams	Not supported (no difference)
	H3c	Review teams generate **more** categories of risk (**breadth**) than interacting teams	Supported
	H3d	Review teams generate **more** categories of risk (**breadth**) than nominal teams	Not supported (no difference)
	H3e	Review teams generate **more** risks within categories (**depth**) than interacting teams	Supported
	H3f	Review teams generate **fewer** risks within categories (**depth**) than nominal teams	Not supported (no difference)
Risk Selection: Types of risks	H4	Compared to non-accountant practitioners, accountant practitioners **select** a higher proportion of risks associated with **comparing the subject matter with suitable criteria** and a lower proportion of risks associated with **the measurement of the subject matter.**	Not supported (no difference)

4.6 Sensitivity Analysis

Sensitivity analysis was conducted in four ways to examine whether the results for Hypotheses 1 to 4 are robust. First, five participants who conducted the experiments via teleconferencing or videoconferencing were excluded. These five participants were two accountant practitioners in the review treatment and one accountant and two non-accountant practitioners in the nominal treatment. The analysis was re-run for all the dependent variables without these participants to see if any changes occurred in the results. Overall, the direction and statistical significance of all the results are largely the same as those presented in Sections 4.5.1 to 4.5.3, with the following exceptions. With the removal of these five participants, the difference between the accountant and non-accountant practitioners becomes marginally significant for the types of risks generated (H1: from $p = 0.024$ to $p = 0.066$, un-tabulated). Further, the marginally significant difference in the breadth of risks generated between the interacting and nominal teams is no longer significant (H2c: from $p = 0.060$ to $p = 0.106$, un-tabulated). These reductions in significance are likely due to the decrease in power from the reduced observations included in the analysis.

Second, as described in footnote 26 in Section 4.4.1, five participants had both accounting and engineering/science backgrounds (i.e., had double degrees or a master's degree in another area). Three of the five participants had financial audit experience and were classified as "accountant" practitioners. The other two practitioners had no financial audit experience and were classified as "non-accountant" practitioners. The sensitivity analysis results after excluding these five participants with mixed backgrounds remain unchanged for every dependent variable (un-tabulated).

Third, as mentioned earlier in footnote 28 in Section 4.4.1, one interacting team comprised two accountant practitioners (one with GHG experience and another with no GHG experience) and three review teams comprised accountant practitioners with an economics or commerce background but no financial audit experience. Sensitivity analysis was conducted by excluding these teams, and the statistical significance and direction of all results remain unchanged (un-tabulated).

Fourth, although all the participants were randomly allocated to each of the three treatments, significant differences are found for GHG assurance experience (in years) between the review and nominal teams (2.94 vs. 2.00, $z = 2.708$, $p = 0.009$, two-tailed, un-tabulated) and for the GHG assurance training in days between the review and interacting teams (7.10 vs. 2.75, $z = 1.967$, $p = 0.049$, two-tailed, un-tabulated). Because GHG assurance experience and training could potentially confound the results, additional analysis was conducted on each dependent variable to examine for experience effects. First, one participant in the review treatment have 50 days of GHG assurance training, which is significantly more than the rest of the participants. Excluding this participant from the analysis in sensitivity tests resulted in the difference in days of GHG assurance training between the nominal versus interacting teams becoming less (i.e., marginally) significant ($z = 1.689$, $p = 0.091$, two-tailed, un-tabulated). The results for all the hypotheses remain the same after excluding this participant. Second, univariate analysis was performed, and no significant correlations were found between each dependent variable and GHG assurance experience either for years of GHG assurance experience or days of GHG assurance training (all $p > 0.10$, un-tabulated). Third, analyses of covariance were used to control the results for the GHG assurance experience effect. The results for all the hypotheses comparing between the review and nominal teams are unchanged after controlling for years of GHG assurance experience (un-tabulated).

4.7 Additional Analysis

4.7.1 The Breadth and Depth of Risks Selected: Research Question 1

While the differences in the breadth and depth of risks *generated* between the three team formats are addressed in Hypotheses 2 and 3, additional analysis is conducted to obtain a preliminary understanding of the effect of team formats on the breadth and depth of risks *selected.* Research Question 1a compares the breadth of risks selected between treatments, and Research Question 1b compares the depth of risks selected between treatments.

To compare the breadth and depth of risks selected between the three conditions, the four risks that received the highest ranking within each group are included. For the interacting teams, the top four subcategories with rankings of 1, 2, 3, and 4 are included. For the nominal teams, items in the top four subcategories with rankings of 1 and 2 (from each team member's list) are included. For the review teams, the top four subcategories are derived by including those ranked 1, 2, 3, and 4 from the reviewer's lists.

To address RQ1a, the number of risk categories selected (breadth) is compared between interacting versus nominal, review versus interacting, and review versus nominal teams. As shown in Table 4.12, Panel A, interacting and review teams select risks with more breadth compared to the nominal teams (3.75 vs. 3.75 and 3.58, respectively). The comparisons in Panels B and C, however, indicate that no significant differences are found between the breadth of risks selected between the nominal, interacting, and review[55] teams (all $p > 0.390$, two-tailed). These results indicate that the breadth of risks selected by MDGHGTs is not affected by team format.

To address RQ1b, the number of the top four risks selected within a category (depth) is compared between the interacting and nominal teams, the review and interacting teams, and the review and nominal teams. Panel A of Table 4.12 provides the mean number of risks selected within a category for all treatments. Among the three treatments, the nominal teams select risks with the greatest depth (1.166), followed by the interacting (1.083) and review teams (1.083). Again, the comparisons between the nominal, interacting, and review teams in Panels B and C show no significant differences in terms of the depth of risks selected (all $p > 0.390$, two-tailed). The results indicate that team format has no effect on the depth of risks selected by MDGHGTs.

[55] Recall that in the later stage, reviewers were asked to re-rank their top four risks from a pooled list of risks (i.e., select risks generated by themselves and the non-accountant practitioners in their team). Thus, the top four risks selected from this pooled list with rankings of 1, 2, 3, and 4 were used to test the sensitivity of results for RQ1a. The analyses shown in Table 4.12 were re-run and yielded similar results. No differences are found in the breadth of risks selected between the review and interacting teams and the review and nominal teams (all $p > 0.400$).

TABLE 4.12 Comparison of Mean Number of the Top Four Risks of Potential Material Misstatements Selected by Teams in the Risk Selection Stage (Stage 2)

Panel A. Descriptive Statistics: Mean

	Mean Number of the Top Four Risks Selected	
Treatment	**Breadth of Risks**	**Depth of Risks**
Nominal (***n*** = 12)	3.58	1.166
Interacting (***n*** = 12)	3.75	1.083
Review (***n*** = 12)	3.75	1.083

Panel B. Independent-Samples Comparison (T-Test)

	t-statistics and p-values*	
Treatment	**Breadth of Risks**	**Depth of Risks**
Interacting vs. Nominal	0.715, p = 0.482	-0.859, p = 0.399
Review vs. Interacting	0.000, p = 1.000	0.000, p = 1.000

Panel C. Paired-Samples Comparison (T-Test)

	t-statistics and p-values*	
Treatment	**Breadth of Risks**	**Depth of Risks**
Review vs. Nominal	0.616, p = 0.551	-0.764, p = 0.461

Notes:
- This table presents the descriptive statistics (Panel A) and analysis of the breadth and depth of risks selected (i.e., top four risks) by the nominal, interacting, and review teams (Panels B and C).
- For the interacting teams, the top four subcategories with rankings of 1, 2, 3, and 4 were included. For the nominal teams, items from the top four subcategories with rankings of 1 and 2 (from each team member's list) were included. For the review teams, the top four subcategories the reviewers selected with rankings of 1, 2, 3, and 4 were included.
- Nominal: team in which no communication occurred between two team members.
- Interacting: team in which verbal communication occurred between two team members.
- Review: team in which no verbal communication occurred between two team members, but an accountant team member has a chance to see the risks generated by a non-accountant team member before adding to the list of risks and selecting their top four risks from their own list.
- *Breadth of risks* is the number of categories selected by each team.
- *Depth of risks* was calculated by dividing the *top four risks* selected by the *breadth* of risks.
- An independent-samples t-test is used to determine the statistical differences between the interacting and nominal teams and the review and interacting teams, while a paired-samples t-test in Panel B is used to determine the statistical differences between the review and nominal teams.

*All reported p-values are two-tailed.

4.7.2 Relationship between Interacting Teams and Elaboration on Task-Relevant Information: Research Question 2

Research Question 2 considers the correlation between the quantity, breadth, and depth of risks generated (RQ2a) and selected (RQ2b) by the interacting teams and the level of elaboration on task-relevant information by the members of the interacting teams. This could only be done for the interacting teams because the accountant and non-accountant practitioners in the nominal and review teams do not interact with each other. Elaboration on task-relevant information is measured using a three-item[56] self-reported measure on a seven-point scale. Higher numbers indicate more elaboration. This measure is based on the definition of information elaboration provided by van Knippenberg et al. (2004) and has been used in Homan et al. (2007, 2008).

The descriptive data analysis (un-tabulated) reveals that the average information elaboration score is 5.93 (ranging from 3.58 to 7.00). The average scores for the three items including the amount of information contributed by the other team member, unique information contributed by the other team member, and level of available information used by all members in the team are 5.96, 5.63, and 6.50, respectively.

Panel A of Table 4.13 provides the Pearson correlations for the elaboration scores and the quantity, breadth, and depth of risks generated. The table shows that the elaboration of task-relevant information is significantly correlated with the number of total risks generated ($r = 0.597$, $p = 0.041$) and marginally significantly correlated with the breadth of risks generated ($r = 0.513$ $p = 0.088$); however, it is not correlated with the depth of risks generated ($r = 0.309$, $p = 0.329$). No significant correlation is found for the breadth and depth of risks selected (all $p > 0.450$). These results suggest that the elaboration on task-relevant information can potentially improve interacting MDGHGTs' risk

[56] The three items are the extent to which the participants in the interacting teams agreed that their team member (1) contributed a lot of information during the team task and (2) contributed unique information during the group task and (3) that they themselves tried to use all the available information during the task. The mean scores for these items are 5.96, 5.63, and 6.50, respectively. Factor analysis was conducted on the three ratings and yielded one factor with an eigenvalue of 1.792. To combine the three scores into an elaboration score, the scores for each rating were multiplied by its factor loading (0.912, 0.881, and 0.430) and scaled by 1. The elaboration scores from the accountant and non-accountant team members were then averaged to form the elaboration score for each team.

generation performance both in terms of the quantity and the breadth of risks generated but not risk selection performance.

TABLE 4.13 Pearson Correlations: Elaboration on Task-Relevant Information, Cross-Understanding, Risks Generated, and Risks Selected in the Interacting Teams

Panel A. Elaboration	**Correlation Coefficient**				
	Risks Generated			**Risks Selected**	
	Total Risks	**Breadth of Risks**	**Depth of Risks**	**Breadth of Risks**	**Depth of Risks**
Elaboration	0.597**	0.513*	0.309	0.239	-0.239
Panel B. Cross-Understanding	**Correlation Coefficient**				
	Risks Generated			**Risks Selected**	
	Total Risks	**Breadth of Risks**	**Depth of Risks**	**Breadth of Risks**	**Depth of Risks**
Cross-Understanding (composite)	0.788***	0.531*	0.640**	0.310	-0.310
- Communication effectiveness	0.634**	0.348	0.643**	0.300	-0.300
- Knowledge elaboration	0.461	0.375	0.278	0.299	-0.299
- Collaboration	0.804***	0.559*	0.614**	0.070	0.070

Notes:
- Panel A of this table presents the correlation coefficient between the level of elaboration on task-relevant information and the number of total risks, breadth, and depth of risks generated as well as the breadth and depth of risks selected by the interacting teams. Panel B presents the correlation coefficient between the level of cross-understanding and the number of total risks, breadth, and depth of risks generated as well as the breadth and depth of risks selected by the interacting teams. *Total risks* are the number of subcategories generated by each team. The *breadth* of risks is the number of categories generated or the categories of the top four risks selected by each team out of the 12 categories. The *depth* of risks was calculated by dividing the total risks generated or the top four risks selected by the breadth of risks.
- *Elaboration* scores are the overall elaboration score calculated by averaging team member scores, with the higher numbers indicating more elaboration.
- *Cross-understanding* scores are the overall cross-understanding score calculated by averaging team member scores, with the higher numbers indicating more cross-understanding. This composite score comprises three components: communication effectiveness, knowledge elaboration, and collaboration.
- Interacting: team in which verbal communication occurred between the two team members.

, **, ** Significant at $p < 0.1$, $p < 0.05$, and $p < 0.01$, two-tailed, respectively.

4.7.3 Relationship between Interacting Teams and Cross-Understanding: Research Question 3

Research Question 3 considers the correlation between the quantity, breadth, and depth of the risks generated (RQ3a) and selected (RQ3b) by the interacting teams and the cross-understanding between the members in the interacting teams. Cross-understanding is constructed using three components (communication effectiveness, knowledge elaboration, and collaboration) and is measured using a 10-item[57] self-reported measure on a seven-point scale suggested by Huber and Lewis (2010). The higher number indicates higher cross-understanding.

The descriptive data analysis (un-tabulated) reveals that the average cross-understanding score is 5.01 (ranging from 3.52 to 5.82). The average scores for the three components including communication effectiveness, knowledge elaboration, and collaboration are 5.95, 4.24, and 5.13, respectively.

As shown in Panel B of Table 4.13, the composite scores of cross-understanding are significantly correlated with the number of total risks ($r = 0.788$, $p = 0.002$) and the depth of risks generated ($r = 0.640$, $p = 0.025$) and marginally significantly correlated

[57] Three items were used to measure communication effectiveness, four items were used to measure knowledge elaboration, and three items were used to measure collaboration. To measure communication effectiveness, participants in the interacting treatment were asked to rate the extent to which they thought their team member (1) chose concepts and words that they understand; (2) tailored communications to refer to concepts, terms, and perspectives they both had in common; and (3) made arguments that were technically, politically, or otherwise unacceptable to them. The mean scores for these items were 6.13, 5.83, and 2.08, respectively. To measure knowledge elaboration, the participants were asked to rate the extent to which they thought their team member (1) inquired about the reasons underlying their knowledge, beliefs, or preferences; (2) often asked for clarification or elaboration on issues related to their knowledge, beliefs, or preferences; (3) prompted them to surface and discuss what they knew, believed, or preferred; and (4) helped them better understand the team's task or task situation. The mean scores for these items were 3.71, 3.75, 4.75 and 5.33, respectively. To measure collaboration, they were asked to rate the extent to which they thought their team member (1) seemed to anticipate what they would do or say; (2) did a good job coordinating his/her actions with theirs; and (3) seemed to recognise when their and their team member's knowledge, beliefs, and preferences differed. The mean scores for these items were 4.71, 5.33, and 5.25, respectively. Factor analyses were conducted on these items and yielded one factor in each component, with eigenvalues of 1.663, 1.948, and 1.343. To combine the items under the same component, the individual scores of each rating were multiplied by its factor loading (scaled by 1). To combine the three components into a composite score (elaboration score*)*, a factor analysis was also conducted on these components and yielded one factor (eigenvalue = 1.614). The scores of each component were multiplied by its factor loadings (0.733, 0.684, and 0.780) and scaled by 1. The elaboration scores from the accountant and non-accountant team members were then averaged to form the elaboration score for each team.

with the breadth of risks generated (r = 0.531, p = 0.076). However, when these correlations are further broken down into three components, the results show that communication effectiveness and collaboration between members are significantly correlated with the quantity and depth of the risks generated (all p > 0.320). These results suggest that having more cross-understanding among MDGHGT members, particularly better communication and collaboration, could enhance the quantity and depth of risks generated by interacting MDGHGTs and thus may warrant further investigation.

4.7.4 The Ability and Expertise of Accountant and Non-Accountant Practitioners to Perform Risk Assessment: Research Question 4

Research Question 4 considers the differences between the accountant and non-accountant practitioners' ability and expertise to perform the risk assessment task, particularly whether accountant and non-accountant practitioners perceive themselves to be different from the other team member in terms of their ability to identify risks (RQ4a), knowledge of the subject matter (RQ4b), and knowledge of the relevant audit criteria and process (RQ4c). All participants were asked to assess their own ability to identify risks, knowledge of the subject matter, and knowledge of the audit criteria process on a seven-point scale (1 = extremely low; 7 = extremely high).[58] Only the participants in the interacting treatment and reviewers in the review treatment were asked to rate the ability and knowledge of the other team member; this was done immediately after assessing their own ability. The mean and median scores of the self-assessment on and the assessment of the other member's ability and knowledge are presented in Panel A of Table 4.14, and the paired-sample comparisons are provided in Panel B of Table 4.14.

[58] Regardless of the treatment group, the accountant practitioners rated themselves significantly higher on knowledge of audit criteria and process than did the non-accountant practitioners (z = -3.394, p = 0.000, two-tailed), while the non-accountant practitioners rated themselves significantly higher on knowledge of the subject matter than did the accountant practitioners (z = -1.967, p = 0.049, two-tailed). No significant difference is found for the self-rated ability to identify risks (z = -1.204, p = 0.228, two-tailed) between the accountant and non-accountant practitioners.

TABLE 4.14 Accountant and Non-Accountant Practitioners' Assessment of Self and the Other Team Member's Ability and Knowledge to Perform the Risk Assessment Task

Panel A. Descriptive Statistics: Mean and Median

		Assessment Scores on Self and the Other Member's Ability and Knowledge to Perform the Risk Assessment Task					
		Ability to Identify Risks		**Knowledge of Subject Matter**		**Knowledge of Audit Criteria and Process**	
Treatment		**Self**	**Other team member**	**Self**	**Other team member**	**Self**	**Other team member**
Interacting team:							
Accountant	Mean	5.42	5.25	4.17	5.75	6.25	4.75
	Median	5.00	5.00	4.00	6.00	6.00	5.00
Non-accountant	Mean	5.17	5.50	5.08	4.50	4.00	6.17
	Median	5.00	6.00	5.50	4.00	4.00	6.00
Review team:							
Accountant	Mean	5.33	5.00	4.83	5.08	5.75	5.00
	Median	5.00	5.00	5.00	5.00	6.00	5.00

Panel B. Paired-Samples Comparison (Wilcoxon Signed Ranks Test)

	z-statistics and p-values*		
	Ability to Identify Risks	**Knowledge of Subject Matter**	**Knowledge of Audit Criteria and Process**
Interacting team:			
Accountant	-0.539, p=0.590	-2.381, p=0.017	-2.630, p=0.009
Non-accountant	-1.100, p=0.271	-0.823, p=0.410	-3.114, p=0.002
Review team:			
Accountant	-0.545, p=0.586	-0.543, p=0.587	-1.638, p=0.101

Notes:
- This table presents descriptive statistics and analysis for the measure of the accountant and non-accountant practitioners' self-assessments and assessments of the other team member's ability and knowledge to perform risk assessment on GHG statements. Assessing the other team member's performance is only applicable to those who interact with their team member (accountant and non-accountant practitioners in the interacting team) or those who review their team member's work (accountant practitioners in the review team).
- The paired-samples comparison (Wilcoxon Signed Ranks Test) is conducted to test whether the accountant/non-accountant practitioners perceived themselves to be different from their team members in three aspects: ability to identify risks, knowledge of the subject matter, and knowledge of the audit criteria and process.
- The three dimensions are coded on a seven-point scale, with 1 indicating "extremely low" and 7 indicating "extremely high".

*Reported p-value is two-tailed.

Overall, participants in both the interacting and review treatments do not perceive themselves to be different from their team members in terms of their ability to identify risks (all $p > 0.270$, two-tailed). The main differences, however, exist in perceived knowledge. The accountant practitioners in the interacting teams perceive themselves to be significantly more knowledgeable about the audit criteria and process ($z = -2.630$, $p = .0009$, two-tailed) than non-accountant practitioners on their team. Similarly, the non-accountant practitioners on the interacting teams perceive themselves to have significantly less knowledge of the audit criteria and process than their accountant team members ($z = -3.114$, $p = 0.002$, two-tailed). The results also show that the accountant practitioners on the interacting teams think that their non-accountant team members are more knowledgeable about the subject matter than themselves ($z = -2.381$, $p = 0.017$, two-tailed), while non-accountant practitioners on the interacting teams do not perceive themselves to be different from their accountant team members in terms of knowledge of the subject matter ($z = -0.823$, $p = 0.410$, two-tailed). The accountant practitioners in the review treatment also generally perceive themselves as having more knowledge of the audit criteria and process and less knowledge of the subject matter, but these differences are not statistically significant (all $p > 0.10$).

4.7.5 The Risk Generation Performance of Accountant Practitioners in the Nominal versus Review Teams

Recall that the review and nominal teams shared the same non-accountant practitioners. The only difference between these two teams was that the accountant practitioners in the review teams saw the list of risks generated by the non-accountant team member before they generated their own list, while accountant practitioners in the nominal teams had to work individually without seeing the other team member's ideas. Instead of comparing the review and nominal teams' performances, the analysis was re-run by comparing the performance of accountant practitioners in the review and nominal treatments to see if the reviewers' performance is actually stimulated by seeing the reviewees' ideas. The un-tabulated results show that no significant difference is found for the breadth of risks generated between accountant practitioners in the nominal and review treatments (6.83 vs. 6.50, respectively). However, significant differences are found in the number of risks generated (11.75 vs. 9.00; $z = -2.179$, $p = 0.040$, two-tailed) and the depth of risks generated (1.71 vs. 1.39; $z = -2.364$, $p = 0.027$, two-tailed), which indicates that the accountant practitioners in the nominal treatment generated more risks in total and generated more risks within the subcategories (i.e., greater depth) than those in the review treatment.

TABLE 4.15 Summary of Results for Research Questions

Stage	Research Questions	Question	Answer
Risk Selection: Breadth and Depth of Risks	RQ1a	Does the type of team format affect the **breadth of risks selected** by teams comprising accountant and non-accountant practitioners?	No
	RQ1b	Does the type of team format affect the **depth of risks selected** by teams comprising accountant and non-accountant practitioners?	No
Risk Generation And Selection: Interacting Teams	RQ2a	Is the level of **elaboration of task-relevant information** correlated with the quantity, breadth, and depth of risks **generated** by interacting teams of accountant and non-accountant practitioners?	Yes (positively correlated with the quantity and breadth of risks generated)
	RQ2b	Is the level of **elaboration of task-relevant information** correlated with the quantity, breadth, and depth of risks **selected** by interacting teams of accountant and non-accountant practitioners?	No
	RQ3a	Is the level of **cross-understanding** correlated with the quantity, breadth, and depth of risks **generated** by interacting teams of accountant and non-accountant practitioners?	Yes (positively correlated with the quantity, breadth and the depth of risks generated)
	RQ3b	Is the level of **cross-understanding** correlated with the quantity, breadth, and depth of risks **selected** by interacting teams of accountant and non-accountant practitioners?	No
Perceived Ability and Knowledge: Interacting and Review Teams	RQ4a	Do accountant and non-accountant practitioners perceive themselves to be different from the other team member in terms of their **ability to identify risks**?	No
	RQ4b	Do accountant and non-accountant practitioners perceive themselves to be different from the other team member in terms of their **knowledge of the audit criteria and process**?	Yes
	RQ4c	Do accountant and non-accountant practitioners perceive themselves to be different from the other team member in terms of their **knowledge of the subject matter**?	Yes

4.8 Discussion and Limitations

This study compares the performance of nominal, interacting, and review teams to determine the effects of interacting and review processes on MDGHGTs' risk generation and selection performance. Three key questions are addressed in this study: (1) is there cognitive diversity in MDGHGTs working together in the planning stage to assess risks in the entity's GHG statements; (2) how do different team formats affect the ability of MDGHGTs to generate risks; and (3) how do different team formats affect MDGHGTs' utilisation of diverse information and perspectives. To address these questions, the types of risks generated and selected by accountant and non-accountant practitioners as well as the quantity, breadth, and depth of risks generated by teams are examined.

