THE EFFECTS OF SCALE AND INFORMATION DISTRIBUTION ON GROUP DECISION-MAKING PROCESSES AND OUTCOMES

A Dissertation
presented to the
Faculty of the Graduate School,
University of Missouri-Columbia

In Partial Fulfillment
of the Requirements for the Degree

Doctor of Philosophy

by

MARK LANDIS

Dr. Vairam Arunachalam, Dissertation Chair

DECEMBER 2006

The undersigned, appointed by the Dean of the Graduate School, have examined the dissertation entitled:

THE EFFECTS OF SCALE AND INFORMATION DISTRIBUTION ON GROUP DECISION-MAKING PROCESSES AND OUTCOMES

Presented by Mark Landis

A candidate for the degree of Doctor of Philosophy

And hereby certify that in their opinion it is worthy of acceptance.

Professor Vairam Arunachalam, Dissertation Supervisor

Professor Thomas Dougherty

Professor Elaine Mauldin

Professor Christopher Robert

Professor Patrick Wheeler

ACKNOWLEDGMENTS

First and foremost, I would like to thank my advisor, Vairam Arunachalam, for his support and mentoring throughout graduate school and on this dissertation. His guidance and enthusiasm have been inspirational, and I will always value the time that I spent studying with him. And of course, his patience with me was greatly appreciated.

I would also like to thank the other accounting faculty members on my dissertation committee: Elaine Mauldin and Patrick Wheeler. Both have been supportive and helpful during the dissertation, and they have made the past four years of graduate school a much more enjoyable experience. Finally, I would like to thank the outside members of the committee, Chris Robert and Tom Dougherty, whose comments and insights have contributed much to this dissertation.

I am also indebted to Jere Francis, the director of the Ph.D. program in the School of Accountancy, who has been supportive of me and my chosen research paradigm. He will always serve as an inspiration for me to be a better researcher.

Finally, I would like to personally thank my friends and family, starting with my parents, Margaret and Bill, without whose support none of this would have been possible, and also Jennifer Urbanski, who has been a close friend for many years. I also owe a sincere thanks to Chad Burton whose friendship through challenging times will not be forgotten.

TABLE OF CONTENTS

LIST OF TABLES

Table
Page

THE EFFECTS OF SCALE AND INFORMATION DISTRIBUTION ON GROUP DECISION-MAKING PROCESSES AND OUTCOMES

Mark Landis

Dr. Vairam Arunachalam, Dissertation Supervisor

ABSTRACT

This dissertation bridges two research streams in the group decision making research: the weighted opinion research stream and social decision scheme research stream. In two experiments, the scale of the decision outcome – which is thought to create the differences between the two research streams – does not affect the underlying behaviors as expected. Results indicate that models of interpersonal influence can generally best explain decisions in both continuous scale and discrete scale settings, in both simple and complex tasks. However, if the network of interpersonal influences is not known, then in simple (information-scarce) tasks, central tendency behaviors can explain group decisions across both discrete and continuous scale settings. In complex (information-rich) tasks, faction-based behaviors tend to describe decision making processes. Additionally, and contrary to expectations, scale does not affect information processing when the information set is different among group members.

CHAPTER 1

INTRODUCTION

The importance of understanding group-decision making is growing in importance as recent auditing standards are mandating group discussions as part of the audit. In addition to brainstorming sessions (see Carpenter, 2005, for the implications of brainstorming sessions), SAS 99 notes the following group discussions should occur at the beginning of the audit when assessing the potential for misstatement due to fraud:

> **¶14.** Prior to or in conjunction with the information-gathering procedure described in paragraphs .19 through .34 of this section, members of the audit team should discuss the potential for material misstatement due to fraud …

Such group discussions should continue throughout the audit:

> **¶18.** Communication among the audit team members about the risks of material misstatement due to fraud should also continue throughout the audit – for example, in evaluating the risks of material misstatement due to fraud at or near the completion of the field work.

The standards also suggest that auditors should engage in group discussions near the end of the audit to reevaluate prior judgments which had already been made regarding the