This study provides a number of important findings. The results show that accountant and non-accountant practitioners generate different types of risks, with accountant practitioners generating more risks associated with *comparing the subject matter with suitable criteria* and fewer risks associated with *the measurement of the subject matter*. These findings are significant because they provide evidence for the existence of cognitive diversity between assurers with different educational backgrounds, specifically accounting versus engineering/science. It demonstrates that MDGHGTs could potentially benefit from the complementary knowledge and perspectives generated by practitioners with accounting and scientific backgrounds. As such, this result supports the ISAE 3410 suggestion that GHG assurance engagements be performed by MDTs, particularly relatively complex engagements (IFAC 2012a, para. A42).

This study finds that nominal teams generate a significantly greater number of risks than interacting teams. In addition, nominal teams focus their risk generation more deeply within specific risk categories compared to interacting teams. While interacting teams may benefit from exchanging and integrating their diverse knowledge and perspectives, they are unlikely to have enough overlapping in their frames of reference and are likely to spend time connecting different pieces of information, thereby lacking enough time to explore these issues in depth (Dahlin et al. 2005). This is an example of production blocking referred to in both the psychology (Diehl and Stroebe 1991) and accounting (Chen et al. 2014) literature. ISAE 3410 requires MDGHGTs to interact through team discussion to perform risk assessment tasks. The results in this thesis show that there are potential disadvantages of this discussion as the interacting group is outperformed by both non-interacting team formats. However, the results suggest that interacting multidisciplinary teams could be used to perform tasks that are highly dependent on accessing a broad range of information, specifically when dealing with complex issues that lie across different disciplines.

Given that potential process losses in interacting MDGHGTs are likely to stem from the relatively high level of diversity in their educational backgrounds, this study explores two important constructs the social psychology literature suggests drive the positive effect of cognitive diversity on team performance: *information elaboration* (van Knippenberg et al. 2004) and *cross-understanding* (Huber and Lewis 2010). Because MDGHGTs are highly diverse in their educational backgrounds, they may find it difficult to understand each other's perspectives (Dunbar 1997; West 2000) or may have a number of disagreements on task-relevant issues (Jehn 1995). The need to explain and reconcile different knowledge bases and perspectives may force them to thoroughly elaborate on different or conflicting views (van Knippenberg et al. 2004). This notion is supported by the fact that the level of elaboration between accountant and non-accountant practitioners in the interacting teams is relatively high (5.9 of 7, on average). Although information elaboration is positively correlated with the quantity and breadth of risks generated by interacting teams, elaboration processes require time (van Knippenberg et al. 2004). Thus, the need to elaborate on different information and perspectives could potentially explain why interacting MDGHGTs generated a much lower number of risks than nominal teams. Because the team may not have enough time left to cover all the risk issues and explore a category in sufficient depth, the need to elaborate could also explain why interacting teams generate risks with a lower breadth and depth than nominal teams. Another potential explanation for why nominal teams outperform interacting teams in this setting is the level of cross-understanding between MDGHGT members. According to Huber and Lewis (2010), when the level of cross-understanding is high, it could enhance communication, elaboration effectiveness, and collaborative behaviours. The results of this study support Huber and Lewis' (2010) proposition by showing that cross-understanding between MDGHGTs in the interacting teams is positively correlated with the quantity and depth of risks generated. These findings indicate that information elaboration and cross-understanding between MDGHGT members are important underlying mechanisms for explaining differential outcomes between nominal and interacting teams. Future research on these issues is warranted.

The results show that review teams generate a similar number of risks and that the risks have a similar breadth and depth to those generated by the nominal teams. These findings are unexpected because the review teams have the opportunity to be stimulated by diverse perspectives while the nominal teams do not. In hierarchical audit teams, Trotman (1985) finds that review teams outperform nominal teams in terms of the accuracy of auditor judgments. In a hypotheses generation task, Ismail and Trotman (1995) use hierarchical audit teams and find that the review process, either with or without discussion between the reviewers and reviewees, increases the number of hypotheses generated. The present study uses multidisciplinary assurance teams and finds no differences between the review and nominal teams, while the review teams outperform interacting teams in the risk generation task. Unlike in Trotman (1985) and Ismail and Trotman (1995), the reviewers (accountant practitioners) and reviewees (non-accountant practitioners) in this study have much more diverse educational backgrounds. Thus, the reviewers in this study could find it difficult to build on the reviewees' ideas, particularly if they do not have enough overlap in mental representations (e.g., Dunbar 1997; West 2000). This difficulty could lead them to focus their risk generation more deeply within specific risk categories they are familiar with or are knowledgeable about. These problems are likely to be exacerbated by the fact that the review teams in this study are not given the opportunity to reconcile their different views through face-to-face discussion. The findings from more recent audit review studies (Brazel et al. 2004; Agoglia et al. 2009; Payne et al. 2010) support this view by showing that reviewers and reviewees face some difficulties when they are not allowed to have a face-to-face discussion of written review notes. Face-to-face review teams outperform electronic review teams in terms of the effectiveness of the workpaper documentation and the quality of the reviewees' judgments (Brazel et al. 2004; Agoglia et al. 2009). These studies suggest that face-to-face reviews allow reviewers and reviewees to communicate more effectively (Brazel et al. 2004) while increasing the reviewers' ability to recognise and mitigate lower quality workpapers (Agoglia et al. 2009) compared to when a face-to-face review is not allowed.

This study also shows that while accountant and non-accountant practitioners may generate different types of risks, they do not select different types of risks. Further, no differences are found in the breadth and depth of risks selected between the nominal, interacting, and review teams. However, the fact that all MDGHGTs in this study tend to select risks across a range of categories rather than focusing on a particular risk issue is promising. Because not all risks generated by practitioners can be addressed under limited resources, prioritising decisions to focus attention on a broad range of significant risks could increase the chance that material misstatements in different areas will be addressed.

This study contributes to the audit brainstorming literature by testing the effects of different team formats suggested by previous fraud brainstorming studies (Carpenter 2007; Hoffman and Zimbelman 2009; Lynch et al. 2009; Trotman et al. 2009; Chen et al. 2014) on MDGHGTs' performance. In hierarchical audit teams, Carpenter (2007) and Chen et al. (2014) find that nominal teams outperform interacting brainstorming teams in the number of risks generated. Carpenter et al. (2011) show that this is also the case in non-hierarchical audit teams. Literature in psychology (Nijstad and De Dreu 2002) suggests that previous brainstorming studies fail to find evidence that interacting teams outperform nominal teams because participants in most studies are homogeneous in their educational backgrounds and thus are less likely to generate different ideas (Nijstad and De Dreu 2002). This notion is in line with Trotman's (1985) and Libby et al.'s (1987) suggestion that interacting teams will outperform nominal teams if enough variation is present in team members' performance and if they are able to recognise the differences in their expertise. This could also explain the findings in audit brainstorming because while teams in previous audit brainstorming literature are different in their hierarchical natures, they are similar in their educational backgrounds (i.e., accounting/financial audit). To address this limitation, the present study tests the findings from audit brainstorming studies in a GHG assurance setting, in which practitioners from different disciplines (i.e., accounting, environmental science, and engineering) are required to work together to assess the risks of material misstatements. The results show that nominal teams also outperform interacting teams in the number of risks generated in the multidisciplinary assurance team setting.

These results shed new light on the suggestions in prior social psychology literature (Stroebe and Diehl 1994; Nijstad and De Dreu 2002; Paulus and Brown 2003) that the complementary knowledge members with diverse educational backgrounds bring to interacting teams will counteract the common process loss due to production blocking. However, the results suggest that the effect of multidisciplinarity on idea generation tasks is more complicated than expected. Stroebe and Diehl (1994) find no difference between nominal and interacting groups in terms of the quantity and breadth of ideas generated when participants have diverse knowledge structures (but not educational backgrounds).[59] The present study uses multidisciplinary practitioners and finds that nominal teams outperform interacting teams in the quantity, breadth, and depth of risks generated. The differences in the results between Stroebe and Diehl (1994) and the present study can potentially be explained by the differences in participants' educational backgrounds. Multidisciplinary team members bring not only a broad range of knowledge and perspectives to a given task but also different frames of reference, concepts, and professional languages (van Knippenberg et al. 2004). Thus, educational diversity can often be a source of misunderstandings, misinterpretations (van Someren et al. 1998; van Asselt 2000), and communication difficulties (e.g., Sheehan et al. 2007). These findings suggest that diversity among MDGHGT members has both a beneficial and a deleterious effect on the risk generation performance of interacting teams.

Moreover, this study examines important issues found by O'Dwyer (2011), which is the only prior study to examine multidisciplinary assurance teams in a sustainability assurance setting. While O'Dwyer (2011) finds that non-accountant assurers perceive accountant assurers to have insufficient subject matter knowledge to work on the task, the accountant and non-accountant practitioners in the present study perceived themselves to be different from each other in terms of knowledge but not in their ability

[59] Participants in Stroebe and Diehl (1994) are students with similar educational backgrounds who are different in their dominant associations regarding environmental concerns.

to identify risks. Thus, although they recognised the differences in their expertise, GHG assurance practitioners perceive their team member as having sufficient knowledge to be able to work together on the task, which is contrary to the findings reported previously in O'Dwyer (2011).

The results discussed above, however, should be considered in light of the study's limitations. First, the teams in this study comprise two members, one with an accounting background and one with a non-accounting background. GHG assurance teams in practice comprise, on average, five members, as reported in Study One. Thus, our study only captures part of the overall team interaction. Other team phenomena that cannot be tested in dyads, such as the effect of the informational minority or majority (e.g., Lau and Murnighan 1998, 2005) on MDGHGTs' performance, is worth investigating in the future.

Second, only accountant practitioners (reviewers) in the review teams see the ideas of the non-accountant practitioners (reviewees). It is unknown whether the results would change if the non-accountant practitioners were to review the accountant practitioners' ideas. Since these two sets of practitioners have different knowledge and mindsets, future research is needed to investigate how MDGHGTs' performance will be affected by having non-accountant practitioners review the work of accountant practitioners. Further, no feedback is provided and no discussion occurs between the reviewers and reviewees in this study. Future research can manipulate the level of discussion (review with and without discussion, e.g., Ismail and Trotman 1995) or the review format (electronic review vs. face-to-face review, e.g., Brazel et al. 2004; Agoglia et al. 2009; Payne et al. 2010) to investigate the impact of discussion on the performance of multidisciplinary review teams.

Third, as this study aims to understand how MDGHGTs utilise their diverse information and perspectives, the aspects of quality examined are related to the breadth and depth of the risks generated and selected. To completely understand the effect of team format on the performance of MDGHGTs, future research can use other criteria to measure the quality of risks generated and selected by MDGHGTs, such as quality measured by a comparison with a group of experts (Trotman et al. 2009; Hammersley et al. 2011; Chen et al. 2014).

Finally, the study is limited by the small number of practitioners currently working in the field of GHG assurance in Australia, specifically in the Big Four audit firms. Accountant practitioners in this study need to have not only financial audit experience/training but also GHG assurance experience/training to perform the tasks. The number of traditional financial auditors who have migrated to providing assurance on GHG statements is currently still relatively small; therefore, the number of eligible participants was small. This study obtained the involvement of a large percentage of those practitioners qualified to complete the task. As this area of assurance expands, the conditions under which different team formats work better than others can be considered in future research.

CHAPTER 5

SUMMARY AND CONCLUSION

5.1 Introduction

This chapter summarises the aims of this dissertation and the results reported from the two studies undertaken to achieve these aims. The implications for practice and potential for future research are also discussed. The motivation for the dissertation and major findings from the two studies are presented in Section 5.2, and contributions and implications of the dissertation are discussed in Section 5.3. Finally, Section 5.4 considers limitations of the two studies and highlights opportunities for future research.

5.2 Motivation and Research Findings

The increasing attention to climate change in recent years had led to an escalating demand for corporations to report non-financial corporate responsibility (CR) information, including greenhouse gas (GHG) emissions (KPMG 2008, 2013; Simnett et al. 2009b; Cohen et al. 2012). To ensure the credibility of this information, ISAE 3410 ("Assurance engagement on greenhouse gas statements") (IFAC 2012a) was issued to provide comprehensive guidance for practitioners undertaking GHG assurance engagements. Given the high level of subject matter expertise required for these types of engagements, assurance on GHG statements is provided by practitioners from various backgrounds, including accounting, engineering, and environmental science (Nugent 2008). Further, ISAE 3410 requires that these engagements are undertaken by multidisciplinary greenhouse gas assurance teams (MDGHGTs) comprising both accountant and non-accountant practitioners and indicates that these teams are to discuss, in the planning stage, the susceptibility of the entity to material misstatements in GHG statements due to fraud or error (IFAC 2012a, para. 29). While the interactions among MDGHGT members are expected to improve the assurance quality because of the necessity of integrating both accounting and scientific competencies to complete GHG assurance engagements (Huggins et al. 2011), the literature on social psychology shows that educational diversity in work teams can either benefit or hinder team performance (van Knippenberg and Schippers 2007). These results have implications for accounting and assurance services firms as they consider the processes to apply to

GHG assurance teams, which are much more multidisciplinary in nature than traditional audit or assurance teams. In particular, the effectiveness of the GHG assurance teams depends on the functioning of the teams, which in turn is likely reflected by both the format of these teams and the processes employed.

The first motivation for this dissertation is that very little known about current accounting and assurance services firm practices employed around GHG emissions assurances. GHG assurance is a new, rapidly emerging assurance area and to date, information on how assurance firms operationalise the multidisciplinary teams engaged to conduct this assurance remains internal to audit firms. Current practices also remain undocumented, such as team composition and the processes involved in addressing collective competence and capabilities associated with having a multidisciplinary team.

The second motivation for this dissertation is to explore a number of team formats that may be utilised to improve current GHG assurance practices so MDGHGT effectiveness is increased and audit outcomes are enhanced. Despite extensive theories in the social psychology literature on strategies to enhance multidisciplinary group performance (see Kerr and Tindale 2004 and van Knippenberg and Schippers 2007 for reviews), auditing research addressing the effects of such strategies on multidisciplinary team performance is currently absent. As such, this dissertation first identifies various factors associated with MDGHGTs' effectiveness and then examines whether three different team formats (nominal, interacting, and review teams) identified from the social psychology (Stroebe and Diehl 1994; Dennis et al. 1999; Rietzschel et al. 2006) and audit literature (Trotman 1985; Ismail and Trotman 1995; Carpenter 2007; Chen et al. 2014) have different effects on MDGHGTs' performance. In this dissertation, the measure of performance employed is the generation and selection of risks of material misstatements due to fraud and error.

Study One utilised a retrospective recall experiential questionnaire (Gibbins et al. 2001; Gibbins and Trotman 2002) to ascertain which of a number of various factors could be associated with the of MDGHGTs' effectiveness. Specifically, participants were asked to recall a GHG engagement they worked on that thought was effective and a second engagement they thought did not operate as effectively. This study employed an Input-Process-Output Framework as suggested by team effectiveness frameworks in psychology (Cohen and Bailey 1997; Kozlowski and Ilgen 2006; Mathieu et al. 2008) (Figure 5.1).

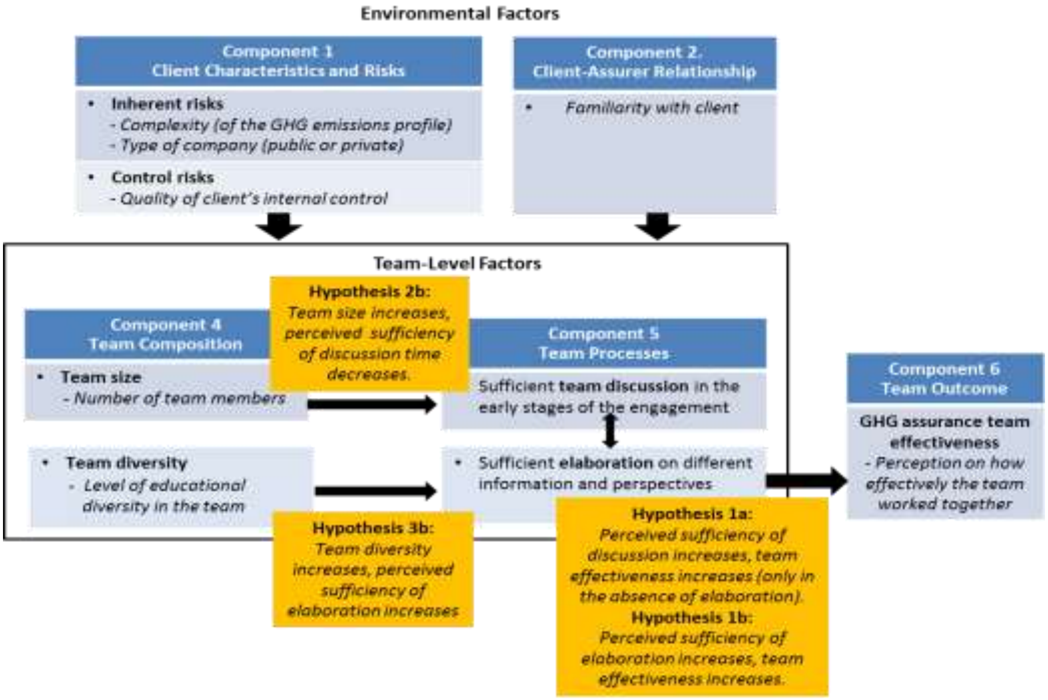


FIGURE 5.1 Summary of Findings from Study One

As shown in Figure 5.1, team processes (including discussion and elaboration) are crucial to the success of MDGHGTs. Specifically, having sufficient elaboration between team members on different information and perspectives significantly increases perceived MDGHGT effectiveness. While having sufficient discussion time in the early stages of engagement is also positively related to MDGHGTs' effectiveness, this relationship becomes insignificant in the presence of elaboration, thus suggesting that discussion is one component of information elaboration[60]. In addition to indicating the importance of team processes, the findings demonstrate that team composition (including team size and diversity) is also an important factor that influences the effectiveness of MDGHGTs. The most salient feature of MDGHGTs—level of educational diversity—enhances team effectiveness through the perceived sufficiency of elaboration on different information and perspectives. When the level of diversity in the team increases, accountant and non-accountant practitioners in the team are perceived as elaborating more on task-relevant information, which increases the effectiveness of the

[60] Elaboration is defined as "the exchange of information and perspectives, individual-level processing of information and perspectives, the process of feeding back the results of this individual-level processing into the group, and discussion and integration of its implications" (van Knippenberg et al. 2004, p. 1011).

MDGHGTs. The number of members in the team is not directly associated with perceived MDGHGT effectiveness; however, the sufficiency of discussion time in the early stages of engagement is perceived to decrease as the size of the team increases.

The findings from Study One also reveal that environmental factors (e.g., the complexity of the GHG emissions profile, whether the client is a public company or not, the quality of the client's internal control, and familiarity with the client) either have a direct perceived impact on MDGHGTs' effectiveness or act indirectly by influencing team composition and team processes. However, this study does not find a significant relationship between MDGHGTs' effectiveness and a number of other environmental factors (e.g., number of facilities and client importance) or task characteristics (e.g., the level of task interdependence between the accountant and non-accountant assurer and the task type). Therefore, these variables are not shown in Figure 5.1.

The relationships between educational diversity, team processes, and team effectiveness found in Study One highlight the important role of team processes (e.g., sufficient discussion time and elaboration) play in the success of MDGHGTs. However, increasing MDGHGT effectiveness through team processes is not necessarily a straightforward task. On one hand, team members' diverse educational backgrounds allow a large pool of knowledge and expertise to be shared and integrated, which thus enhances team creativity and decision making (Guzzo and Dickson 1996; Williams and O'Reilly 1998). On the other hand, individuals in MDTs may have different frames of reference, professional language, and problem-solving styles that impede optimum sharing and the integration of diverse ideas and information (van Knippenberg and Schippers 2007). In fact, the only empirical study to date exploring team effectiveness in the sustainability assurance setting provides evidence of such difficulties (O'Dwyer 2011). These results suggest that to improve the performance of MDGHGTs, team processes that can facilitate the use of the diverse information and perspectives from team members, while at the same time minimising the communication problems caused by educational diversity are needed.

Through a controlled experiment, Study Two extends the findings from Study One by investigating team processes underlying the differential effects of three different team formats: nominal, interacting, and review teams. This study also addresses a limitation in Study One (i.e., the use of recalled and perceived effectiveness) by using a more objective outcome to measure the effectiveness of MDGHGTs. The experiment employs two risk assessment tasks: risk generation and risk selection. Three main research questions are addressed in this study: (1) is there cognitive diversity in MDGHGTs working together on risk generation and selection tasks; (2) how do different team formats affect their ability to generate risks; and (3) how do different team formats affect the utilisation of diverse information and perspectives. To address the first question, the type of risks generated and the type of risks included in the four key risks selected are used to measure the cognitive diversity between accountant and non-accountant practitioners. To address the second question, the quantity of risks generated is used to measure the teams' ability to generate risks (Carpenter 2007; Trotman 2009; Chen et al. 2014). To address the third question, the number of different categories of risks generated (i.e., breadth) and the number of risks generated within a specified risk category (i.e., depth) (Asare 2000; Green and Trotman 2003; Dahlin et al. 2005; Kohn and Smith 2011) are used. The summary of findings from Study Two is illustrated in Figure 5.2.

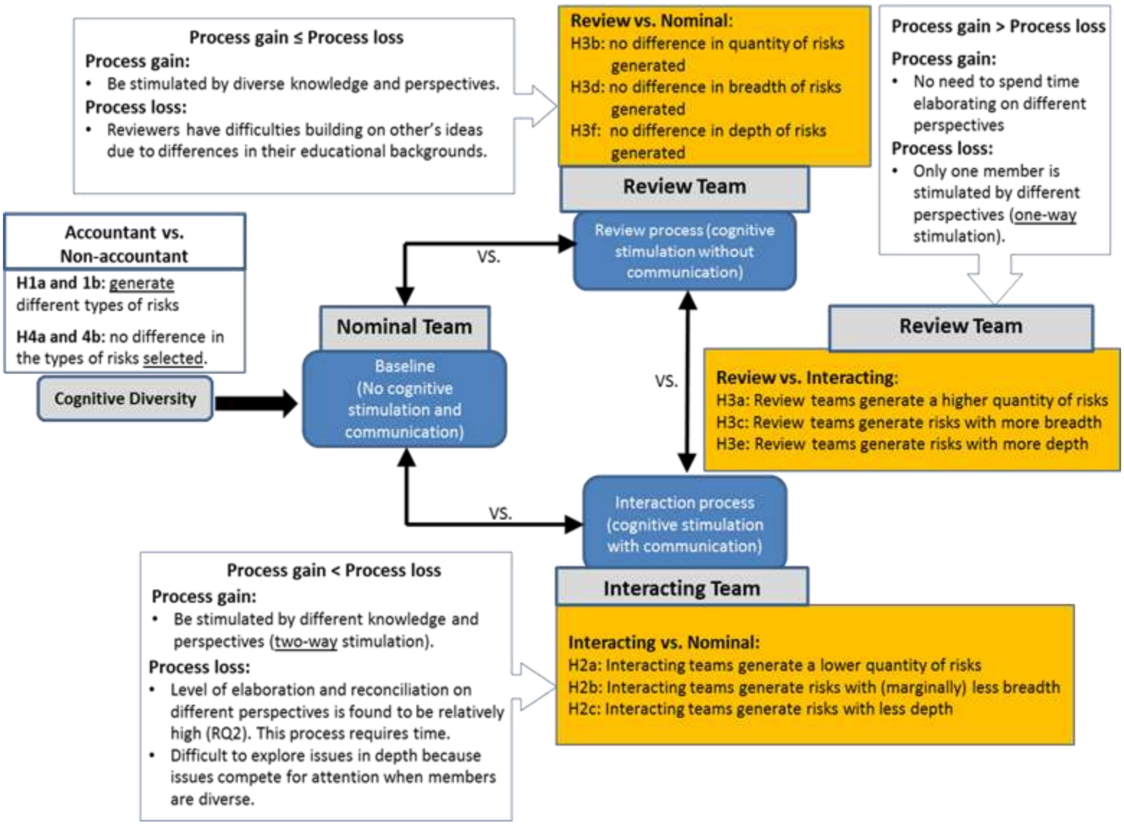


FIGURE 5.2 Summary of Findings from Study Two

Study Two provides evidence that the existence of cognitive diversity between accountant and non-accountant practitioners in MDGHGTs leads to different risk generation characteristics. Specifically, accountant practitioners generate more risks associated with comparing the subject matter with suitable criteria and fewer risks associated with the measurement of the subject matter. These findings demonstrate that accountant and non-accountant practitioners can complement each other by working together on GHG assurance engagements, specifically on a risk generation task.

The effect of the three different types of team formats on the quantity, breadth, and depth of risks generated are addressed. First, interacting and nominal teams are compared using the nominal team, which simply combines outputs from accountants and non-accountants, as the benchmark to understand the consequences of MDGHGT members engaging in interaction. The key difference between these two team formats is that interacting teams have the opportunity to be stimulated by different perspectives through team discussions while nominal teams do not. Compared to nominal teams, the interacting teams not only generated a lower quantity of risks but also generated a smaller range of risk categories (i.e., less breadth) with fewer risks generated within each category (i.e., less depth). The differences in risk generation performance between nominal and interacting teams can potentially be explained by the high level of information elaboration in the interacting teams. Given that accountant and non-accountant practitioners have very different educational backgrounds (accounting vs. science or engineering), they are likely to have conflicting views on task-relevant issues, which forces them to thoroughly process different information (van Knippenberg et al. 2004). Although information elaboration between interacting MDGHGTs is positively correlated with quantity and marginally correlated with the breadth of risks generated by interacting teams, it is also a time-consuming process. This could explain why interacting teams in this study are not as productive in risk generation as nominal teams. Further, elaboration on different information and perspectives could lead interacting teams to think of a wide range of risk issues but may not leave enough time for interacting teams to explore those issues in depth (Dahlin et al. 2005). This is an example of production blocking referred to in both the psychology (Diehl and Stroebe 1991) and accounting (Chen et al. 2014) literatures. While the results show the potential disadvantages of MDGHGT discussions during risk generation tasks, interacting teams could be useful when a broad range of issues needs to be considered. Future research that manipulates the amount of time available for interaction would be useful to further address these issues.

Second, the review and nominal teams are compared to examine the effect of cognitive stimulation through the review process. The key difference between these two teams is that the nominal team does not have the opportunity to be stimulated by their contributing team members while the accountant practitioners in the review teams do have the opportunity to see the risks generated by their non-accountant team members. By having the chance to build on the ideas of others, review teams are expected to outperform nominal teams. However, no differences are found in the quantity, breadth, and depth of risks generated between these two team formats. One possible reason for this result is the differences in their knowledge bases, which could make it difficult for accountant practitioners to understand the reasons behind non-accountant practitioners' ideas (van Someren et al. 1998; van Asselt 2000), particularly when the accountant reviewers have to build on those ideas without having the chance to clarify and discuss them with the non-accountant reviewees (Brazel et al. 2004; Agoglia et al. 2009; Payne et al. 2010). This difficulty could lead the accountant reviewers to focus their risk generation deeper than expected within specific risk categories they are familiar with or are knowledgeable about. Thus, MDGHGTs may not fully benefit from the cognitive stimulation offered by the review process unless it occurs with discussion.

Third, the review and interacting teams are compared to understand the consequences of being stimulated by seeing others' ideas versus being stimulated by discussing the ideas with others. Unlike nominal teams, both the review and interacting teams have the opportunity to be exposed to diverse ideas and perspectives. However, the difference between the review and interacting teams is that while the former is stimulated by another team members' perspective, the team does not have the chance to discuss these perspectives; the latter not only has the stimulation of generating together but also the chance to discuss each other's perspectives. The results show that review teams generate a greater number of risks than interacting teams. These risks are also greater in breadth and have a greater depth than the risks generated by interacting teams. These findings are likely because of process losses in the interacting teams rather than cognitive stimulation in the review teams, given that review teams do not outperform nominal teams.

The results also show that while accountants and non-accountants generate different types of risks, they do not select different types of risks when asked to nominate their four key risks. Moreover, no difference is found in the breadth and depth of risks selected between nominal, interacting, and review teams. Further, this study finds that MDGHGTs in all treatments select their top four risks across a broad range of categories. This is important because it could increase the chance that risks in different areas will be addressed.

5.3 Key Contributions

The results of this dissertation have several important contributions. First, Study One identifies team composition (educational diversity) and team process (information elaboration) variables as important factors in enhancing the perception of MDGHGT effectiveness. This finding is consistent with the team effectiveness frameworks suggested by previous studies in social psychology (e.g., Cohen and Bailey 1997; Ilgen et al. 2005; Mathieu et al. 2008). Given the very limited research on assurance teams composed of practitioners from different disciplines, this dissertation makes an important contribution to auditing research and assurance practice by providing insights into the factors underlying the success of MDTs in the GHG assurance context, in which practitioners from accounting backgrounds work with practitioners from scientific backgrounds.

Second, the findings from Study One add to the existing audit literature and team effectiveness frameworks by identifying environmental variables related to the GHG assurance setting. Since team effectiveness depends heavily on context, it is important to understand team effectiveness in different disciplines. The prior audit literature finds that factors such as client inherent and control risks (i.e., complexity, whether the client is a public company, and the quality of the client's internal control) and the client assurer–relationship (familiarity with the client and client importance) are associated with audit effectiveness (DeAngelo 1981; O'Keefe et al. 1994; Hackenbrack and

Knechel 1997; Reynolds and Francis 2001; Gibbins and Trotman 2002; Knechel et al. 2009). Consistent with the prior literature, the findings from this thesis indicate that the quality of the client's internal control and familiarity with the client are associated with team effectiveness. Further, this study provides new evidence that the level of educational diversity in MDGHGTs differs depending on factors such as complexity of the emissions profile and whether the client is a public company.

Third, the results of Study One show how two different strategies, which the social psychology literature suggests optimise the performance of diverse teams (i.e., extending the discussion time in the start-up phase and information elaboration), affect MDGHGT effectiveness. Larson et al. (1994) suggest that extending discussion times for informational diverse groups in the start-up phase provides a greater opportunity for diverse information to be shared and used, while previous studies on work team diversity (van Knippenberg et al. 2004; van Ginkel and van Knippenberg 2008) suggest that information elaboration improves the performance of cognitively diverse teams. Consistent with the previous literature, Study One reveals that MDGHGTs work more effectively together when they perceive that they have sufficient discussion time in the early stages of engagement and when they perceive that sufficient elaboration on different information and perspectives occurred. However, the fact that sufficient discussion time is not significant in the presence of sufficient elaboration indicates that discussion is one component of information elaboration. This finding fits the definition of information elaboration provided by van Knippenberg et al. (2004): "the exchange of information and perspectives, the process of feeding back the results of this individual-level processing into the group, and discussion and integration of its processes" (p.1011). This study also shows that the diversity level is positively related to MDGHGTs' effectiveness but only in the absence of information elaboration. This finding lends some support to van Knippenberg et al.'s (2004) proposition that information elaboration drives the positive effect of diversity on team performance.

Fourth, Study Two contributes to the limited empirical evidence on multidisciplinary assurance teams by examining how the cognitive diversity between accountant and non-accountant practitioners affects assurance team performance in the GHG assurance setting. Using interview data, O'Dwyer (2011) shows that tensions are found between accountant and non-accountant practitioners in the sustainability assurance setting because of their different concepts of materiality and their different mindsets on how to approach the data. While O'Dwyer (2011) demonstrates the deleterious effect of cognitive diversity, Study Two adds to the existing evidence by demonstrating how MDGHGTs' performance can benefit from cognitive diversity. In the first instance, cognitive diversity is measured by directly comparing the types of risks accountant and non-accountant practitioners generate and select. The results show that accountant practitioners generate more risks associated with comparing subject matter with suitable criteria and fewer risks associated with measuring the subject matter compared to non-accountant practitioners. These findings are important because they suggest the need for accountant and non-accountant practitioners to complement each other when working together on GHG assurance engagements.

Fifth, Study Two contributes to the audit brainstorming literature by examining how cognitive diversity affects team performance under different team formats suggested by previous fraud brainstorming studies (Carpenter 2007; Hoffman and Zimbelman 2009; Lynch et al. 2009; Trotman et al. 2009; Carpenter et al. 2011; Chen et al. 2014). Because the prior studies mainly focus on teams composed of auditors who typically have accounting backgrounds, this study contributes to the literature on audit brainstorming by providing the first empirical evidence on the effect of interaction on MDGHGT performance. The prior literature suggests that interacting teams in previous brainstorming studies do not outperform nominal teams because of lack of diversity in their participants' educational backgrounds (Nijstad and De Dreu 2002), which then leads to lack of variation in their performance and ability to differentiate relative expertise (Trotman 1985; Libby et al. 1987). This dissertation addresses this limitation by testing the findings from previous audit brainstorming studies in a GHG assurance setting, in which practitioners from different disciplines (i.e., accounting, environmental science, and engineering) are required to work together to assess the risks of material misstatements. Similar to the hierarchical audit teams (Carpenter 2007; Chen et al. 2014) and non-hierarchical audit teams (Lynch et al. 2009; Carpenter et al. 2011), interacting teams in this study do not outperform nominal teams on the quantity of risks generated.

Sixth, Study Two also contributes to the audit review process literature by investigating the effect of the review process on MDGHGTs' performance. While accountant practitioners are increasingly required to review non-accountant practitioners' work (Griffith 2014), auditing research has not examined the effect of the review process on multidisciplinary assurance teams. In hierarchical financial audit teams in which reviewers and reviewees have similar educational backgrounds, the review process has been found to improve the accuracy of audit judgment (Trotman 1985) and the number of hypotheses generated (Ismail and Trotman 1995). However, inconsistent with previous audit review process studies, this study finds that MDGHGTs do not benefit from having accountant practitioners review non-accountant practitioners' work because they generate a similar number and breadth of risks as the nominal teams. One possibility for the lack of cognitive stimulation suggested by this result is that differences in the knowledge base between reviewers and reviewees make it difficult for reviewers to build on ideas from non-accountant practitioners in their team because they do not fully understand these ideas. The possibility that an interacting review process will overcome this is left for future research.

Seventh, Study Two examines two aspects of quality—the breadth and depth of risks generated and selected—to gain insights on how MDGHGTs in different team formats utilise their diverse knowledge and perspectives. This is important because increasing the breadth of risks generated/selected increases the chance that material misstatements in different areas will be detected/addressed, while increasing the depth of risks generated/selected increases the chance that important issues will be explored more completely. Dahlin et al. (2005) find that educationally diverse teams are more likely to generate ideas over a range of areas and explore these ideas in greater depth than homogeneous teams. They suggest that educationally diverse teams can quickly process a wide range of both familiar and unfamiliar issues because they have access to a larger pool of cognitive resources. Thus, the teams have more time to explore these issues in depth. Study Two in this dissertation demonstrates that this is not always the case, particularly when educationally diverse teams interact with each other. This study suggests that the need to spend time elaborating on different perspectives could reduce the breadth of risks generated. Further, this reduces the depth of risks generated because various issues compete for attention during MDGHGT discussions.

Eighth, Study Two contributes to the social psychology literature by exploring the relationship between information elaboration and MDGHGTs' risk generation performance. Previous studies on information elaboration use students with similar educational backgrounds to examine the effect of information elaboration exclusively on decision-making tasks (van Ginkel and van Knippenberg 2009; van Ginkel et al. 2009). Study Two extends these studies by showing that information elaboration can both positively and negatively affect idea generation. On one hand, Study Two finds an increase in the quantity and breadth of risks generated because of the need to thoroughly exchange, discuss, and integrate all relevant information and perspectives. On the other hand, the elaboration process requires time (van Knippenberg et al. 2004); therefore, under the tight audit time constraints, information elaboration is likely to decrease the quantity and breadth of the risks generated. In addition, Study Two explores the relationship between cross-understanding and MDGHGTs' risk generation performance. Huber and Lewis (2010) propose that a high level of cross-understanding between members could cancel out the negative effects of cognitive diversity found in MDTs by increasing communication, elaboration, and collaboration effectiveness. However, very limited research has examined this proposition either in teams comprising members from different disciplines or in an idea-generation task. This dissertation finds that cross-understanding is positively correlated with the quantity and depth of risks generated; in doing so, the dissertation contributes to the social psychology literature by testing the findings from the social psychology literature in the context of GHG assurance.

5.4 Implications for Assurance Practice

This dissertation has several implications for assurance practice. First, Study One provides evidence on various factors that could increase MDGHGTs' effectiveness. In particular, this study pays attention to the factors that are under the control of the assurance firms, including team composition and team process variables. The results suggest that MDGHGTs work more effectively together when they perceive they have sufficient elaboration on different information and perspectives. Further, the educational diversity level affects the perception of elaboration sufficiency in MDGHGTs. While team size does not affect team effectiveness, MDGHGTs are less likely to perceive they have sufficient discussion to share their information and perspectives in the early stages of engagement as the size of the team increases. These findings will be of interest to assurance firms that currently provide GHG assurance services. By encouraging MDGHGTs to thoroughly elaborate on their different knowledge and perspectives, MDGHGTs should be able to improve their assurance quality. In addition, assurance firms should be aware of the diversity level in the team and the team size when composing teams. Having too many accountant practitioners or non-accountant practitioners on one team or too many team members overall can decrease the level of elaboration and discussion in the team.

Second, ISAE 3410 requires MDGHGTs to discuss the susceptibility of an entity's GHG statement to material misstatements at the planning stage (IFAC 2012a, para. 29). However, the findings from Study Two suggest that interacting MDGHGTs are less able to generate a high number of risks or deeply explore a risk category compared to nominal teams. Decreasing the quantity of risks generated will decrease the chance that more quality risks are generated, including the primary risk (Osborn 1953; Hammersley et al. 2011). Further, decreasing the breadth of risks generated reduces the chance that material misstatements in different areas are detected, whereas decreasing the depth of risks generated reduces the chance that important risk issues are completely explored (Dahlin 2005). Study Two explores the possibility of improving MDGHGTs' performance through interaction and finds that the quantity and the depth of risks generated can be increased through a higher level of cross-understanding. Cross-understanding refers to "the extent to which group members have an accurate understanding of one another's mental models" (Huber and Lewis 2010, p. 7). By understanding "what others know, believe, are sensitive to, and prefer" (Huber and Lewis 2010, p. 9), MDGHGTs can communicate, elaborate, and collaborate more effectively. Cross-understanding can evolve through prior interactions and communications with each other, from observations of members' communications or behaviours, and from access to members' biographical information (Huber and Lewis 2010). Therefore, MDGHGTs' risk generation performance could be enhanced through several methods, including informing members about others' expertise and encouraging members to meet each other before their task begins.

Third, ISAE 3410 suggests that GHG assurance engagements may be performed by a multidisciplinary team, particularly when the engagement is relatively complex (IFAC 2012a, para. A42). The findings from both studies in this dissertation provide evidence to support the use of multidisciplinary teams in the GHG assurance setting. The findings from Study One show a positive relationship between the level of educational diversity and the perceived effectiveness of MDGHGTs, while the findings from Study Two demonstrate how accountant and non-accountant practitioners can complement each other when generating risks of material misstatements. However, the additional analyses in Study One show that MDGHGTs experience some difficulties working together because the roles and responsibilities between accountant and non-accountant practitioners in the team are unclear. Therefore, assurance firms should clearly define the roles and responsibilities of all team members to provide overall direction and to enhance the functioning of MDGHGTs.

5.5 Limitations and Future Research Directions

The two studies in this dissertation suggest several opportunities for future research. Team tasks and team processes are dynamic because they can change over time (McGrath 1991). In particular, team processes develop as team members gain more experience working together (Kozlowski and Ilgen 2006). Given the newness of the GHG assurance service, MDGHGTs are currently in the early stages of development. As such, members of these teams are in the process of learning how to deal more effectively with other members from different disciplines. This provides an opportunity for future research to examine the effect on MDGHGT effectiveness of team tenure and over time stable process constructs, such as shared mental models (Mathieu et al. 2000). Van Ginkel and van Knippenberg (2008) investigate informational diverse groups' performance and find that diverse groups make better decisions when they share mental models emphasising elaboration, compared to when they hold such a shared metal model to a lesser extent. Future research could examine whether this holds for MDGHGTs, where the team members are diverse in their educational background rather than their information.

Study One does not find the expected relationship between task characteristic variables and team effectiveness. The first variable (task interdependence) is measured in that study by the proportion of direct measurement used by clients as a measure of complexity requiring subject-matter expertise. Future research could measure task interdependence using self-assessment measures (e.g., Stewart and Barrick 2000; Van der Vegt et al. 2000). The second variable (task type) is categorised into two types: reasonable and limited assurance engagements. Although these two engagements differ in the level of assurance and are expected to differ in the levels of interaction and coordination required between accountant and non-accountant practitioners, there may not be enough differences between them to explain the variance in the effectiveness of MDGHGTs. Future research can categorise GHG assurance tasks in different ways. For example, team tasks could be categorised into eight different tasks using McGrath's (1984) Group Task Circumplex, which includes planning tasks, creativity tasks, intellectual tasks, decision-making tasks, cognitive conflict tasks, mixed-motive tasks, competitive tasks, and psycho-motor tasks. Further, the five task characteristics (task variety, task identity, task significance, autonomy, and feedback) suggested by Hackman and Oldham (1975) could also be explored.

The additional analyses in Study One reveal two additional factors that MDGHGT members perceived as contributing to team effectiveness: (1) sufficient communication with the client and (2) team members with sufficient technical skills and experience. The results also show two major factors perceived as inhibiting team effectiveness: (1) unclear roles and responsibilities of accountant/non-accountant practitioners and (2) lack of time to prepare and complete work. These findings provide an opportunity for future research to investigate these variables' effect on MDGHGT effectiveness.

Future research can also examine other dimensions of diversity in MDGHGTs using large teams. The use of larger teams (more than two members) allows the testing of the informational minority/majority's (Lau and Murnighan 1998, 2005) effect on the performance of MDGHGTs. Homan et al. (2007) show that the level of information elaboration is higher in informationally diverse groups than in informationally homogeneous groups. However, this result does not hold when informationally diverse groups experience a strong subgroup categorisation (i.e., male members hold information A and female members hold information B). Future research could respond to the call for more research investigating the effect of other kinds of diversity on diverse teams (i.e., Van Knippenberg and Schippers 2007).

In Study Two, accountant practitioners were asked to review the risks generated by non-accountant practitioners. Because this study provides evidence for the existence of cognitive diversity between these practitioners, future research could investigate how MDGHGT performance will be affected by having non-accountant practitioners review accountant practitioners' work. In addition, other review formats, such as face-to-face review/review with discussion, could be considered (Ismail and Trotman 1995; Brazel et al. 2004; Agoglia et al. 2009; Payne et al. 2010). Future research can manipulate the level of discussion in the review teams to examine when review with discussion works better than review without discussion in a multidisciplinary team setting.

Study Two in this dissertation aims to understand how MDGHGTs utilise their diverse information and perspectives. Thus, only two aspects of quality are examined: the breadth and depth of risks generated and selected. Future research can adopt other quality criteria to compare MDGHGTs' risk assessment performance in different team formats, e.g., quality as measured by the comparison with a group of experts (Trotman et al. 2009; Hammersley et al. 2011; Chen et al. 2014). This will help obtain a more complete understanding of team format's effect on the performance of MDGHGTs.

Finally, both studies in this dissertation are limited by the small number of practitioners currently working in the field of GHG assurance in Australia, specifically in the Big Four audit firms. To answer the retrospective research instrument in Study One, all participants had to have undertaken at least two GHG assurance engagements. To perform the tasks in Study Two, accountant and non-accountant practitioners had to have GHG assurance experience or some training in GHG assurance. Given that audit firms are in the early stages of implementing MDGHGTs, the number of financial auditors who had moved across to provide assurance on GHG emissions was relatively small when these studies were conducted. As this area of assurance expands, future research can consider the conditions under which different team formats work better than others.

REFERENCES

Agoglia, C. P., R. C. Hatfield, and J. F. Brazel. 2009. The effects of audit review format on review team judgments. *Auditing: A Journal of Practice & Theory* 28 (1):95-111.

Amason, A. C., and H. J. Sapienza. 1997. The effects of top management team size and interaction norms on cognitive and affective conflict. *Journal of Management* 23 (4):495-516.

American Institute of Certified Public Accountants (AICPA). 2002. *Statement on Auditing Standards* (*SAS*) *No. 99: Consideration of fraud in a financial statement audit*. New York: AICPA.

Argote, L., and A. A. Kane. 2003. Learning from direct and indirect experience in organizations: The effects of experience content, timing, and distribution. In *Group Creativity: Innovation Through Collaboration*, edited by P. B. Paulus and B. A. Nijstad. New York: Oxford University Press, 277-303.

Arruñada, B., and C. Paz-Ares. 1997. Mandatory rotation of company auditors: A critical examination. *International Review of Law and Economics* 17 (1):31-61.

Asare, S. K., and A. Wright. 1997. Hypothesis revision strategies in conducting analytical procedures. *Accounting, Organizations and Society* 22 (8):737-755.

———, J. Cohen, and G. Trompeter. 2005. The effect of non-audit services on client risk, acceptance and staffing decisions. *Journal of Accounting and Public Policy* 24 (6):489-520.

———, C. Haynes, and G. Jenkins. 2007. The effect of client and preparer risk factors on workpaper review effectiveness. *Behavioral Research in Accounting* 19:1-17.

———, A. M. Wright, and G. M. Trompeter. 2000. The effect of accountability and time budgets on auditors' testing strategies. *Contemporary Accounting Research* 17 (4):539-560.

Australian Government Com Law 2009. National Greenhouse and Energy Reporting (NGER) Act 2007. Available at: http://www.comlaw.gov.au/Details/C2009C00122.

Australian Government Clean Energy Regulator (Australian CER). 2014. About the carbon pricing mechanism. Available at: http://www.cleanenergyregulator.gov.au/Carbon-Pricing-Mechanism/About-theMechanism/Pages/default.aspx

Australian Government Department of Climate Change and Energy Efficiency (Australian DCCEE). 2010. National Greenhouse and Energy Reporting System Measurement: Technical guidelines. Available at: http://www.climatechange.gov.au/sites/climatechange/files/documents/03_2013/nger-technical-guidelines-2010-pdf.pdf

Balkundi, P., and D. A. Harrison. 2006. Ties, leaders, and time in teams: strong inference about network structure's effects on team viability and performance. *Academy of Management Journal* 49 (1):49-68.

Bantel, K. A., and S. E. Jackson. 1989. Top management and innovations in banking: Does the composition of the top team make a difference? *Strategic Management Journal* 10:107-124.

Baron, R. M., and D. A. Kenny. 1986. The moderator-mediator variable distinction in scoial psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology* 51 (6):1173-1182.

Beasley, M. S., and J. G. Jenkins. 2003. A primer for brainstorming fraud risks. *Journal of Accountancy* 196 (6):32-38.

Bedard, J., and S. Biggs. 1991. Pattern recognition, hypotheses generation, and auditor performance in an analytical task. *The Accounting Review* (July):622-642.

Bhappu, A. D., T. L. Griffith, and G. B. Northcraft. 1997. Media effects and communication bias in diverse groups. *Organizational Behavior and Human Decision Processes* 70 (3):199-205.

Blau, P. M. 1977. *Inequality and heterogeneity*. New York: Free Press.

Boritz, J. E., L. A. Robinson, C. Wong, and N. Kochetova-Kozloski. 2014. Use of specialists during an audit. *University of Waterloo and Saint Mary's University Working Paper*.

Bowers, C. A., J. A. Pharmer, and E. Salas. 2000. When member homogeneity is needed in work team: a meta-analysis. *Small Group Res.* 31:305-327.

Brammer, S., and S. Pavelin. 2008. Factors influencing the quality of corporate environmental disclosure. *Business Strategy and the Environment* 17 (2):120-136.

Brass, D. J. 1985. Technology and the structuring of jobs: Employee satisfaction, performance, and influence. *Organizational Behavior and Human Decision Processes* 35:216-240.

Brazel, J. F., C. P. Agoglia, and R. C. Hatfield. 2004. Electronic versus face-to-face review: The effects of alternative forms of review on auditors' performance. *The Accounting Review* 79 (4):949-966.

Brown, S. L., and K. M. Eisenhardt. 1995. Product development: past research, present findings, and future directions. *Academy of Management Review* 20 (2):343-378.

Bunderson, J. S., and K. M. Sutcliffe. 2002. Comparing alternative conceptualizations of functional diversity in management teams: Process and performance effects. *The Academy of Management Journal* 45 (5):875-893.

Camacho, L. M., and P. B. Paulus. 1995. The role of social anxiousness in group brainstorming. *Journal of Personality and Social Psychology* 68 (6):1071-1080.

Campion, M. A., G. J. Medsker, and A. C. Higgs. 1993. Relations between work group characteristics and effectiveness: Implications for designing effective work groups. Personnel psychology. 46:823-850.

———, E. M. Papper, and G. J. Medsker. 1996. Relations between work team characteristics and effectiveness: Replication and extension. Personnel Psychology. 49:429-452.

Carbon Disclosure Project (CDP). 2013. Sector insights: what is driving climate change action in the world's largest companies? Global 500 Climate Change Report

2013. Available at:https://www.cdp.net/cdpresults/cdp-global-500-climate-change-report-2013.pdf

Carcello, J. V., and A. L. Nagy. 2004. Audit firm tenure and fraudulent financial reporting. *Auditing: : A Journal of Practice & Theory* 23 (2):5-5.

Carey, P., and R. Simnett. 2006. Audit partner tenure and audit quality. *The Accounting Review* 81 (3):653-676.

Carlile, P. R. 2004. Transferring, translating, and transforming: An integrative framework for managing knowledge across boundaries. *Organization Science* 15 (5):555-568.

Carpenter, T. 2007. Audit team brainstorming, fraud risk identification, and fraud risk assessment: Implications of sas no.99. *The Accounting Review* 82 (5):1119-1140.

———, J. L. Reimers, and P. Z. Fretwell. 2011. Internal auditors' fraud judgments: The benefits of brainstorming in groups. *Auditing: A Journal of Practice & Theory* 30 (3):211-224.

Chatman, J. A., and F. J. Flynn. 2001. The influence of demographic heterogeneity on the emergence and consequences of cooperative norms in work teams. *Academy of Management Journal* 44:956-974.

Corporate Register. 2008. Assure view: The CSR assurance statement report. Available at: http://www.corporateregister.com/pdf/AssureView.pdf.

Chen, G., B. Kirkman, R. Kanfer, D. Allen, and B. Rosen. 2007. A multilevel study of leadership, empowerment, and performance in teams. *Journal of Applied Psychology* 92 (2):331-346.

Chen, X. C., K. T. Trotman, and F. H. Zhou. 2014. Nominal versus interacting electronic fraud brainstorming in hierarchical audit teams. *Accounting Review* Forthcoming.

Chen, C., D. Young, and Z. Zhuang. 2013. Externalities of mandatory ifrs adoption: Evidence from cross-border spillover effects of financial information on investment efficiency. *The Accounting Review* 88 (3):881-914.

Chung, H., and S. Kallapur. 2003. Client importance, nonaudit services, and abnormal accruals. *The Accounting Review* 78 (4):931-955.

Cohen, J. 1960. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement* 20 (1):37-46.

Cohen, S. G., and D. E. Bailey. 1997. What makes teams work: Group effectiveness Reseaarch from the Shop Floor to the Executive Suite. *Journal of Management* 23 (3):239-290.

Cohen, J. R., L. L. Holder-Webb, L. Nath, and D. Wood. 2012. Corporate reporting of nonfinancial leading indicators of economic performance and sustainability. *Accounting Horizons* 26 (1):65-90.

Cole, M. S., M. Z. Carter, and Z. Zhang. 2013. Leader–team congruence in power distance values and team effectiveness: The mediating role of procedural justice climate. *Journal of Applied Psychology* 98 (6):962-973.

Craswell, A., D. J. Stokes, and J. Laughton. 2002. Auditor independence and fee dependence. *Journal of Accounting and Economics* 33 (2):253-275.

Curral, L. A., R. H. Forrester, J. F. Dawson, and M. A. West. 2001. It's what you do and the way that you do it: Team task, team size, and innovation-related group processes. *European Journal of Work and Organizational Psychology* 10 (2):187-204.

Dahlin, K. B., L. R. Weingart, and P. J. Hinds. 2005. Team diversity and information use. *The Academy of Management Journal* 48 (6):1107-1123.

De Dreu, C. K. W., and L. R. Weingart. 2003. Task versus relationship conflict, team performance, and team member satisfaction: a meta-analysis. *Journal of Applied Psychology* 88 (4):741-749.

———, and M. A. West. 2001. Minority dissent and team innovation: the importance of participation in decision making. *Journal of Applied Psychology* 86 (6):1191-1201.

DeAngelo, L. E. 1981. Auditor size and audit quality. *Journal of Accounting and Economics* 3 (3):183-199.

Deis Jr, D. R., and G. Giroux. 1996. The effect of auditor changes on audit fees, audit hours, and audit quality. *Journal of Accounting and Public Policy* 15 (1):55-76.

Deloitte. 2011. Australia's carbon pricing mechanism: Key issues for business. Available at: http://www.deloitte.com/assets/DcomAustralia/Local%20Assets/Documents/Services/Climate%20change%20and%20sustainability/Deloitte_carbon_pricing_mechanism%20.pdf.

Department for Environment Food and Rural Affairs (Defra). 2009. Guidance on how to measure and report your greenhouse gas emissions. Available at : https://www.gov.uk/government/publications/guidance-on-how-to-measure-and-report-your-greenhouse-gas-emissions.

Dennis, A. R., J. E. Aronson, W. G. Heninger, and E. D. Walker. 1999. Sructuring time and task in electronic brainstorming. *Management Information Systems Quarterly* 23 (1):95-108.

———, and J. S. Valacich. 1993. Computer brainstorms: More heads are better than one. *Journal of Applied Psychology* 78 (4):531-537.

Dhaliwal, D. S., S. Radhakrishnan, A. H. Tsang, and Y. G. Yang. 2011. Nonfinancial disclosure and analyst forecast accuracy: International evidence on corporate social responsibility disclosure. *The Accounting Review* 87 (3):723-759.

Diehl, M., and W. Stroebe. 1987. Productivity loss in brainstorming groups: Toward the solution of a riddle. *Journal of Personality and Social Psychology* 53 (3):497-509.

———. 1991. Productivity loss in idea-generating groups: tracking down the blocking effect. *Journal of Personality and Social Psychology* 61 (3):392-403.

Dunbar, K. 1997. How scientists think: On-line creativity and conceptual change in science. In *Creative Thought: An investigation of conceptual structures and processes*, edited by T. B. Ward, S. M. Smith and J. Vaid. Washington D.C: American Psychological Association Press, 461-493.

Fargher, N. L., D. Mayorga, and K. T. Trotman. 2005. A field-based analysis of audit workpaper review. *Auditing: A Journal of Practice & Theory* 24 (2):85-110.

Faure, C. 2004. Beyond brainstorming: effects of different group procedures on selection of ideas and satisfaction with the process. *The Journal of Creative Behavior* 38 (1):13-34.

Favere-Marchesi, M., and C. E. N. Emby. 2005. The Impact of Continuity on Concurring Partner Reviews: An Exploratory Study. *Accounting Horizons* 19 (1):1-10.

Fay, D., C. Borrill, Z. Amir, R. Haward, and M. A. West. 2006. Getting the most out of multidisciplinary teams: A multi-sample study of team innovation in health care. *Journal of Occupational & Organizational Psychology* 79 (4):553-567.

Fong, P. S. W. 2003. Knowledge creation in multidisciplinary project teams: an empirical study of the processes and their dynamic interrelationships. *International Journal of Project Management* 21 (7):479-486.

Francis, J. R. 2004. What do we know about audit quality? *The British Accounting Review* 36 (4):345-368.

Fritz, M. S., and D. P. MacKinnon. 2007. Required sample size to detect the mediated effect. *Psychological Science* 18 (3):233-239.

Fukukawa, H., and T. J. Mock. 2011. Audit risk assessments using belief versus probability. *Auditing: A Journal of Practice & Theory* 30 (1):75-99.

Gibbins, M., S. A. McCracken, and S. E. Salterio. 2007. The chief financial officer's perspective on auditor-client negotiations. *Contemporary Accounting Research* 24 (2):387-422.

———, and J. D. Newton. 1994. An empirical exploration of complex accountability in public accounting. *Journal of Accounting Research* 32 (2):165-186.

———, S. Salterio, and A. Webb. 2001. Evidence about auditor–client management negotiation concerning client’s financial reporting. *Journal of Accounting Research* 39 (3):535-563.

———, and K. T. Trotman. 2002. Audit review: managers' interpersonal expectations and conduct of the review. *Contemporary Accounting Research* 19 (3):411-444.

Global Reporting Initiative (GRI). 2013. The external assurance of sustainability reporting. Available at: https://www.globalreporting.org/resourcelibrary/GRI-Assurance.pdf

Goodman, P. S. 1996. Impact of task and technology and group performance. In *Designing effective work groups*, edited by P. S. Goodman. San Francisco: Jossey-Bass, Inc.

Graetz, K. 2011. Understanding Group Work. Available at: http://elearning.winona.edu/wikiworkshop/wikiworkshop2.html.

Gray, R. 2000. Current developments and trends in social and environmental auditing, reporting and attestation: A review and comment. *International Journal of Auditing* 4:247-268.

Green, W. J., and Q. Li. 2009. Evidence of an expectiation gap for greenhouse gas emissions assurance. *Accounting, Auditing & Accountability Journal* 25 (1):146-173.

———, and K. T. Trotman. 2003. An examination of different performance outcomes in an analytical procedures task. *Auditing: A Journal of Practice & Theory* 22 (2):219-235.

Griffith, E. E. 2014. Auditing complex estimates: the interaction of audit-team specialists' caveats and client source credibility. *University of Wisconsin-Madison Working Paper.*

Guzzo, R. A., and M. W. Dickson. 1996. Teams in organizations: recent research on performance and effectiveness. *Annual Review of Psychology* 47:307-338.

Haas, M. R. 2010. The double-edged swords of autonomy and external knowledge: analyzing team effectiveness in a multinational organization. *Academy of Management Journal* 53 (5):989-1008.

Hackenbrack, K., and W. R. Knechel. 1997. Resource allocation decisions in audit engagements. *Contemporary Accounting Research* 14 (3):481-499.

Hackman, J. 1987. The design of work teams. *Handbook of organizational behavior*:315-342.

———, and G. R. Oldman. 1975. Development of the job diagnostic survey. *Journal of Applied Psychology* 60 (2):159-170.

Hammersley, J. S., K. M. Johnstone, and K. Kadous. 2011. How do audit seniors respond to heightened fraud risk? *Auditing: A Journal of Practice & Theory* 30 (3):81-101.

Haukoos, J. S., and R. J. Lewis. 2005. Advanced statistics: Bootstrapping confidence intervals for statistics with “difficult” distributions. *Academic Emergency Medicine* 12 (4):360-365.

Hill, G. W. 1982. Group versus individual performance: Are N+1 heads better than one? *Psychological Bulletin* 91 (3):517-539.

Hirst, D. E., and L. Koonce. 1996. Audit analytical procedures: a field investigation. *Contemporary Accounting Research* 13 (2):457-486.

Hoffman, V. B., and M. F. Zimbelman. 2009. Do strategic reasoning and brainstorming help auditors change their standard audit procedures in response to fraud risk? *The Accounting Review* 84 (3):811-837.

Holland, J. L. 1973. *Making vocational choices: A theory of carrers*. Englewood Cliffs, N.J.: Prentice-Hall.

Homan, A. C., J. R. Hollenbeck, S. E. Humphrey, D. Van Knippenberg, D. R. Ilgen, and G. A. Van Kleef. 2008. Facing differences with an open mind: openness to experience, salience of intragroup differences, and performance of diverse work groups. *Academy of Management Journal* 51 (6):1204-1222.

———, D. van Knippenberg, G. A. Van Kleef, and C. K. W. De Dreu. 2007. Interacting dimensions of diversity: Cross-categorization and the functioning of diverse work groups. *Group Dynamic: Theory, Research, and Practice* 11 (2):79-94.

Horwitz, S. K., and I. B. Horwitz. 2007. The effects of team diversity on team outcomes: A meta-analytic review of team demography. *Journal of Management* 33 (6):987-1015.

Hu, J., and R. C. Liden. 2011. Antecedents of team potency and team effectiveness: An examination of goal and process clarity and servant leadership. *Journal of Applied Psychology*, *96*(4), 851-862.

Huber, G. P., and K. Lewis. 2010. Cross-Understanding:Implications for Group Cognition and Performance. *Academy of Management Review* 35 (1):6-26.

Huggins, A., W. J. Green, and R. Simnett. 2011. The competitive market for assurance engagements on greenhouse gas statements: Is there a role for assurers from the accounting profession? *Current Issues in Auditing* 5 (2):A1-A12.

Humphrey, S. E., J. D. Nahrgang, and F. P. Morgeson. 2007. Integrating motivational, social, and contextual work design features: A meta-analytic summary and theoretical extension of the work design literature. *Journal of Applied Psychology* 92 (5):1332-1356.

Iacobucci, D. 2009. Everything you always wanted to know about SEM (structural equations modeling) but were afraid to ask. *Journal of Consumer Psychology* 19 (4):673-680.

Ilgen, D. R., J. R. Hollenbeck, M. Johnson, and D. Jundt. 2005. Teams in organizations: From input-process-output models to imoi models. *Annual Review of Psychology* 56 (1):517-543.

International Auditing and Assurance Standard Board (IAASB). 2009. Consultation paper. Assurance on a Greenhouse Gas Statement. Available at: http://web.ifac.org/download/ ISQC1_standalone_2009 _Handbook.pdf .

International Federation of Accountants (IFAC). 2010a. International Standard on Auditing 200: Overall objectives of the independent auditor and the conduct of an audit in accordance with international standards on auditing. Available at: http://www.ifac.org/sites/default/files/downloads/a008-2010-iaasb-handbook-isa-200.pdf

———. 2010b. International Standard on Auditing 620: Using the work of an auditor's expert. Available at: http://www.ifac.org/sites/default/files/ downloads/a035-2010-iaasb-handbook-isa-620.pdf

———. 2012a. International standard on assurance engagements 3410: Assurance engagements on greenhouse gas statements. Available at: http://www.ifac.org/publications-resources/2012-handbook-international-quality-control-auditing-review-other-assurance-a.

———. 2012b. International standard on auditing 315 (revised): Identifying and assessing the risks of material misstatement through understanding the entity and tis environment. Available at: http://www.ifac.org/sites/default/ files/publications/files/A046%202012%20IAASB%20Handbook%20ISA%20315%20(Revised).pdf

———. 2013. International Standard on Assurance Engagements 3000: Revised, assurance engagements other than audits or reviews of historical financial information. Available at: http://www.ifac.org/publications-resources/international-standard-assurance-engagements-isae-3000-revised-assurance-enga

Ismail, Z., and K. T. Trotman. 1995. The impact of the review process in hypothesis generation tasks. *Accounting, Organizations and Society* 20 (5):345-357.

Jackson, S. 1996. *The consequences of diversity in multidisciplinary work teams*. edited by M. A. West: John Wiley & Sons Ltd.

———, and A. Joshi. 2011. Work team diversity. In *APA Handbook of Industrial and Organizational Psychology*, edited by S. Zedeck, 651-686.

———, A. Joshi, and N. L. Erhardt. 2003. Recent research on team and organizational diversity: SWOT analysis and implications. *Journal of Management* 29 (6):801-830.

Jehn, K. A. 1995. A multimethod examination of the benefits and detriments of intragroup conflict. *Administrative Science Quarterly* 40 (2):256-282.

———, G. B. Northcraft, and M. A. Neale. 1999. Why differences make a difference: A field study of diversity, conflict, and performance in workgroups. *Administrative Science Quarterly* 44 (4):741-763.

Johnson, E., I. K. Khurana, and J. K. Reynolds. 2002. Audit-firm tenure and the quality of financial reports. *Contemporary Accounting Research* 19 (4):637-660.

Kaplan, S., K. LaPort, and M. J. Waller. 2013. The role of positive affectivity in team effectiveness during crises. *Journal of Organizational Behavior* 34 (4):473-491.

Kearney, E., D. Gebert, and S. C. Voelpel. 2009. When and how diversity benefits teams: The importance of team members' need for cognition. *Academy of Management Journal* 52 (3):581-598.

Kerr, N. L., and R. S. Tindale. 2004. Group performance and decision making. *Annual Review of Psychology* 55:623.655.

Kiggundu, M. N. 1981. Task interdependence and the theory of job design. *Academy of Management Review* 6 (3):499-508.

Kilduff, M., R. Angelmar, and A. Mehra. 2000. Top management-team diversity and firm performance: Examining the role of cognitions. *Organization Science* 11 (1):21-34.

Kirkman, B. L., B. Rosen, P. E. Tesluk, and C. B. Gibson. 2004. The impact of team empowerment on virtual team performance: The moderating role of face-to-face interaction. *Academy of Management Journal* 47 (2):175-192.

Knechel, R.W., P. Rouse, and C. Schelleman. 2009. A modified audit production framework: Evaluating the relative efficiency of audit engagements. *The Accounting Review* 84 (5):1607-1638.

Kohn, N. W., P. B. Paulus, and Y. Choi. 2011. Building on the ideas of others: An examination of the idea combination process. *Journal of Experimental Social Psychology* 47:554-561.

———, and S. M. Smith. 2011. Collaborative fixation: effects of others' ideas on brainstorming. *Applied Cognitive Psychology* 25:359-371.

Kozlowski, S. W., and D. R. Ilgen. 2006. Enhancing the effectiveness of work groups and teams. *Psychological Science in the Public Interest* 7:77-124.

KPMG. 2008. KPMG international survey of corporate sustainability reporting. Available at: http://www.kpmg.com/aglobal/en/issuesandinsights/articlespublications/pages/sustainability-corporate-responsibility-reporting-2008.aspx.

———. 2011. KPMG international survey of corporate responsibility reporting 2011. Available at: http://www.kpmg.com/PT/pt/IssuesAndInsights/Documents/corporate-responsibility 2011.pdf

———. 2013. The KMPG Survey of Corporate Responsibility Reporting 2013. Available at: http://www.kpmg.com/global/en/issuesandinsights/articlespublications/corporate-responsibility/pages/corporate-responsibility-reporting-survey-2013.aspx

Krishnan, J., and P. C. Schauer. 2000. The differentiation of quality among auditors: Evidence from the not-for-profit sector. *Auditing: A Journal of Practice & Theory* 19 (2):9-25.

Larcker, D. F., and S. A. Richardson. 2004. Fees paid to audit firms, accrual choices, and corporate governance. *Journal of Accounting Research* 42 (3):625-658.

Larson, J. R., P. G. Foster-Fishman, and C. B. Keys. 1994. Discussion of shared and unshared informatoin in decision making groups. *Journal of Personality and Social Psychology* 67:446-461.

LePine, J. A., R. F. Piccolo, C. L. Jackson, J. E. Mathieu, and J. R. Saul. 2008. A meta-analysis of teamwork processes: Tests of a multidimensional model and relationships with team effectiveness criteria. *Personnel Psychology* 61 (2):273-307.

Lau, D. C., and J. K. Murnighan. 1998. Demographic diversity and faultlines: The compositional dynamics of organizational groups. *Acadamy of Management Review* 23:325-340.

———. 2005. Factional groups:a new vantage on demographic faultlines: the compositional dynamics of organizational groups. *Acad. Manage. Rev.* 23:325-340.

Laughlin, P. R., and A. B. Hollingshead. 1995. A theory of collective induction. *Organization Behavior and Human Decision Process* 61:94-107.

Li, C. 2009. Does client importance affect auditor independence at the office level? Empirical evidence from going-concern opinions. *Contemporary Accounting Research* 26 (1):201-230.

Lim, C.Y., and H.T. Tan. 2010. Does auditor tenure improve audit quality? Moderating effects of industry specialization and fee dependence. *Contemporary Accounting Research* 27 (3):923-957.

Libby, R., K. T. Trotman, and I. Zimmer. 1987. Member variation, recognition of expertise, and group performance. *Journal of Applied Psychology* 72 (1):81-87.

Low, K.-Y. 2004. The effects of industry specialization on audit risk assessments and audit-planning decisions. *The Accounting Review* 79 (1):201-219.

Luippold, B. L., and T. E. Kida. 2012. The impact of initial information ambiguity on the accuracy of analytical review judgments. *Auditing: A Journal of Practice & Theory* 31 (2):113-129.

Lynch, A. L., U. S. Murthy, and T. J. Engle. 2009. Fraud brainstorming using computer-mediated communication: The effects of brainstorming technique and facilitation. *The Accounting Review* 84 (4):1209-1232.

Mathieu, J. E., L. L. Gilson, and T. M. Ruddy. 2006. Empowerment and team effectiveness: An empirical test of an integrated model. *Journal of Applied Psychology* 91 (1):97-108.

———, T. S. Heffner, G. F. Goodwin, E. Salas, and J. A. Cannon-Bowers. 2000. The influence of shared mental models on team process and performance. *Journal of Applied Psychology* 85 (2):273-283.

———, M. T. Maynard, T. Rapp, and L. Gilson. 2008. Team effectiveness 1997-2007: A review of recent advancements and a glimpse into the future. *Journal of Management* 34 (3):410-476.

McGrath, J. E. 1964. *Social psychology: A brief introduction*. New York: Holt, Rinehart & Winston.

———. 1984. *Group interaction and performance*. Englewood Cliffs, NJ: Prentice-Hall.

———. 1991. Time, interaction, and performance: A theory of groups. *Small Group Research* 22 (2):147-174.

Milliken, F. J., and L. L. Martins. 1996. Searching for common threads: Understanding the multiple effects of diversity in organizational groups. *The Academy of Management Review* 21 (2):402-433.

———, C. A. Bartel, and T. R. Kurtzberg. 2003. Diversity and creativity in work groups: A dynamic perspective on the affective and cognitive processes that link diversity and performance. In *Group Creativity: Innovation Through Collaboration*, edited by P. B. Paulus and B. A. Nijstad. New York: Oxford University Press, 32-62.

Moreno, K. K., S. Bhattacharjee, and D. M. Brandon. 2007. The effectiveness of alternative training techniques on analytical procedures performance. *Contemporary Accounting Research* 24 (3):983-1014.

Morgeson, F. P., and S. E. Humphrey. 2006. The Work Design Questionnaire (WDQ): Developing and validating a comprehensive measure for assessing job design and the nature of work. *Journal of Applied Psychology* 91 (6):1321-1339.

Mullen, B., C. Johnson, and E. Salas. 1991. Productivity loss in brainstorming groups: A meta-analytic integration. *Basic and Applied Social Psychology* 12 (1):3-23.

Mumford, M., M. Mobley, C. Uhlman, and R. Reiter-Palmon. 1991. Process analytic models of creative capacities. *Creativity Research Journal* 4:91-122.

Myers, J. N., L. A. Myers, and T. C. Omer. 2003. Exploring the term of the auditor-client relationship and the quality of earnings: A case for mandatory auditor rotation? *The Accounting Review* 78 (3):779-799.

Nelson, M., and H. T. Tan. 2005. Judgment and decision making research in auditing: A task, person, and interpersonal interaction perspective. *Auditing: A Journal of Practice & Theory* 24:41-71,73-87.

Nijstad, B. A., and C. K. W. De Dreu. 2002. Creativity and group innovation. *Applied Psychology: An International Review* 51:400-406.

———, and P. B. Paulus. 2003. Group Creativity: common themes and future directions. In *Group Creativity: Innovation Through Collaboration*, edited by P. B. Paulus and B. A. Nijstad. New York: Oxford University Press, 326-339.

———, W. Stroebe, and H. F. M. Lodewijkx. 2002. Cognitive stimulation and interference in groups: Exposure effects in an idea generation task. *Journal of Experimental Social Psychology* 39 (6):531-548.

Nugent, M. 2008. IAASB Main Agenda (December 2008). In *Issues Papers-Emissions Assurance*.

O'Dwyer, B. 2011. The case of sustainability assurance: Constructing a new assurance service. *Contemporary Accounting Research* 28 (4):1230-1266.

O'Keefe, T. B., D. A. Simunic, and M. T. Stein. 1994. The production of audit services: Evidence from a major public accounting firm. *Journal of Accounting Research* 32 (2):241-261.

Osborn, A. F. 1953. *Applied Imagination*. New York: Scribner's.

———. 1957. *Applied Imagination: Principles and procedures of crative thinking*. New York, NY: Charles Scribner's Sons.

Parnes, S. J., and A. Meadow. 1959. Effect of "brainstorming" instructions on creative problem-solving by trained and untrained subjects. *Journal of Educational Psychology* 50 (4):171-176.

Paulus, P. B. 2000. Groups, teams, and creativity: the creative potential of idea-generating groups. *Applied Psychology: An International Review* 49 (2):237.

———. 2008. Forstering creativity in groups and teams. In *Handbook of organizational creativity*, edited by J. Zhou and C. Shalley. New York: Taytor & Francis Group.

———, and V. R. Brown. 2003. Enhancing ideational creativity in groups: lessons from research on brainstorming. In *Group Creativity: Innovation Through Collboration*, edited by P. B. Paulus and B. A. Nijstad. New York: Oxford University Press, 110-136.

———, and H.-C. Yang. 2000. Idea generation in groups: A basis for creativity in organizations. *Organizational Behavior and Human Decision Processes* 82 (1):76-87.

———, M. T. Dzindolet, and N. W. Kohn. 2012. Collaborative creativity-group creativity and team innovation. In *Handbook of Organizational Creativity*, edited by D. M. Michael. San Diego: Academic Press, 327-357.

———, M. T. Dzindolet, G. Poletes, and L. M. Camacho. 1993. Perception of performance in group brainstorming: The illusion of productivity. *Personality and Social Psychology Bulletin* 19:78-89.

Payne, E. A., R. J. Ramsay, and E. M. Bamber. 2010. The effect of alternative types of review on auditors’ procedures and performance. *Auditing: A Journal of Practice & Theory* 29 (1):207-220.

Pelled, L. H., K. M. Eisenhardt, and K. R. Xin. 1999. Exploring the black box: An analysis of work group diversity, conflict, and performance. *Administrative Science Quarterly* 44 (1):1-28.

Petersen, M. A. 2009. Estimating standard errors in finance panel data sets: Comparing approaches. *Review of Financial Studies* 22 (1):435-480.

Pike, B. J., M. B. Curtis, and L. Chui. 2013. How does an initial expectation bias influence auditors' application and performance of analytical procedures? *The Accounting Review* 88 (4):1413-1431.

Poulton, B. C., and M. A. West. 1999. The determinants of effectiveness in primary health care teams. *Journal of Interprofessional Care* 13 (1):7-18.

Preacher, K., and A. Hayes. 2004. SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, & Computers* 36 (4):717-731.

———, and A. F. Hayes. 2008. Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods* 40 (3):879-891.

PricewaterhouseCoopers (PwC). 2007. Building trust in emissions reporting: Global trends in emissions tradings schemes. Available at: http://www.pwc.com/la/en/ publications/ building-trust-in-emissions-reporting.jhtml.

Putman, V., L., and P. B. Paulus. 2009. Brainstorming, brainstorming rules and decision making. *Journal of Creative Behavior* 43 (1):29-45.

Randel, A. E., and K. S. Jaussi. 2003. Functional background identity, diversity, and individual performance in cross-functional teams. *Academy of Management Journal* 46 (6):763-774.

———, and J. R. Francis. 2001. Does size matter? The influence of large clients on office-level auditor reporting decisions. *Journal of Accounting and Economics* 30 (3):375-400.

Rich, J. S., I. Solomon, and K. T. Trotman. 1997a. The audit review process: A characterization from the persuasion perspective. *Accounting, Organizations and Society* 22 (5):481-505.

———. 1997b. Multi-auditor judgment/decision making research: A decade later. *Journal of Accounting Literature* 16:86-126.

Reinig, B. A., R. O. Briggs, and J. F. Nunamaker Jr. 2007. On the measurement of ideation quality. *Journal of Management Information Systems* 23 (4):143-161.

Reiter-Palmon, R., B. Wigert, and T. d. Vreede. 2012. Chapter 13 - team creativity and innovation: The effect of group composition, social processes, and cognition. In *Handbook of Organizational Creativity*, edited by D. M. Michael. San Diego: Academic Press, 295-326.

Rietzschel, E. F., B. A. Nijstad, and W. Stroebe. 2006. Productivity is not enough: A comparison of interactive and nominal brainstorming groups on idea generation and selection. *Journal of Experimental Social Psychology* 42 (2):244-251.

———, B. A. Nijstad, and W. Stroebe. 2010. The selection of creative ideas after individual idea generation: Choosing between creativity and impact. *British Journal of Psychological* 101:47-68.

Salas, E., N. J. Cooke, and M. A. Rosen. 2008. On teams, teamwork, and team performance: Discoveries and developments. *Human Factors* 50:540-547.

———, K. C. Stagl, C. S. Burke, and G. F. Goodwin. 2007. Fostering team effectiveness in organizations: Toward an integrative theoretical framework, of team performance. In *Modeling Complex systems: Motivation, cognition and social processes: Nebraska Symposium on Motivation*, edited by R. A. Dienstbier, J. W. Shuart, W. Spaulding and J. Poland. Lincoln: University of Nebraska Press, 185-243.

Selley, D. 1999. Use of specialists in assurance engagements. *Canadian Accounting and Assurance Reference Service*.+

Shaw, J. D., J. Zhu, M. K. Duffy, K. L. Scott, H.-A. Shih, and E. Susanto. 2011. A contingency model of conflict and team effectiveness. *Journal of Applied Psychology* 96 (2):391-400.

Sheehan, D., L. Robertson, and T. Ormond. 2007. Comparison of language used and patterns of communication in interprofessional and multidisciplinary teams. *Journal of Interprofessional Care* 21 (1):17-30.

Shin, S. J., and J. Zhou. 2007. When is educational specialization heterogeneity related to creativity in research and development teams? Transformational leadership as a moderator. *Journal of Applied Psychology* 92 (6):1709-1721.

Simnett, R. 1996. The effect of information selection, information processing and task complexity on predictive accuracy of auditors. *Accounting, Organizations and Society* 21 (7–8):699-719.

———, M. Nugent, and A. Huggins. 2009a. Developing an international assurance standard on greenhouse gas statements. *Acccounting Horizons* 23 (4):347-363.

———, and K. Trotman. 1989. Auditor versus model: Information choice and information processing. *The Accounting Review* 64 (3):514-528.

———, A. Vanstraelen, and W. F. Chua. 2009b. Assurance on general purpose non-financial reports: An international comparison. *The Accounting Review* 84 (3): 937–967.

Smith, C. M. 2008. Adding minority status to a source of conflict: an examination of influence processes and product quality in dyads. *European Journal of Social Psychology* 38 (1):75-83.

Smith, K. G., K. A. Smith, J. D. Olian, H. P. Sims, D. P. O'Bannon, and J. A. Scully. 1994. Top management team demography and process: The role of social integration and communication. *Administrative Science Quarterly* 39:412-438.

Sobel, M. E. 1982. Asymptotic confidence intervals for indirect effects in structural equation models. *Sociological Methodology* 13 (1982):290-312.

Sparrowe, R. T., R. C. Liden, S. J. Wayne, and M. L. Kraimer. 2001. Social networks and the performance of individuals and groups. *Academy of Management Journal* 44 (2):316-325.

Stasser, G., D. D. Stewart, and G. M. Wittenbaum. 1995. Expert roles and information exchange during discussion: The importance of knowing who knows what. *Journal of Experimental Social Psychology* 31 (3):244-265.

———, and W. Titus. 1985. Pooling of unshared information in group decision making: Biased information sampling during discussion. *Journal of Personality and Social Psychology* 53:81-93.

Steiner, I. 1972. *Group Process and Productivity*. New York, NY: Academic Press.

Stewart, G. L. 2006. A meta-analytic review of relationships between team design features and team performance. *Journal of Management* 32 (1):29-54.

———, and M. R. Barrick. 2000. Team structure and performance: Assessing the mediating role of intrateam process and the moderating role of task type. *The Academy of Management Journal* 43 (2):135-148.

Straus, S. G., and J. E. McGrarth. 1994. Does the medium matter? The interaction of task type and technology on group performance and member reactions. *Journal of Applied Psychology* 79 (1):87-97.

Stroebe, W., and M. Diehl. 1994. Why groups are less effective than their members: On productivity losses in idea-generating groups. *European Review of Social Psychology* 5 (1):271-303.

Sundstrom, E., M. McIntyre, T. Halfhill, and H. Richards. 2000. Work groups: From the hawthorne studies to work teams of the 1990s and beyond. *Group Dynamic: Theory, Research, and Practice* 4 (1):44-67.

Tan, H.T. 1995. Effects of expectations, prior involvement, and review awareness on memory for audit evidence and audit judgment. *Journal of Accounting Research* 33 (1):113-135.

Triandis, H. C., L. L. Kurowski, and M. J. Gelfand. 1994. Workplace diversity. In *Handbook of Industrial and Organizational Psychology*, edited by T. H. C., M. D. Dunnette and L. M. Hough. Palo Alto, CA: Consult. Psychol. Press., 769-827.

Trotman, K. T. 1985. The review process and the accuracy of auditor judgments. *Journal of Accounting Research* 23 (2):740-752.

———, R. Simnett, and A. Khalifa. 2009. Impact of the type of audit team discussions on auditors' generation of material frauds. *Contemporary Accounting Research* 26 (4):1115-1142.

———, A. M. Wright, and S. Wright. 2005. Auditor negotiations: An examination of the efficacy of intervention methods. *The Accounting Review* 80 (1):349-367.

———, and P. W. Yetton. 1985. The effect of the review process on auditor judgments. *Journal of Accounting Research* 23 (1):256-267.

———, P. W. Yetton, and I. R. Zimmer. 1983. Individual and group judgments of internal control systems. *Journal of Accounting Research* 21 (1):286-292.

Tschan, F., and M. von Cranach. 1996. Group task structure, processes, and outcome. In *Handbook of work group psychology*, edited by M. A. West. Chichester, UK: John Wiley, 95-121.

Urbach, N., and F. Ahlemann. 2010. Structural equation modeling in information systems research using partial least sqaures. *Journal of Information Technology Theory and Application* 11 (2):5-40.

Van Asselt, M. B. A. 2000. *Perspectives on Uncertainty and Risk*. Dordrecht: Kluwer.

Van der Vegt, G. S., and J. S. Bunderson. 2005. Learning and performance in multidisciplinary teams: the importance of collective team identification. *Academy of Management Journal* 48 (3):532-547.

Van der Vegt, G., B. Emans, and E. Van De Vliert. 2000. Team members' affective responses to patterns of intragroup interdependence and job complexity. *Journal of Management* 26 (4):633-655.

Van de Ven, A., and A. L. Delbeco. 1971. Nominal versus interacting group processes for committee decision-making effectiveness. *The Academy of Management Journal* 14 (2):203-212.

van Ginkel, W. P., R. S. Tindale, and D. van Knippenberg. 2009. Team reflexivity, development of shared task representations, and the use of distributed information in group decision making. *Group Dynamics: Theory, Research, and Practice* 13 (4):265-280.

———, and D. van Knippenberg. 2008. Group information elaboration and group decision making: The role of shared task representations. *Organizational Behavior and Human Decision Processes* 105 (1):82-97.

———, and D. van Knippenberg. 2009. Knowledge about the distribution of information and group decision making: When and why does it work? *Organization Behavior and Human Decision Process* 108:218-229.

van Knippenberg, D., C. K. W. De Dreu, and A. C. Homan. 2004. Work group diversity and group performance: an integrative model and research agenda. *Journal of Applied Psychology* 89 (6):1008-1022.

———, and M. C. Schippers. 2007. Work group diversity. *The Annual Review of Psychology* 58 (2007):515-541.

van Someren, M. W., P. Reimann, H. P. A. Boshuizen, and T. De Jong. 1998. *Learning With Multiple Representations*. Oxford, UK: Elsevier.

Wageman, R. 1995. Interdependence and group effectiveness. *Administrative Science Quarterly* 40 (1):145-180.

Warner, R. M. 2013. Mediation. In *Applied Statistics: From Bivariate Through Multivariate Techniques* SAGE Publications Inc., 645-687.

West, M. A. 2000. Reflexivity, revolution, and innovation in work teams. *Advances in Interdisciplinary Studies of Work Teams* 5:1-29.

———. 2002. Sparkling fountains or stagnant ponds: An integrative model of creativity and innovation implementation in work groups. *Applied Psychology: An International Review* 51 (3):355-424.

———, and N. R. Anderson. 1996. Innovation in top management teams. *Journal of Applied Psychology* 8 (6):680-693.

———, C. S. Borrill, J. F. Dawson, F. Brodbeck, D. A. Shapiro, and B. Haward. 2003. Leadership clarity and team innovation in health care. *The Leadership Quarterly* 14 (4–5):393-410.

Wiersema, M. F., and K. A. Bantel. 1992. Top management team demography and corporate strategic change. *Academy of Management Journal* 35 (1):91-121.

Williams, K. Y., and C. A. O'Reilly. 1998. Demography and diversity in organizations: A review of 40 years of research. In *Research in Organizational Behavior, Vol 20, 1998*. Stamford: Jai Press Inc, 77-140.

Wittenbaum, G. M. 2000. The bias toward discussing shared information: Why are high-status group members immune? *Communication Research* 27:379-400.

World Bank. 2012. State and trends of carbon pricing 2014. World Bank, Washington, DC. Available at: http://www.wds.worldbank.org/external/default/WDSContentServer/WDSP/IB/2013/04/22/000445729_20130422140212/ Rendered/PDF/768370AR0State00Box374391B00PUBLIC0.pdf

———. 2013. Mapping carbon pricing initiatives: developments and prospects, 2013. World Bank, Washington, DC. Available at:https://www.thepmr.org/system/files/documents/ Mapping % 20 Carbon%20Pricing%20Initiatives-%20Developments%20 and% 20 Prospects.pdf

———. 2014. State and trends of carbon pricing 2014. World Bank, Washington, DC. Available at: http://www.wds.worldbank.org/external/default/WDSContentServer/WDSP/IB/2014/05/27/000456286_20140527095323/Rendered/PDF/882840AR0REPLA00EPI2102680Box385232.pdf

World Business Council for Sustainable Development and World Resources Institute (WBCSD and WRI).2004. The greenhouse gas protocol—A corporate accounting and reporting standard. Available at:http://www.GHGprotocol.org/standards/publications.

———. 2005. The greenhouse gas protocol—The GHG protocol for project accounting. Available at: http://www.ghgprotocol.org/files/ghgp/ ghg_project_protocol.pdf.

Zhou, S., R. Simnett, and W. Green. 2013. The effect of legal environment and corporate governance on the decision to assure and assurance provider choice: Evidence from the GHG assurance market. *UNSW Working Paper.*

APPENDIX 1
Study One Research Instrument

PARTICIPANT INFORMATION STATEMENT

<u>Evaluating the composition and functioning of multidisciplinary greenhouse gas assurance engagement teams</u>

You are invited to participate in a study of multidisciplinary greenhouse gas (GHG) assurance teams. This is an important area on which to gain insights as the multidisciplinary nature of GHG assurance teams is a major distinguishing factor of this type of engagement. The importance of this issue has been recognised by the International Auditing and Assurance Standards Board's attempts to develop an international assurance standard for GHG assurance engagements. This research on the factors affecting the functioning of GHG assurance teams has the ability to inform the development of this assurance standard. We are therefore grateful for your help to learn more about this important and evolving area.

If you decide to participate, we will ask you to fill out the attached questionnaire. It should take about 30 minutes to complete.

Any information that is obtained in connection with this study and that can be identified with you will remain confidential and will be disclosed only with your permission, except as required by law. We plan to discuss/publish the results at conferences and in accounting/auditing academic journals. In any publication, information will be provided in such a way that you cannot be identified.

Complaints may be directed to the Ethics Secretariat, The University of New South Wales, SYDNEY 2052 AUSTRALIA (phone 9385 4234, fax 9385 6648, email ethics.sec@unsw.edu.au) quoting approval no. 106037. Any complaint you make will be investigated promptly and you will be informed of the outcome.

Completion and return of the questionnaire will be regarded as your consent for participating in this study. Your decision whether or not to participate will not prejudice your future relations with the University of New South Wales. If you decide to participate, you are free to withdraw your consent and discontinue participation at any time, without prejudice.

If you have any questions, please feel free to ask us. If you have any questions later, Professor Roger Simnett (ph. (02) 9385 5825) will be happy to answer them. If you would like a summary of the results of this research phone or fax Roger on (02) 9662 5815, or email him at R.Simnett@unsw.edu.au.

You may keep this form for your records.

INSTRUCTIONS FOR ANSWERING QUESTIONNAIRE

The purpose of this study is to learn about the factors affecting the functioning of multidisciplinary assurance teams. The types of engagements we are focusing on are GHG/sustainability assurance engagements, where multidisciplinary teams (including both assurance and subject matter experts) are common. This research is expected to inform the development of the relevant international standards: for GHG assurance ISAE 3410; and for a review ISAE 3000.

Please take your time in responding to this questionnaire. Your responses are important because we are contacting only a limited number of people to participate in this study. The attached questionnaire has been thoroughly tested with professionals like you who perform GHG/sustainability assurance engagements to make sure that it is clear and will not waste your time. It should take approximately 30 minutes to complete. Please answer **all** of the attached questions in the order they are asked. **Do not change** any answer once you have written it – using hindsight may invalidate the process. Please answer on your own, without discussing the questions with anyone.

Please note that there are **two tasks** in this questionnaire. The first involves answering questions regarding a recent GHG/sustainability assurance engagement in which you feel the GHG/sustainability assurance team of which you were a part **did work effectively together**. The second task is similar to the first but involves a GHG/sustainability assurance engagement in which you feel the team **did not work as effectively together** as in the first task. The final part of the questionnaire elicits demographic information.

To reduce your time, some questions have a response scale with 9 points, some identified and others representing mid positions, such as:

Much less busy than normal	As busy as normal	Much busier than normal

|____|____|____|____|____|____|____|____|

In such scales, please just circle which one of the nine points is closest to your view.

We would appreciate if you answer each question frankly and anonymously. Do not identify yourself, your firm, your clients or anyone else. We assure you that all replies will be kept completely confidential – our responsibility to you parallels yours to your clients. You can be assured that **only** aggregate results will be reported.

If you have any questions while you are completing the materials, please do not hesitate to ask us.

Once again, thank you for your participation.

EVALUATING THE FUNCTIONING OF MULTIDISCIPLINARY GHG/SUSTAINABILITY ASSURANCE ENGAGEMENT TEAMS

PART 1: EFFECTIVE MULTIDISCIPLINARY TEAM

Consider a recent GHG/sustainability assurance engagement you were involved in. Choose an engagement where at least one GHG/sustainability assurance team member was from a financial audit background and at least one team member had other than a financial audit background. Select an engagement in which you feel the team worked *effectively* together. Please answer the following questions in relation to this engagement.

(1A) Your role in this assurance engagement

What was your role in the engagement identified above?
Assurance Team Leader (please tick if applicable)_____
Assurance Team Member (please specify role) ________________________________

(1B) Client and engagement GHG characteristics

(a) How large was the client? Estimated annual revenue: $______________________
and/or estimated annual GHG emissions (Tonnes/$CO_{2\text{-}e}$) _____________________

(b) Approximately how many facilities does the client have (please tick)?
Single facility____________; 2-5 facilities_____________; 5+ facilities_________

(c) What industry sector was the client in?
Production________; Utilities________; Finance__________; Mining_________;
Other (e.g. services, etc. Please specify)___________________________________

(d) Was the client a public company? _____; or private _____; or something else
(if something else, identify generally, e.g. facility only, government organization, etc.)____

(e) Reason the client undertook the assurance engagement:
Regulatory requirement (please specify)_____________________________________
Voluntary reporting, e.g. sustainability report (please specify)__________________

(f) Was this (please tick): (1) a limited assurance engagement________________?; or
(2) a reasonable assurance engagement________________________________?

(g) Rate the complexity of the client's GHG emissions profile compared to other similar GHG/sustainability assurance engagements you have undertaken:

Much lower profile complexity than others	Average profile complexity	Much higher profile complexity than others

|______|______|______|______|______|______|______|______|

(1C) Client-assurer relationship

(a) How many previous GHG/sustainability assurance engagements have you/your assurance firm undertaken for this client?___________________________________

(b) Does your assurance firm act in any other capacity for this client (e.g. as auditor of their financial statements)? If so, please specify______________________________

(c) How would you rate the relative importance of this client to your assurance firm?

Low importance	Average importance	High importance

|______|______|______|______|______|______|______|______|

(1D) Client's systems to capture and record GHG/sustainability data

(a) Which measurement methodologies did the client use in quantifying their GHG/sustainability data?

Direct measurement	_______	%
Estimation techniques	_______	%
Total	100	%

If estimation techniques were used, please specify in general terms how this was done:____

(b) Please estimate **the proportion of each type** of scope emissions in this entity's report **that was assured**:

Scope 1 emissions______%;Scope 2 emissions______%;Scope 3 emissions______%

(c) How well developed were the client's systems to capture and record GHG/sustainability data?

Not at all developed	Adequately developed	Very well developed

|______|______|______|______|______|______|______|______|

(1E) Client's report preparers

(a) To your knowledge, how many report preparers worked on this inventory?_______

(b) Rate the availability of the client's report preparer(s) to the assurance team for this engagement:

Low availability	Satisfactory availability	High availability

|_____|_____|_____|_____|_____|_____|_____|_____|

(c) Assess the capabilities of the report preparer(s) compared to other similar GHG/sustainability assurance engagements you have undertaken:

Much lower capabilities than others	Average capabilities	Much higher capabilities than others

|_____|_____|_____|_____|_____|_____|_____|_____|

(d) Rate the quality of the work of the report preparer(s) compared to other similar GHG/sustainability assurance engagements you have undertaken:

Much lower quality work than others	Average quality work	Much higher quality work than others

|_____|_____|_____|_____|_____|_____|_____|_____|

(e) Rate the quality of the report preparer(s) documentation compared to other similar GHG/sustainability assurance engagements you have undertaken:

Much lower quality documentation than others	Average quality documentation	Much higher quality documentation than others

|_____|_____|_____|_____|_____|_____|_____|_____|

(1F) GHG/sustainability assurance engagement team background information

(a) In accordance with assurance standards, members of the GHG/sustainability assurance team can be classified as firm employees, internal experts or external experts. On the table on the following page please fill out the following details for the members of your GHG/sustainability assurance team (if known). If the Team Leader for this engagement had a different title please write (T.L.) next to their title.

Firm employees other than internal experts (please list by team role title only – do not include names)	**Professional background (Financial audit/Engineering /Scientific/Other)**	**Degree of overall involvement in the engagement (High (H)/Medium (M)/Low (L))**	**Familiarity with team members (First time worked with; Unfamiliar (U) / Familiar (F))**
1)			
2)			
3)			
4)			
5)			
Firm employees categorised as internal experts (please list by team role title only)			
1)			
2)			
3)			
External experts (please list by team role title only)			
1)			
2)			
3)			

(b) If an external expert was, or external experts were, part of the team, identify why such as an expert was required (please tick):

He/she possessed specialised knowledge not possessed by firm employees________

Firm employees with this specialist knowledge were not available at the time of the engagement____

Specific request of the client____

Other (please specify)____

(1G) Evaluation of GHG /sustainability assurance engagement team

(a) Approximately how long did the assurance engagement take? _______hours

(b) What is your overall rating of how well the GHG/sustainability assurance team worked together?

Did not work well together	Worked together satisfactory	Worked very well together

|______|______|______|______|______|______|______|______|

(c) List the factors that you feel contributed to the GHG/sustainability assurance team working well together:

____

____

(d) List the factors that you feel inhibited the GHG/sustainability assurance team's ability to work well together:

____

____

(1H) Evaluation of the GHG/sustainability assurance engagement team process

The following two questions relate to the preliminary team discussions (if any) at the planning stage that involved key members of the GHG/sustainability assurance team.

(a) Given the multi-disciplinary nature of the GHG/sustainability assurance team, did you feel there was sufficient discussion time in the early stages of the engagement to share diverse information and perspectives?

There was not enough time	There was sufficient time	There was more than enough time

|______|______|______|______|______|______|______|______|

(b) Given the multi-disciplinary nature of the GHG/sustainability assurance team, did you feel there was sufficient elaboration and integration of different information and perspectives from different team members?

There was not enough	It was sufficient	There was more than enough

|______|______|______|______|______|______|______|______|

The following three questions relate to the evidence gathering and evaluation stages of the assurance engagement

(c) To what extent were the following assurance procedures used to gather evidence?

Tests of controls	_______%
Detailed substantive testing	_______%;
Analytical procedures	_______%;
Total	100 %

If other, please specify____

(d) To what extent was there a clear separation between the information search/collection stage and the information processing/decision making stage?

There was no separation	There was a partial separation	There was a clear separation

|_____|_____|_____|_____|_____|_____|_____|_____|

(e) To what extent did the GHG/sustainability assurance team discuss the information collected **before** final evaluations and decisions were made?

Our team did not do this	Our team partially did this	Our team did this

|_____|_____|_____|_____|_____|_____|_____|_____|

(f) Was there any other mechanism by which different information and perspectives from different team members was shared and integrated **at any stage** of the engagement? If so, please specify:

____

____

(g) In your opinion, what factors, if they had been present, would have made the team work together better? These may be some of the factors listed above or others.

____

____

____

____

____

(1I) Final comments

Please comment on any issues not properly covered in the above questions, or anything else you wish to raise regarding the team for this GHG/sustainability assurance engagement.

____

____

____

____

____

____

____

PART 2: LESS EFFECTIVE MULTIDISCIPLINARY TEAM

Consider another recent GHG/sustainability assurance engagement you were involved in. Choose an engagement where at least one GHG/sustainability assurance team member was from a financial audit background and at least one team member had other than a financial audit background. Select an engagement in which you feel the team *did not work as effectively* together as the team covered in **Part 1**. Please answer the following questions in relation to this engagement.

(2A) Your role in this assurance engagement

What was your role in the engagement identified above?
Assurance Team Leader (please tick if applicable)______________________________
Assurance Team Member (please specify role) ________________________________.

(2B) Client and engagement GHG characteristics

(a) How large was the client? Estimated annual revenue: $___________________ **and/or** estimated annual GHG emissions (Tonnes/$CO_{2\text{-}e}$) ________________________

(b) Approximately how many facilities does the client have (please tick)?
Single facility___________; 2-5 facilities___________; 5+ facilities_____________

(c) What industry sector was the client in?
Production_________; Utilities_________; Finance_________; Mining_________;
Other (e.g. services, etc. Please specify)_____________________________________

(d) Was the client a public company? _____; or private _____; or something else (if something else, identify generally, e.g. facility only, government organisation, etc.)____

(e) Reason the client undertook the assurance engagement:
Regulatory requirement (please specify)______________________________________
Voluntary reporting, e.g. sustainability report (please specify)____________________

(f) Was this (please tick): (1) a limited assurance engagement________________?; or (2) a reasonable assurance engagement______________________________________?

(g) Rate the complexity of the client's GHG emissions profile compared to other similar GHG assurance engagements you have undertaken:

Much lower profile complexity than others	Average profile complexity	Much higher profile complexity than others

|______|______|______|______|______|______|______|______|

(2C) Client-assurer relationship

(a) How many previous GHG/sustainability assurance engagements have you/your assurance firm undertaken for this client?___________________________________

(b) Does your assurance firm act in any other capacity for this client (e.g. as auditor of their financial statements)? If so, please specify_______________________________

(c) How would you rate the relative importance of this client to your assurance firm?

Low importance Average importance High importance

|______|______|______|______|______|______|______|______|

(2D) Client's systems to capture and record GHG/sustainability data

(a) Which measurement methodologies did the client use in quantifying their GHG/ sustainability data?

Direct measurement _______%
Estimation techniques _______ %
Total 100 %

If estimation techniques were used, please specify in general terms how this was done:

____

(b) Please estimate **the proportion of each type** of scope emissions in this entity's report **that was assured**:

Scope1 emissions______%; Scope2 emissions______%; Scope 3 emissions______%

(c) How well developed were the client's systems to capture and record GHG/sustainability data?

Not at all developed Adequately developed Very well developed

|______|______|______|______|______|______|______|______|

(2E) Client's report preparers

(a) To your knowledge, how many report preparers worked on this inventory?________

(b) Rate the availability of the client's report preparer(s) to the GHG/sustainability assurance team for this engagement:

Low availability Satisfactory availability High availability

|______|______|______|______|______|______|______|______|

(c) Assess the capabilities of the report preparer(s) compared to other similar GHG/sustainability assurance engagements you have undertaken:

Much lower capabilities than others	Average capabilities	Much higher capabilities than others

|______|______|______|______|______|______|______|______|

(d) Rate the quality of the work of the report preparer(s) compared to other similar GHG/sustainability assurance engagements you have undertaken:

Much lower quality work than others	Average quality work	Much higher quality work than others

|______|______|______|______|______|______|______|______|

(e) Rate the quality of the report preparer(s) documentation compared to other similar GHG/sustainability assurance engagements you have undertaken:

Much lower quality documentation than others	Average quality documentation	Much higher quality documentation than others

|______|______|______|______|______|______|______|______|

(2F) GHG/sustainability assurance engagement team background information

(a) In accordance with assurance standards, members of the GHG/sustainability assurance team can be classified as firm employees, internal experts or external experts. On the table on the following page please fill out the following details for the members of your GHG/sustainability assurance team (if known). If the Team Leader for this engagement had a different title please write (T.L.) next to their title.

Firm employees other than internal experts (please list by team role title only – do not include names)	**Professional background (Financial audit/Engineering/Scientific/Other)**	**Degree of overall involvement in the engagement (High (H)/Medium (M)/Low (L))**	**Familiarity with team members (First time worked with; Unfamiliar (U) / Familiar (F))**
1)			
2)			
3)			
4)			
5)			
Firm employees categorised as internal experts (please list by team role title only)			
1)			
2)			
3)			
External experts (please list by team role title only)			
1)			
2)			
3)			

(b) If an external expert was, or external experts were, part of the team, identify why such as an expert was required (please tick):

He/she possessed specialised knowledge not possessed by firm employees________

Firm employees with this specialist knowledge were not available at the time of the engagement________

Specific request of the client________

Other (please specify)________

(2G) Evaluation of GHG/sustainability assurance engagement team

(a) Approximately how long did the assurance engagement take? ________hours

(b) What is your overall rating of how well the GHG/sustainability assurance team worked together?

Did not work well together	Worked together satisfactorily	Worked very well together

|______|______|______|______|______|______|______|______|

(c) List the factors that you feel contributed to the GHG/sustainability assurance team working well together:

____

____

(d) List the factors that you feel inhibited the GHG/sustainability assurance team's ability to work well together:

____

____

(2H) Evaluation of the GHG/sustainability assurance engagement team process

The following two questions relate to the preliminary team discussions (if any) at the planning stage that involved key members of the GHG/sustainability assurance team.

(a) Given the multi-disciplinary nature of the GHG/sustainability assurance team, did you feel there was sufficient discussion time in the early stages of the engagement to share diverse information and perspectives?

There was not enough time	There was sufficient time	There was more than enough time

|______|______|______|______|______|______|______|______|

(b) Given the multi-disciplinary nature of the GHG/sustainability assurance team, did you feel there was sufficient elaboration and integration of different information and perspectives from different team members?

There was not enough	It was sufficient	There was more than enough

|______|______|______|______|______|______|______|______|

The following three questions relate to the evidence gathering and evaluation stages of the assurance engagement

(c) To what extent were the following assurance procedures used to gather evidence?

Tests of controls	_______%
Detailed substantive testing	_______%;
Analytical procedures	_______%;
Total	100 %

If other, please specify____

(d) To what extent was there a clear separation between the information search/collection stage and the information processing/decision making stage?

There was no separation	There was a partial separation	There was a clear separation

|______|______|______|______|______|______|______|______|

(e) To what extent did the GHG/sustainability assurance team discuss the information collected **before** final evaluations and decisions were made?

Our team did not do this	Our team partially did this	Our team did this

|______|______|______|______|______|______|______|______|

(f) Was there any other mechanism by which different information and perspectives from different team members was shared and integrated **at any stage** of the engagement? If so, please specify:

____

____

(g) In your opinion, what factors, if they had been present, would have made the team work together better? These may be some of the factors listed above or others.

____

____

____

____

(2I) Final comments

Please comment on any issues not properly covered in the above questions, or anything else you wish to raise regarding the team for this GHG/sustainability assurance engagement.

____

____

____

____

PTO

PART 3:

(3) Demographic details:

(a) What is your designated title within in your firm?___________________________

(b) What is your tertiary education background?
Accounting/Auditing______; Engineering_______; Environmental Science ______;
Law _____; Other (please specify) _______________________________________

(c) How long have you been involved in conducting environmental/GHG/sustainability assurance engagements? ____

(d) Approximately how many GHG/sustainability assurance engagements have you undertaken, either individually or as part of a team? __________________________

(e) For how many of these GHG/sustainability assurance engagements were you the assurance team leader?____

(f) Have you attended training courses on assurance for GHG emissions?
Yes_____; No___________.
If yes, how many hours of training courses?______________hours.

Thank you for your time. If you would like a summary of the results of this research please inform the KSAM training coordinator.

APPENDIX 2
Study One Definitions of Variables

Category	Variable	Definition
Variables of interest:		
Team process	*DISCUSS*	Rating of the perceived sufficiency of discussion time in the early stages of the engagement to share diverse information, measured on a nine-point scale;
	ELABORATE	Rating of the perceived sufficiency of elaboration and integration of different information from different team members, measured on a nine-point scale;
Team composition	*TEAMSIZE*	Number of members in the team;
	DIVERSITY[a]	Level of educational diversity in the team;
Control variables:		
Task characteristics	*DIRECT*	Proportion of direct measurement the client used in quantifying their GHG data;
	TASK	Dummy variable =1 if the engagement is a reasonable level of assurance and 0 otherwise;
Client characteristics and risks	*SIZE*	Dummy variable =1 if the client is large (have more than five facilities) and 0 otherwise;
	COMPLEX	Rating of the complexity of client's GHG emissions profile compared to other similar GHG assurance engagements, measured on a nine-point scale;
	PUBLIC	Dummy variable =1 if the client is a public company and 0 otherwise;
	AVGIC[b]	Average of the five internal control quality ratings, measured on a nine-point scale;
Client-assurer relationship	*PREVIOUS*	Number of previous GHG assurance engagements undertaken for the client;
	IMPORTANCE	Rating of the client's relative importance to the respondent's assurance firm, measured on a nine-point scale;
Demographic variables:		
	GHGYEAR	Number of years involved in conducting GHG/sustainability assurance;
	TRAINING	Number of training hours on GHG assurance.
Additional/alternative variables:		
	MIXBG	Dummy variable = 1 if teams have members with mixed backgrounds and 0 otherwise;
	REG	Dummy variable = 1 if the client undertook GHG assurance on a regulatory basis and 0 otherwise.

[a] Educational diversity is measured by using the modified-Blau's (1977) index of heterogeity, $1 - \Sigma(Pi)^2$, where P*i* is the proportion of a team's members in the i^{th} category (accounting, others, and mixed) weighted by the level of involvement of each member (low = 1, medium = 2, and high = 3).

[b] As the five ratings of internal control quality were strongly correlated with each other (Spearman's correlation coefficient in the range of 0.462 to 0.744, all p=0.000), it is considered to be more appropriate to combine them together as one variable. The factor analysis has been conducted on these five ratings, and has yielded one factor. To combine these five quality ratings, the individual scores of each rating were multiplied by its factor loading (factor loadings = 0.109, 0.092, 0.296, 0.296, and 0.208, respectively). Since all factor loadings were scaled by 1, the five ratings can then be combined into AVGIC.

APPENDIX 3
Study Two Case Material

Audit Planning Documents
SteelCo

1. Understanding the Entity and its environment

Industry	Electric Arc Furnace – steel making
Background	**SteelCo** manufacture steel sections for use in the construction industry. The company's operation consists of the following sequential steps; melting scrap steel (and some iron) in an electric arc furnace, adjusting the steel chemistry to exact specifications in a ladle furnace and then casting into basic billets. The billets are then transferred to a hot rolling mill where they are rolled through a series of mill stands to obtain the correct length and cross section as required by the customer orders. **<u>Electric Arc furnace</u>** An **Electric Arc Furnace (EAF)** is used to produce new steel from scrap metal. This method can be lower cost than the traditional blast furnace method of making steel, and it conserves raw materials like iron ore, coke and fluxes. The EAF is a large circular steel shell lined with a refractory material. A charge for a typical melt would consist of 86 percent scrap steel, and 14 percent iron. Power is supplied to the furnace through the **electrodes**. The electrodes are placed in the furnace and when the power is applied it produces an arc of electricity from the electrode to the scrap steel. The energy from the arc raises the temperature to 1600°C, melting the scrap. One of the advantages of the EAF technology is that it can use scrap material or variable quality and chemical composition as a feedstock. The operators are able to closely monitor the chemical composition of the melt and adjust it to the required quality through the addition of carbon, nickel, chrome and other elements. The EAF process also uses oxygen, hydrogen, nitrogen and fluxes to control the temperature, to remove small amounts of impurity and prevent oxidation from the air. After about 80 minutes, the molten steel is tapped into a ladle and transferred to the **ladle furnace**. The purpose of the ladle

	furnace is to allow for a more exact control of the chemical composition prior to casting. The main source of scrap steel is from cars, but washing machines, fridges, bicycles and steel from demolished buildings can also be recycled using the EAF. **<u>Casting</u>** The liquid steel is then cast into steel billets which are then ready to be used in the rolling mills. **<u>Hot Rolling</u>** Hot rolling is the main method to shape steel into different products for the construction industry. At the rolling mills at SteelCo, the steel billets produced in the caster are re-heated in a furnace to about 1200^{o}C and allowed to "soak" at that temperature to ensure a uniform heating. The re-heated billet is then drawn from the furnace and passed through a series of mill stands. A mill stand consists of two rolls revolving at the same speed but in opposite directions. The gap between the rolls is smaller than the steel being rolled, so that the steel is reduced in thickness and at the same time lengthened to form the desired end products. **<u>Sources of Emissions and energy use</u>** The major source of emissions for SteelCo is from **electricity**, due to the high energy used in the EAF. **Natural gas** is the main fuel used in the rolling mills. Often the EAF and the rolling mill are on the same site, but in NSW the rolling mill is located offsite. Gas includes energy used by on-site contractors, who use natural gas for welding and forklifts. Within the steel making process, emissions also arise from the chemical reaction (industrial process emissions). The natural gas pipeline is actually shared with GlassCo, a plant next door who make glass products. This is due to legacy issues, and so the invoiced natural gas amount is actually reduced by the natural gas supplied to GlassCo to get the natural gas consumed by SteelCo. Process emissions from the production of steel also occur and are calculated using a **"Carbon Mass Balance".** The Carbon mass balance method effectively measures the carbon content of all inputs (i.e., coke) into the process, less the carbon content of all products and waste, with the balance being the carbon emissions from the process.

	Currently SteelCo uses **Method 1** in the measurement of their carbon emissions from electricity and gas (energy emissions), and **Method 2** for the carbon mass balance emissions instead of the default factors in the measurement determination (please see page 6 for the definitions of Method 1 and 2). **<u>Boundary</u>** The recycled scrap metal is shredded by RecycleCo, who operate at a site next door to SteelCo. SteelCo own the land and equipment, however, RecycleCo manage the site, and are the employer of the workers on the site.
Emissions intensive Trade Exposed	Steel making is an Emissions-Intensive Trade Exposed (EITE) process. The activity definition of the manufacture of carbon steel from cold ferrous feed under the EITE assistance program includes the stages of the heating and melting (generally achieved using electricity) of a cold ferrous feed, such as ferrous scrap and pig iron, into liquid steel and the subsequent casting of liquid steel into solid carbon steel products.
Reporting period	The reporting period runs from July 1st to June 30th of each year.
Level of assurance	Reasonable assurance is required

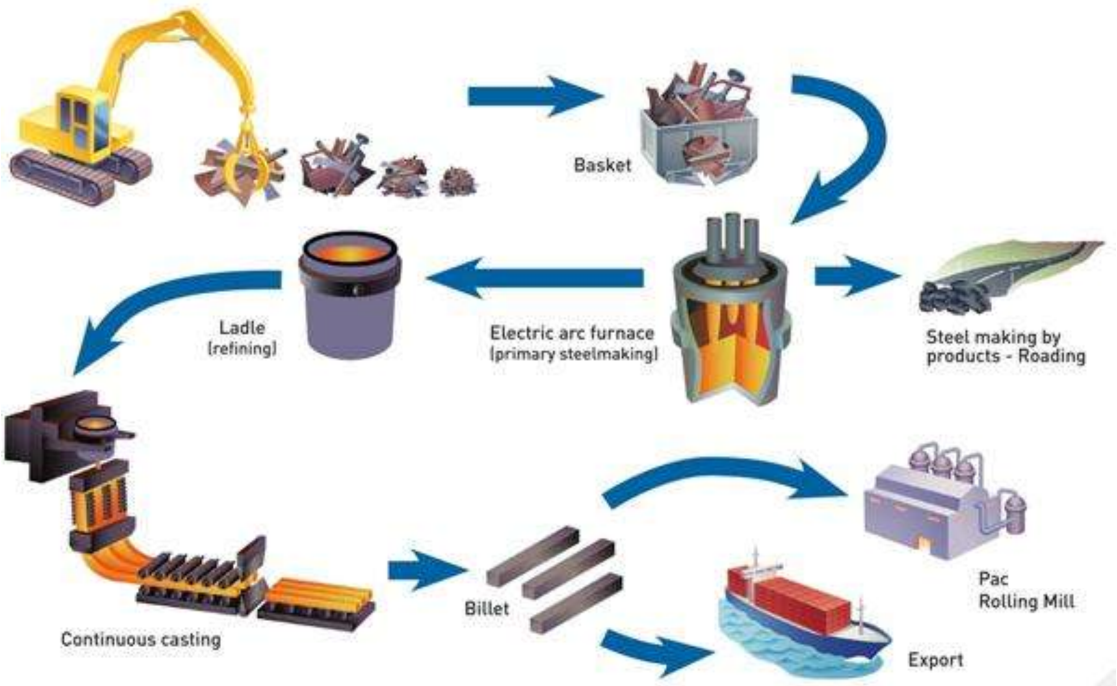


Figure 1: EAF steel making process
(Source: www.pacificsteel.co.nz/process)

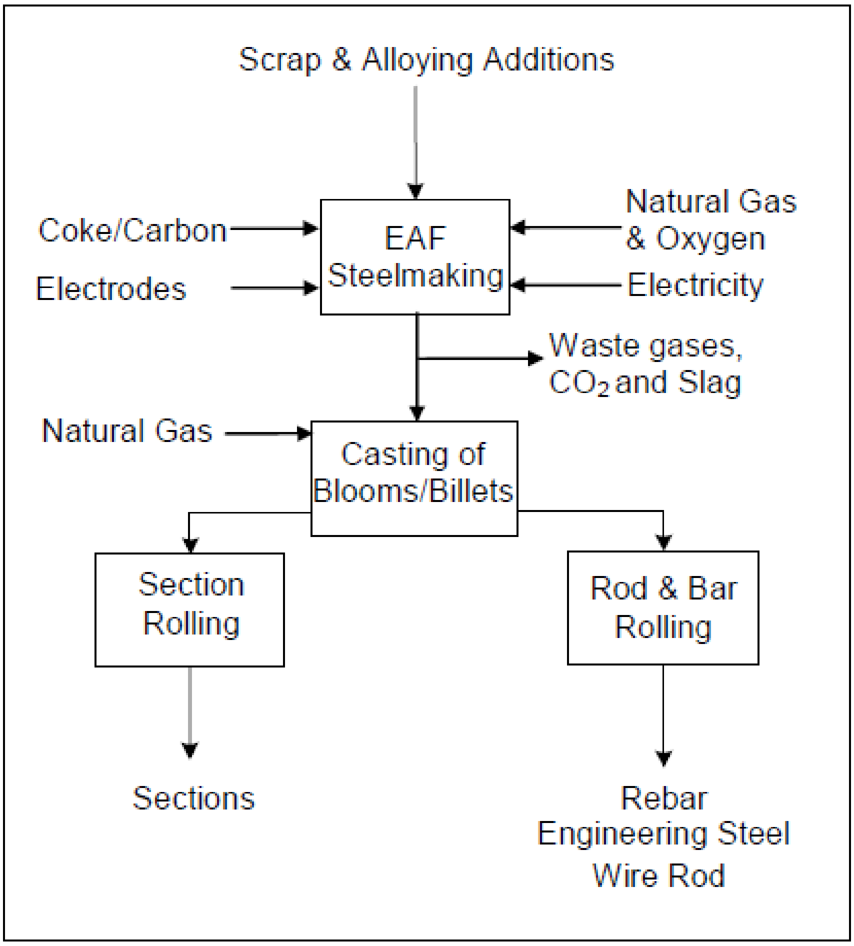


Figure 2: EAF Process Flow Overview Diagram

2. Entity's emissions data - prepared by the engagement team

Site	Activity	State	Production		Electricity						Natural Gas				Total Energy	
			Current Year	Prior Year		Current Year			Current Year	Prior Year		Current Year		Prior Year	Current Year	Prior Year
			tonnes	tonnes	Method	MWh	$	conversion	GJ	GJ	Method	$	GJ	GJ	GJ	GJ
1	EAF and rolling mill	Tasmania	324,890	353,805	1	176,565	$7,945,425	3.60	635,634	699,769	1	2,029,014	362,324	394,933	997,958	1,094,703
2	Rolling Mill	NSW	241,787	360,021	1	26,454	$1,190,430	3.60	95,234	127,509	1	1,799,521	321,343	507,368	416,577	634,878
3	EAF	NSW	249,321	160,313	1	189,789	$8,540,505	3.60	683,240	477,517	1	866,186	154,676	135,960	837,916	613,477
4	EAF and rolling mill	Victoria	532,567	262,769	1	298,456	$13,430,520	3.60	1,074,442	609,853	1	3,678,007	656,787	432,823	1,731,229	1,042,676

Site	Electricity related emissions				Natural Gas related emissions					Total energy emissions		Industrial process emissions			Total emissions	
	Emission	Current Year		Prior Year	Emission	Current Year		Current Year	Prior Year	Current Year	Prior Year		Current Year	Prior Year	Current Year	Prior Year
	Factor	MWh	tonnes CO2	tonnes CO2	Factor	GJ	conversion	tonnes CO2	tonnes CO2	tonnes CO2	tonnes CO2	Method	tonnes CO2	tonnes CO2	tonnes CO2	tonnes CO2
1	0.30	176,565	52,970	58,314	51.33	362,324	0.001	18,598	20,272	71,568	78,586	2	17,342	19,231	88,910	97,817
2	0.89	26,454	23,544	31,523	51.33	321,343	0.001	16,495	26,043	40,039	57,566	2	-	-	40,039	57,566
3	0.89	189,789	168,912	118,053	51.33	154,676	0.001	7,940	6,979	176,852	125,032	2	16,832	12,355	193,684	137,386
4	1.21	298,456	361,132	204,978	51.33	656,787	0.001	33,713	22,217	394,845	227,195	2	37,112	19,001	431,957	246,196

3. Entity's emissions data – Assumptions and Ratio analysis

Assumptions	
Electricity cost per mwh	$45
natural gas kgCO2/GJ	51.33
electricity GJ/MWh	3.6
Cost of gas / GJ	$5.60

State, Territory or grid description	**Emission factor (kg CO2-e/kWh)**
New South Wales and Australian Capital Territory	0.89
Victoria	1.21
Queensland	0.88
South Australia	0.68
South West Interconnected System in Western Australia	0.8
Tasmania	0.3
Northern Territory	0.67

Ratio analysis	**Current Year**			**Prior Year**		
	mwh/ tonnes	**gas/ tonnes**	**total tonnes CO2/ tonnes production**	**mwh/ tonnes**	**gas/ tonnes**	**total tonnes CO2/ tonnes production**
Site 1	0.543	1.115	0.274	0.549	1.116	0.276
Site 2	0.109	1.329	0.166	0.098	1.409	0.160
Site 3	0.761	0.620	0.777	0.827	0.848	0.857
Site 4	0.560	1.233	0.811	0.645	1.647	0.937

<u>Definitions</u>

Method 1: the National Greenhouse Accounts default method

This method specifies the use of designated emission factors in the estimation of emissions. These emission factors are national average factors determined by the Department of Climate Change and Energy Efficiency using the Australian Greenhouse Emissions Information System.

Method 2: a facility-specific method using industry sampling and Australian or international standards listed in the Determination or equivalent for equivalent for analysis of fuels and raw mater

Method 2 enables corporations to undertake additional measurements in order to aim more accurate estimates for emissions for that particular facility. This method is likely to be most useful for fuels which exhibit some variability in key qualities, such as carbon content, from source to source.

Under Method 2, representative and unbiased samples of fuels consumed must be obtained for analysis. Analysis of the fuels for carbon, energy, ash or moisture content must be done in accordance with listed Australian or international documentary standards or equivalent.

(**<u>Source:</u>** Technical Guidelines for the estimation of greenhouse gas emissions by facilities in Australia, National greenhouse gas and energy reporting system measurement, *Department of Climate Change and Energy Efficiency,* June 2010.)

APPENDIX 4
Study Two Research Instruments

PARTICIPANT INFORMATION STATEMENT AND CONSENT FORM

Effects of Types of Team Process on the Performance of Multidisciplinary Greenhouse Gas Assurance Teams

Participant selection and purpose of study

You are invited to participate in a study of multidisciplinary team in greenhouse gas (GHG) assurance engagement. We hope to learn which type of team processes may be most beneficial in assessing risks of material misstatement in the GHG assurance engagement. You were selected as a possible participant in this study because your knowledge and work experience as an assurance or subject-matter expert.

Description of study and risks

If you decide to participate, we will require you to read a GHG assurance case and make a number of judgments based on that case. The experiment will last approximately 60 minutes. We do not anticipate any risks to you from participating in this study, other than those you encounter in day-to-day life.

Confidentiality and disclosure of information

Any information that is obtained in connection with this study and that can be identified with you will remain confidential and will be disclosed only with your permission, except as required by law. If you give us your permission by participating in the study, we plan to publish the results at accounting conferences and accounting journals. In any publication/conference presentation, information will be provided in such a way that you cannot be identified.

Recompose to participants

As a token of appreciation, you will be given a $50 gift voucher for your participation in the study.

Complaints may be directed to:

Ethics Secretariat	Phone: +61 2 9385 4234
The University of New South Wales	Fax: +61 2 9385 6648
Sydney 2052 AUSTRALIA	Email: ethics.sec@unsw.edu.au).

Any complaint you make will be investigated promptly and you will be informed out the outcome.

Feedback to participants

You will have the option of getting combined feedback on the results of the study but no individual feedback will be given. Please advise the experiment administrator if you wish to receive a copy of the results.

Your consent

Your decision whether or not to participate will not prejudice your future relations with the University of New South Wales and your company. Your subsequent participation is indication that you consent to participating in our study. If you decide to participate, you are free to withdraw your consent and to discontinue participation at any time without prejudice.

If you have any questions, please feel free to contact us (details below) and we will be happy to answer them.

You will be given a copy of this form to keep.

APPENDIX 4.1
Study Two Research Instruments
(Computer Screens)

Nominal Team Treatment

Some definitions before you begin...

Assurance expert: A financial auditor, typically with an accounting background.

Subject-matter expert: A person who has technical knowledge of GHG measurement processes, typically with an engineering or environmental science background.

Risks of potential material misstatements: The risks that the entity's GHG statements will contain a material misstatement whether caused by error or fraud (i.e. what could go wrong).

Next

THE CASE

Assume that you are a member of a two-person team comprising one assurance and one subject-matter expert which has been asked to provide input to the risks assessment as part of the planning process for the GHG assurance engagement of SteelCo.

An abbreviated version of SteelCo's planning document is provided to you here. Please spend 10 minutes on reading the case material and becoming familiar with the company and its environment.

Please do not start reading until you have been instructed to do so.

Read Paper Version of Case Materials Then Press


UNSW
Australian School of Business
CO_2
EMISSION
PART A
Please read this page carefully.
You have been asked to take 20 minutes to record a list of risks of potential material misstatements (what could go wrong) for SteelCo this year. You will then be asked to take 3 minutes to rank the top risks in order of their significance, and 12 minutes to identify the appropriate procedures for each of your top risks. Please note that your other team member will not be present at this time.
Some participants' answers (without identification to any individual) will be selected for review by other members of your organisation.
It is important that you work independently, concentrate for the full 35 minutes, and document as many risks of potential material misstatements as you can in the first 20 minutes. Do not leave the recording to the end of the time period, i.e., record each idea as it is considered.
YOU CAN REFER TO THE CASE WHILE ANSWERING THE FOLLOWING QUESTIONS
Next


UNSW
Australian School of Business
CO_2
EMISSION
Participant No. ss099
PART A - Task 1
Please document the risks of potential material misstatements that may occur for SteelCo this year.
You have 20 minutes to complete this task
Next

Participant No: ss099

Required: Please enter the risks of potential material misstatements in the boxes provided (please be as specific as possible).

Risk number	Risks of potential material misstatements i.e. what could go wrong Please be as specific as possible
1	
2	
3	
4	
5	

Risk number	Risks of potential material misstatements i.e. what could go wrong Please be as specific as possible
6	
7	
8	
9	
10	
11	
12	
13	

Risk number	Risks of potential material misstatements i.e. what could go wrong Please be as specific as possible
14	
15	
16	
17	
18	
19	
20	

Next

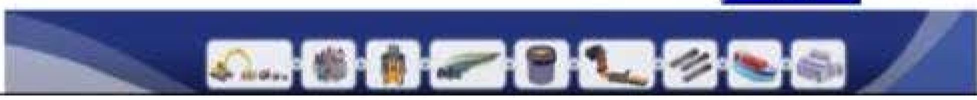

UNSW
Australian
School of
Business
CO_2
EMISSION
Participant No: ss099
PART A - Task 2
Please rank your 'top four' risks of potential material misstatements in order of their significance
(significant risks are the risks of material misstatements that, in your opinion, require special audit consideration)
You have 3 minutes to complete this task
Next

Participant No.: ss099

Required: Please enter a number 1, 2, 3 or 4 in the provided boxes to identify the top 'four' risk (1 = most significant, to 4 = least significant).

	Risks of potential material misstatements	Top 'four' risks (1 to 4)
1	test A	1
2	test B	2
3	test C	3
4	test D	4
5	test E	
6	test F	
7	test G	
8	test H	

Next



Participant No: ss099

PART A - Task 3

Please identify the appropriate procedures to address your 'top four' risks that you have ranked in Task 2.

You have 12 minutes to complete this task

Next

Participant No. ss099

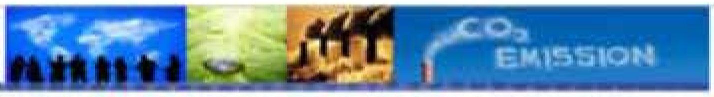
Required: Please enter the procedures for each risk in the boxes provided
(please be as specific as possible).

1	Risks of potential material misstatements
	test A
	Appropriate procedures to address the assessed risks of material misstatements Please be as specific as possible
	procedure A
2	Risks of potential material misstatements
	test B
	Appropriate procedures to address the assessed risks of material misstatements
	Procedure B
3	Risks of potential material misstatements
	test C
	Appropriate procedures to address the assessed risks of material misstatements Please be as specific as possible
	Procedure C
4	Risks of potential material misstatements
	test D
	Appropriate procedures to address the assessed risks of material misstatements Please be as specific as possible
	Procedure D

Next

UNSW
Australian School of Business
CO2
EMISSION
PART B
You have 5 minutes to complete this part
Next

UNSW | Australian School of Business

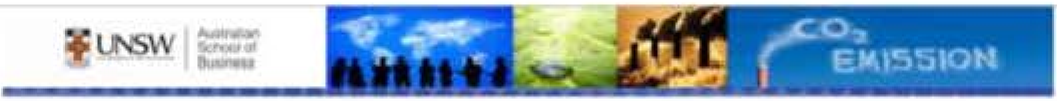


It is important that you complete the remainder of this task as an individual, i.e., please do not discuss this with the others.

1. In Part A, you were asked to document risks and procedures for SteelCo. Please indicate the level you would rank yourself in terms of :

	Low			Medium			High
1.1 Ability to identify risks of material misstatements	1	2	3	4	5	6	7
1.2 Ability to Identify appropriate procedures to address the risks	1	2	3	4	5	6	7
1.3 Knowledge of the subject matter in environmental reports	1	2	3	4	5	6	7
1.4 Knowledge of the audit criteria and processes	1	2	3	4	5	6	7

Next



It is important that you complete the remainder of this task as an individual, i.e., please do not discuss this with the others.

1. In Part A, you were asked to document risks and procedures for SteelCo. Please indicate the level you would rank yourself in terms of :

	Low			Medium			High
1.1 Ability to identify risks of material misstatements	1 ○	2 ○	3 ○	4 ○	5 ○	6 ●	7 ○
1.2 Ability to identify appropriate procedures to address the risks	1 ○	2 ○	3 ○	4 ●	5 ○	6 ○	7 ○
1.3 Knowledge of the subject matter in environmental reports	1 ○	2 ○	3 ○	4 ●	5 ○	6 ○	7 ○
1.4 Knowledge of the audit criteria and processess	1 ○	2 ○	3 ○	4 ○	5 ●	6 ○	7 ○

Next


UNSW Australian School of Business
CO2 EMISSION
Participant No.: ss099
PART C
You have 5 minutes to complete this part
Next

It is important that you complete the remainder of this task as an individual, i.e., please do not discuss this with others.

1. What is your position in the firm?(e.g. senior auditor, sustainability manager)

2. What is your tertiary education background?

- ☐ Accounting/Auditing
- ☐ Engineering
- ☐ Environmental Science
- ☐ Other (please specify)

3. What is your industry specialisation?

4. How many years of financial audit experience do you have? (if any)

years

5. How long have you been involved in conducting environmental/GHG/sustainability assurance engagements?

years

6. Approximately how many environmental/GHG/sustainability assurance engagements have you undertaken?

engagements

7. Have you been involved in conducting financial/GHG/sustainability assurance engagements for clients in manufacturing?

- ○ Yes, I have undertaken engagements
- ○ No

8. How much training on assurance for GHG emissions have you done?

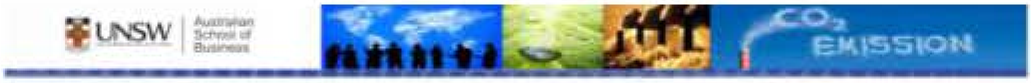
hours, or

days

****

THANK YOU FOR YOUR PARTICIPATION

Next

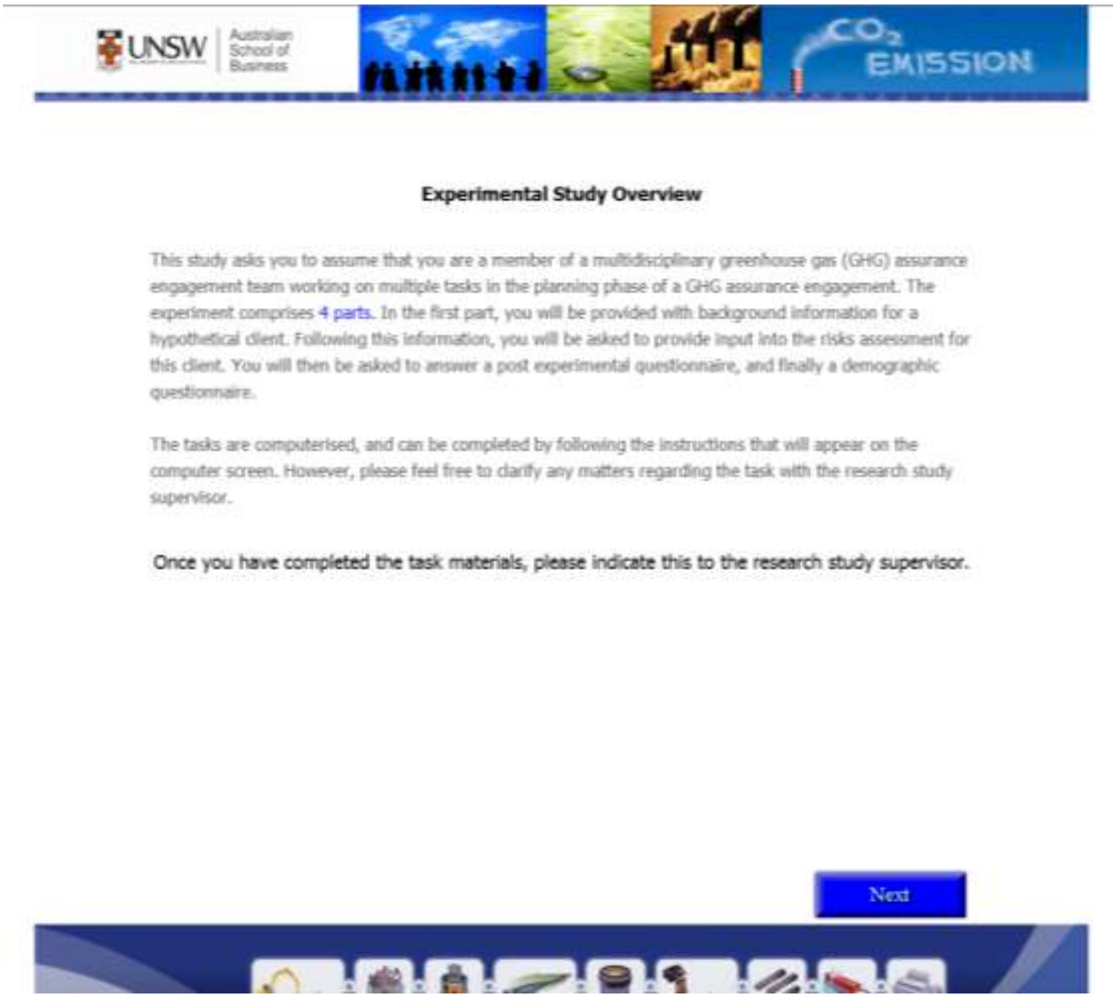


Participant No. 88099

THANK YOU FOR YOUR PARTICIPATION

APPENDIX 4.2
Study Two Research Instruments
(Computer Screens)

Interacting Team Treatment


UNSW | Australian School of Business

CO_2 EMISSION

Experimental Study Overview

This study asks you to assume that you are a member of a multidisciplinary greenhouse gas (GHG) assurance engagement team working on multiple tasks in the planning phase of a GHG assurance engagement. The experiment comprises 4 parts. In the first part, you will be provided with background information for a hypothetical client. Following this information, you will be asked to provide input into the risks assessment for this client. You will then be asked to answer a post experimental questionnaire, and finally a demographic questionnaire.

The tasks are computerised, and can be completed by following the instructions that will appear on the computer screen. However, please feel free to clarify any matters regarding the task with the research study supervisor.

Once you have completed the task materials, please indicate this to the research study supervisor.

Next

UNSW Australian School of Business

CO2 EMISSION

Some definitions before you begin...

Assurance expert: A financial auditor, typically with an accounting background.

Subject-matter expert: A person who has technical knowledge of GHG measurement processes, typically with an engineering or environmental science background.

Risks of potential material misstatements: The risks that the entity's GHG statements will contain a material misstatement whether caused by error or fraud (i.e. what could go wrong).

Next

UNSW Australian School of Business

CO2 EMISSION

THE CASE

Assume that you are a member of a two-person team comprising one assurance and one subject-matter expert which has been asked to provide input to the risks assessment as part of the planning process for the GHG assurance engagement of SteelCo.

An abbreviated version of SteelCo's planning document is provided to you here. Please spend 10 minutes on reading the case material and becoming familiar with the company and its environment.

Please do not start reading until you have been instructed to do so.

Read Paper Version of Case Materials Then Press

Participant No: az001

PART A

Please read this page carefully.

You and another team member have been asked to take 20 minutes to discuss and record a list of risks of potential material misstatements (what could go wrong) for SteelCo this year. Your team will then be asked to take 3 minutes to rank the top risks in order of their significance, and 12 minutes to identify the appropriate procedures for each of your top risks.

Some participants' answers (without identification to any individual) will be selected for review by other members of your organisation.

It is important that you work as a team, concentrate for the full 35 minutes, and document as many risks of potential material misstatements as you can in the first 20 minutes. Do not leave the recording to the end of the time period, i.e., record each idea as it is considered.

YOU CAN REFER TO THE CASE WHILE ANSWERING THE FOLLOWING QUESTIONS

Next

Participant No: az001

This Part should be completed by your team

The person allocated with the participant number xx001 should record

If you are the person allocated with participant number xx002,
please move to your team member's computer to do Part A as a team.

Next

Participant No: az001

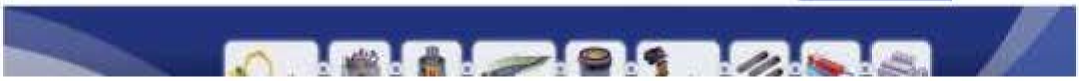
PART A - Task 1

Once your team member has joined you, as a team please document the risks of potential material misstatements that may occur for SteelCo this year.

The person allocated with the participant number xx001 should record

You have 20 minutes to complete this task

Next



Participant No: az001

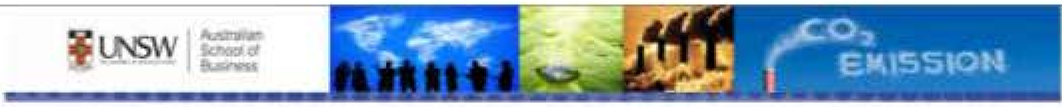
Required: Please enter the risks of potential material misstatements in the boxes provided (please be as specific as possible).

Risk number	Risks of potential material misstatements i.e. what could go wrong Please be as specific as possible
1	
2	
3	
4	
5	

Risk number	Risks of potential material misstatements i.e. what could go wrong Please be as specific as possible
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	

Risk number	Risks of potential material misstatements i.e. what could go wrong Please be as specific as possible
16	
17	
18	
19	
20	

Next



Participant No: az001

PART A- Task 2

As a team please rank your team's 'top four' risks of potential material misstatements in order of their significance

(significant risks are the risks of material misstatements that, in your opinion, require special audit consideration)

The person allocated with the participant number xx001 should record

You have 3 minutes to complete this task

Next

Participant No: az001

Required: Please enter a number 1, 2, 3 or 4 in the provided boxes to identify the top 'four' risk (1 = most significant, to 4 = least significant).

	Risks of potential material misstatements	Top 'four' risks (1 to 4)
1	Risk A	
2	Risk B	
3	Risk C	
4	Risk D	
5	Risk E	
6	Risk F	

Next

UNSW Australian School of Business

CO2 EMISSION

Participant No.: az001

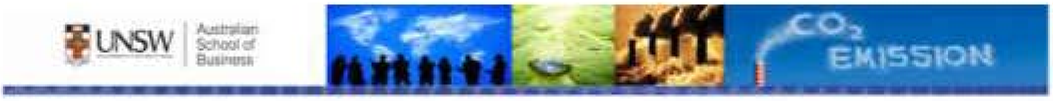
PART A- Task 3

As a team Please identify the appropriate procedures to address your 'top four' risks that you and your team member have ranked in Task 2.

The person allocated with the participant number xx001 should record

You have 12 minutes to complete this task

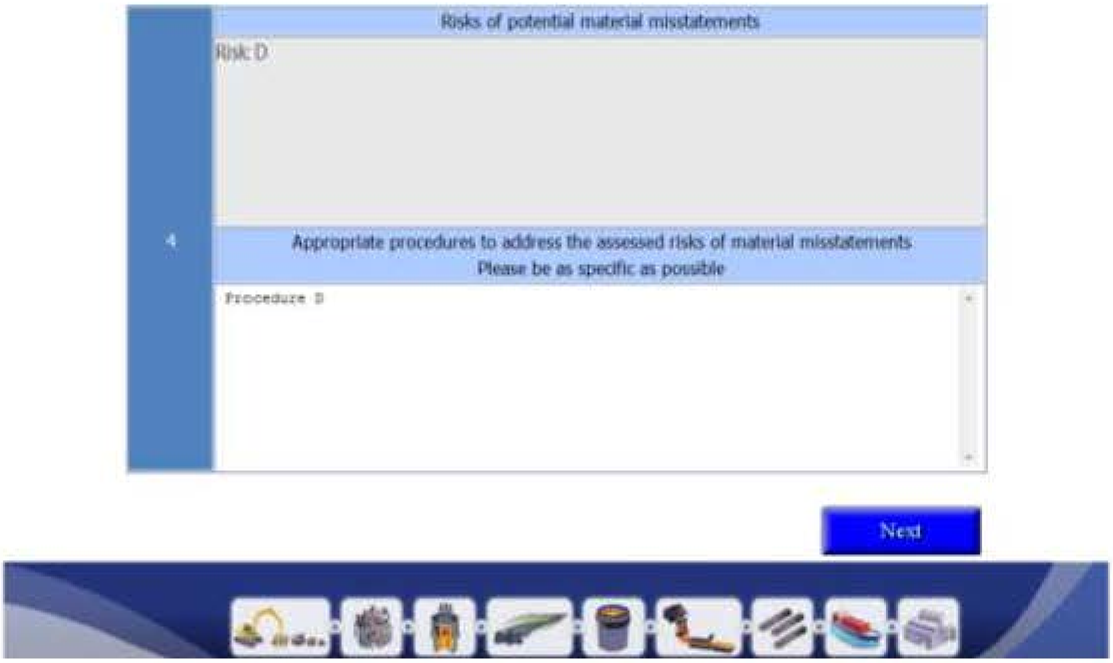
Next



Participant No: az001

<u>Required:</u> Please enter the procedures for each risk in the boxes provided (<u>please be as specific as possible</u>).

1	Risks of potential material misstatements
	Risk A
	Appropriate procedures to address the assessed risks of material misstatements Please be as specific as possible
	Procedure A
2	Risks of potential material misstatements
	Risk B
	Appropriate procedures to address the assessed risks of material misstatements Please be as specific as possible
	Procedure B
3	Risks of potential material misstatements
	Risk C
	Appropriate procedures to address the assessed risks of material misstatements Please be as specific as possible
	Procedure C



4	Risks of potential material misstatements
	Risk D
	Appropriate procedures to address the assessed risks of material misstatements Please be as specific as possible
	Procedure D

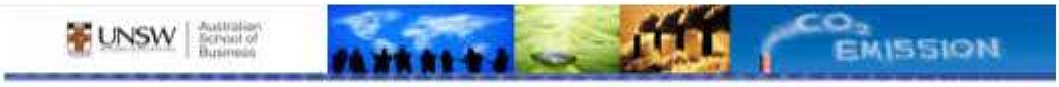
Next

Participant No: az001

The next part should be completed by each individual

If you are the person allocated with participant number xx002,
please move back to your computer to do Part B.

Next



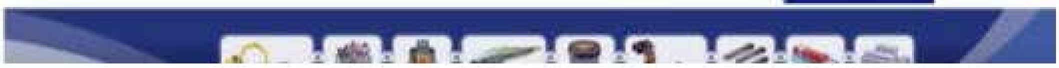
Participant No.: az001

PART B

This part should be completed by <u>each individual</u>

You have 5 minutes to complete this part

Next



Participant No.: az001

It is important that you complete the remainder of this task as an individual, i.e., please <u>do not discuss this with the others.</u>

1. In Part A, you and your team member were asked to discuss SteelCo. Please indicate how you would rank yourself and your team member in terms of :

	you								your team member						
	Low			Medium			High		Low			Medium			High
1.1 Ability to identify risks of material misstatements	1	2	3	4	5	6	7		1	2	3	4	5	6	7
1.2 Ability to Identify appropriate procedures to address the risks	1	2	3	4	5	6	7		1	2	3	4	5	6	7
1.3 Knowledge of the subject matter in environmental reports	1	2	3	4	5	6	7		1	2	3	4	5	6	7
1.4 Knowledge of audit criteria and processes	1	2	3	4	5	6	7		1	2	3	4	5	6	7

Next

Participant No: az001

2. On a scale of 1 to 7, with 1 indicating poor performance and 7 indicating excellent performance, how would you rate the overall performance of your team member in this study?

1	2	3	4	5	6	7
Poor performance			Moderate performance			Excellent performance

3. In Part A, you were asked to discuss SteelCo with your team member. Please indicate to what extent you agree with the statements listed.

To what extent you agree that:	Completely Disagree			Neither agree or disagree			Completely Agree
3.1 Your team member focused on assessing key numerical indicators or quantitative data	1	2	3	4	5	6	7
3.2 Your team member focused on assessing qualitative data or measures	1	2	3	4	5	6	7
3.3 Your team member places emphasis on assessing data accuracy	1	2	3	4	5	6	7
3.4 Your team member places emphasis on assessing data completeness	1	2	3	4	5	6	7

4. To what extent did your team member live up to your expectations for an expert in their area?

1	2	3	4	5	6	7
Did not live up to expectations at all			Moderately lived up to expectations			Fully lived up to expectations

Next

Participant No: az001

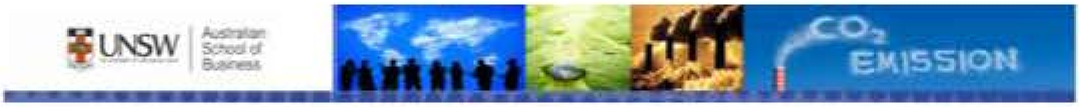
5. In Part A, you and your team member were asked to discuss SteelCo. Please indicate to what extent you agree that:

To what extent you agree that:	Completely Disagree		Neither agree or disagree				Completely Agree
5.1 Your team member contributed a lot of information during the team task	1	2	3	4	5	6	7
5.2 Your team member contributed unique information during the group task	1	2	3	4	5	6	7
5.3 During the task, you and your team member tried to use all available information	1	2	3	4	5	6	7

6. In Part A, you and your team member were asked to discuss SteelCo. Please indicate to what extent you think that:

To what extent you think that:	Almost Never		Sometimes				Almost Always
6.1 Your team member chooses concepts and words that you understand	1	2	3	4	5	6	7
6.2 Your team member tailors communications to refer to concepts, terms, and perspectives that you both have in common	1	2	3	4	5	6	7
6.3 Your team member makes arguments that are technically, politically, or otherwise unacceptable to you	1	2	3	4	5	6	7
6.4 Your team member inquires about the reasons underlying your knowledge, beliefs, or preferences	1	2	3	4	5	6	7
6.5 Your team member often asks for clarification or elaboration on issues related to your knowledge, beliefs, or preferences	1	2	3	4	5	6	7

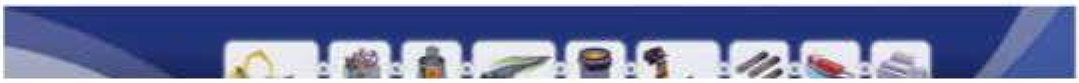
Next



Participant No: az001

To what extent:	Almost Never			Sometimes			Almost Always
6.6 Your team member prompts you to surface and discuss what you know, believe, or prefer	1	2	3	4	5	6	7
6.7 Your team member helps you to better understand the team's task or task situation	1	2	3	4	5	6	7
6.8 Your team member seems to anticipate what you will do or say	1	2	3	4	5	6	7
6.9 Your team member does a good job coordinating his/her actions with yours	1	2	3	4	5	6	7
6.10 Your team member seems to recognize when yours and your team member's knowledge, beliefs and preferences differ	1	2	3	4	5	6	7

Next



PART C

This part should be completed by <u>each individual</u>

You have 5 minutes to complete this part

Next

Participant No. az001

It is important that you complete the remainder of this task as an individual, i.e., please <u>do not discuss this with others.</u>

1. What is your position in the firm?(e.g. senior auditor, sustainability manager)

2. What is your tertiary education background?

- ☐ Accounting/Auditing
- ☐ Engineering
- ☐ Environmental Science
- ☐ Other (please specify)

3. What is your industry specialisation?

4. How many years of financial audit experience do you have? (if any)

years

5. How long have you been involved in conducting environmental/GHG/sustainability assurance engagements?

years

6. Approximately how many environmental/GHG/sustainability assurance engagements have you undertaken?

engagements

7. Have you been involved in conducting financial/GHG/sustainability assurance engagements for clients in manufacturing?

- ○ Yes, I have undertaken engagements
- ○ No

8. How much training on assurance for GHG emissions have you done?

[ ] hours, or

[ ] days

9. How familiar were you with your assigned team member for this experiment before performing the team task?

1	2	3	4	5	6	7
○	○	○	○	○	○	○
Not familiar at all			Moderately familiar			Very familiar

****

THANK YOU FOR YOUR PARTICIPATION

Next

Participant No: az001

THANK YOU FOR YOUR PARTICIPATION

APPENDIX 4.3
Study Two Research Instruments
(Computer Screens)

Review Team Treatment

Experimental Study Overview

This study asks you to assume that you are a member of a multidisciplinary greenhouse gas (GHG) assurance engagement team working on multiple tasks in the planning phase of a GHG assurance engagement. The experiment comprises 4 parts. In the first part, you will be provided with background information for a hypothetical client. Following this information, you will be asked to provide input into the risks assessment for this client. You will then be asked to answer a post experimental questionnaire, and finally a demographic questionnaire.

The tasks are computerised, and can be completed by following the instructions that will appear on the computer screen. However, please feel free to clarify any matters regarding the task with the research study supervisor.

Once you have completed the task materials, please indicate this to the research study supervisor.

Next

Some definitions before you begin...

Assurance expert: A financial auditor, typically with an accounting background.

Subject-matter expert: A person who has technical knowledge of GHG measurement processes, typically with an engineering or environmental science background.

Risks of potential material misstatements: The risks that the entity's GHG statements will contain a material misstatement whether caused by error or fraud (i.e. what could go wrong).

Next

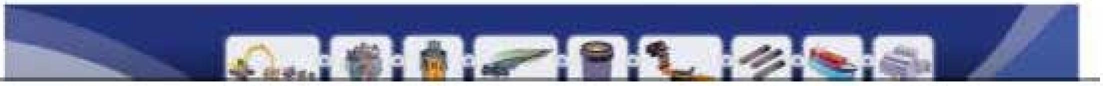
THE CASE

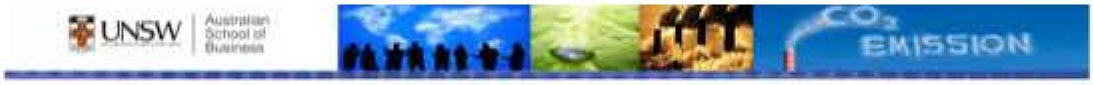
Assume that you are a member of a two-person team comprising one assurance and one subject-matter expert which has been asked to provide input to the risks assessment as part of the planning process for the GHG assurance engagement of SteelCo.

An abbreviated version of SteelCo's planning document is provided to you here. Please spend 10 minutes on reading the case material and becoming familiar with the company and its environment.

Please do not start reading until you have been instructed to do so.

Read Paper Version of Case Materials Then Press





PART A

Please read this page carefully.

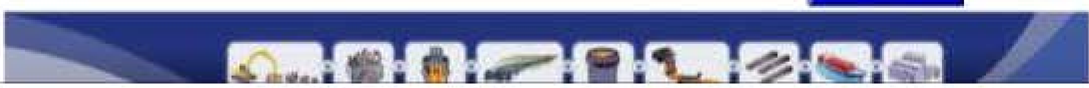
The other member of your team has already listed a set of risks of potential material misstatements for SteelCo. You have been asked to take 20 minutes to review your team member's ideas and record your ideas by adding to your team member's list of risks of potential material misstatements (what could go wrong) for SteelCo this year. You will then be asked to take 3 minutes to rank the top risks in order of their significance, and 12 minutes to identify the appropriate procedures for each of your top risks. Please note that your other team member will not be present at this time.

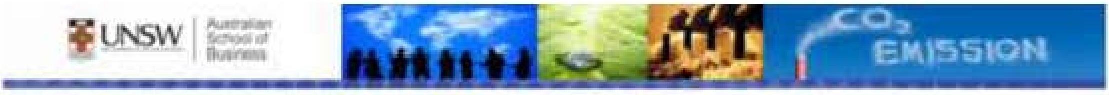
Some participants' answers (without identification to any individual) will be selected for review by other members of your organisation.

It is important that you work independently, concentrate for the full 35 minutes, and document as many additions to your other member's set of risks as you can in the first 20 minutes. Do not leave the recording to the end of the time period, i.e., record each idea as it is considered.

YOU CAN REFER TO THE CASE WHILE ANSWERING THE FOLLOWING QUESTIONS

Next





Participant No.: ra099

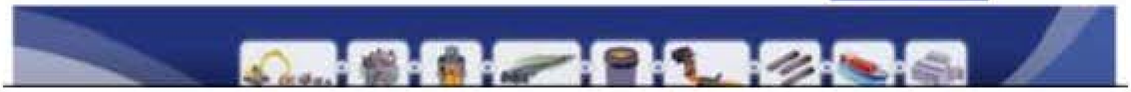
PART A – Task 1

You will be provided with a list of risks of potential material misstatements (what could go wrong) identified by the other member in your team. This member was asked to list risks which may have led to potential material misstatements in SteelCo's GHG statement.

Please list other risks which you believe may have led to potential material misstatements but have not been included by the other team member.

You have 20 minutes to complete this task

Next



Participant No: ra099

Required: The list of risks identified by your team member is provided to you here. Please enter the other additional risks of potential material misstatements in the boxes provided below

(please be as specific as possible).

Risk number	Risks of potential material misstatements identified by the other team member
1	test A
2	test B
3	test C
4	test D
5	test E
6	test F
7	test G
8	test H

Risk number	Other risks that have not been included by the other member Please be as specific as possible
1	Risk A
2	Risk B
3	Risk C
4	Risk D
5	Risk E
6	Risk F
7	Risk G
8	Risk H

Risk number	Other risks that have not been included by the other member Please be as specific as possible
9	Risk I
10	
11	
12	
13	
14	
15	

Risk number	Other risks that have not been included by the other member Please be as specific as possible
16	
17	
18	
19	
20	

Next

Participant No: ra099

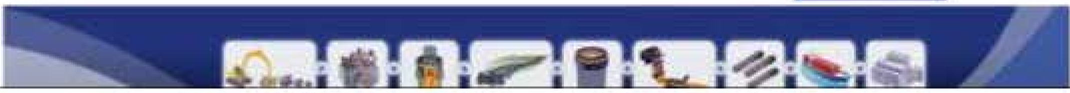
PART A - Task 2

From the set of risks you have added, please rank your 'top four' risks of potential material misstatements in order of their significance

(significant risks are the risks of material misstatements that, in your opinion, require special audit consideration)

You have 3 minutes to complete this task

Next

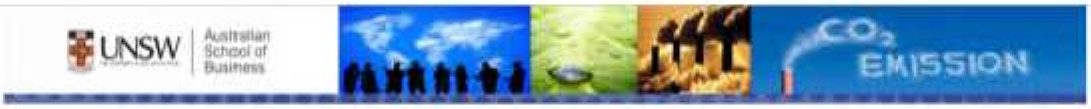


Participant No. ra099

Required: Please enter a number 1, 2, 3 or 4 in the provided boxes to identify the top 'four' risk (1 = most significant, to 4 = least significant).

	Other risks that have not been included by the other member	Top 'four' risks (1 to 4)
1	Risk A	1
2	Risk B	2
3	Risk C	3
4	Risk D	4
5	Risk E	
6	Risk F	
7	Risk G	
8	Risk H	
9	Risk I	

Next



Participant No: ra099

PART A - Task 3

Please identify the appropriate procedures to address your 'top four' risks that you have ranked in Task 2.

You have 12 minutes to complete this task

Next

Participant No: ra099

Required: Please enter the procedures for each risk in the boxes provided (please be as specific as possible).

1	Risks of potential material misstatements
	Risk A
	Appropriate procedures to address the assessed risks of material misstatements Please be as specific as possible
	Procedure A
2	Risks of potential material misstatements
	Risk B
	Appropriate procedures to address the assessed risks of material misstatements Please be as specific as possible
	Procedure B
3	
	Appropriate procedures to address the assessed risks of material misstatements Please be as specific as possible
	Procedure C
4	Risks of potential material misstatements
	Risk D
	Appropriate procedures to address the assessed risks of material misstatements Please be as specific as possible
	Procedure D

Next


UNSW
Australian School of Business
CO2 EMISSION
ra099
PART B
You have 5 minutes to complete this part
Next

Participant No: ra099

1. In Part A, you were asked to rank your top 4 risks from the set of risks you added to the original list. Assume you had been asked to rank the top 4 risks from the whole set (both from your team member's list and the risks you have added). Please rank your 'top four' risks of potential material misstatements in order of their significance (1 = most significant, to 4=least significant).

	Risks of potential material misstatements	Rank Number (1 to 4)
1	test A	1
2	test B	2
3	test C	3
4	test D	4
5	test E	
6	test F	

	Risks of potential material misstatements	Rank Number (1 to 4)
7	test G	
8	test H	
9	Risk A	
10	Risk B	
11	Risk C	
12	Risk D	
13	Risk E	

	Risks of potential material misstatements	Rank Number (1 to 4)
14	Risk F	
15	Risk G	
16	Risk H	
17	Risk I	

Next

Participant No: ra099

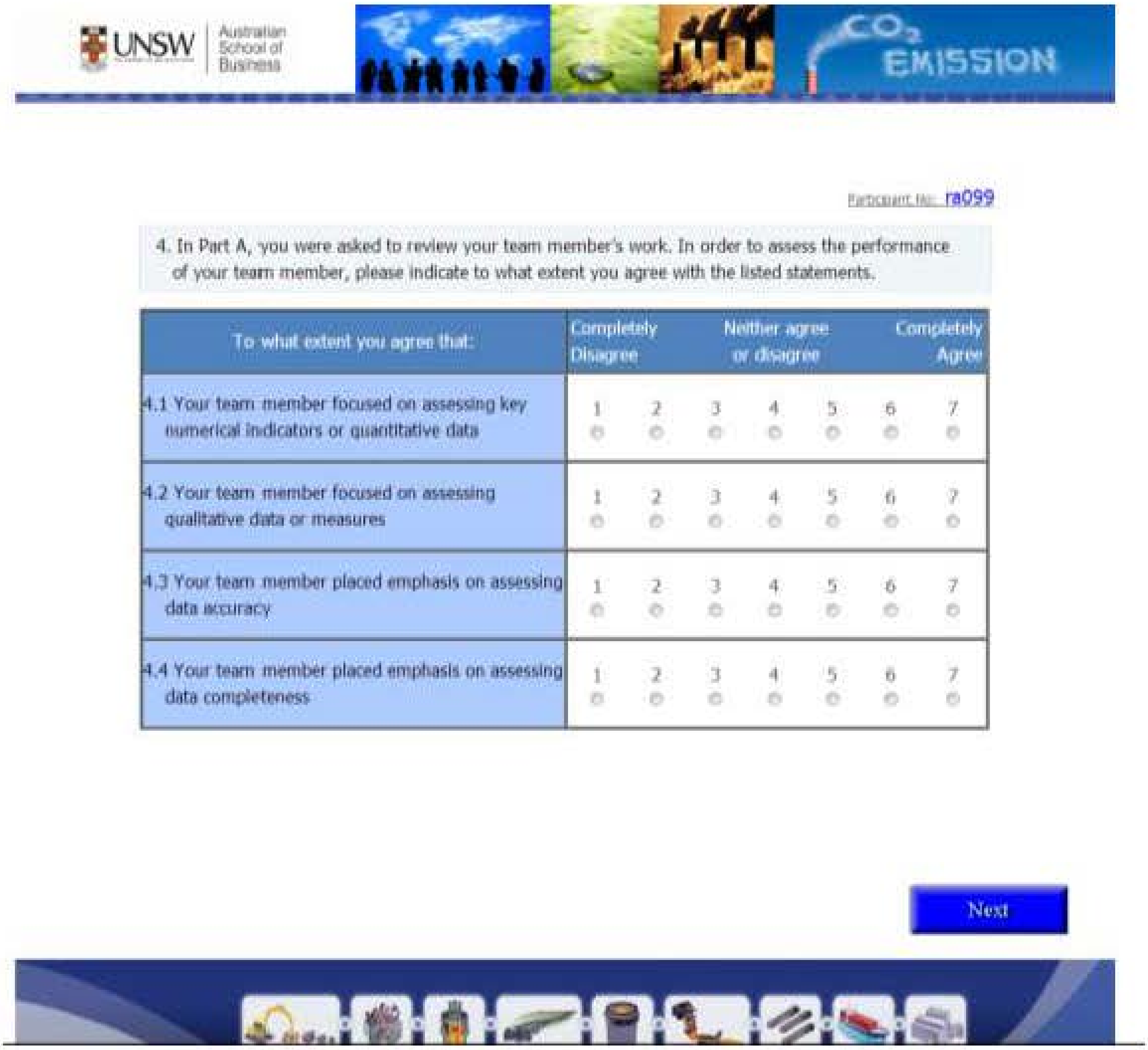
2. In Part A, you were asked to review your team member's work. Please indicate how you would rank yourself and your team member in terms of :

	You							Your team member						
	Low			Medium			High	Low			Medium			High
2.1 Ability to identify risks of material misstatements	1	2	3	4	5	6	7	1	2	3	4	5	6	7
2.2 Ability to Identify appropriate procedures to address the risks	1	2	3	4	5	6	7	1	2	3	4	5	6	7
2.3 Knowledge of the subject matter in environmental reports	1	2	3	4	5	6	7	1	2	3	4	5	6	7
2.4 Knowledge of the audit criteria and processes	1	2	3	4	5	6	7	1	2	3	4	5	6	7

3. On a scale of 1 to 7, with 1 indicating poor performance and 7 indicating excellent performance, how would you rate the overall performance of your team member in this study?

1	2	3	4	5	6	7
Poor performance			Moderate performance			Excellent performance

Next



UNSW Australian School of Business

CO2 EMISSION

Participant No: ra099

4. In Part A, you were asked to review your team member's work. In order to assess the performance of your team member, please indicate to what extent you agree with the listed statements.

To what extent you agree that:	Completely Disagree			Neither agree or disagree			Completely Agree
4.1 Your team member focused on assessing key numerical indicators or quantitative data	1	2	3	4	5	6	7
4.2 Your team member focused on assessing qualitative data or measures	1	2	3	4	5	6	7
4.3 Your team member placed emphasis on assessing data accuracy	1	2	3	4	5	6	7
4.4 Your team member placed emphasis on assessing data completeness	1	2	3	4	5	6	7

Next


UNSW
Australian
School of
Business
CO2
EMISSION
ra099
PART C
You have 5 minutes to complete this part
Next

Participant No: ra099

It is important that you complete the remainder of this task as an individual, i.e., please do not discuss this with others.

1. What is your position in the firm?(e.g. senior auditor, sustainability manager)

2. What is your tertiary education background?

- ☐ Accounting/Auditing
- ☐ Engineering
- ☐ Environmental Science
- ☐ Other (please specify)

3. What is your industry specialisation?

4. How many years of financial audit experience do you have? (if any)

years

5. How long have you been involved in conducting environmental/GHG/sustainability assurance engagements?

years

6. Approximately how many environmental/GHG/sustainability assurance engagements have you undertaken?

engagements

7. Have you been involved in conducting financial/GHG/sustainability assurance engagements for clients in manufacturing?

- ○ Yes, I have undertaken engagements
- ○ No

8. How much training on assurance for GHG emissions have you done?

hours, or

days

****

THANK YOU FOR YOUR PARTICIPATION

Next

Participant No: ra099

THANK YOU FOR YOUR PARTICIPATION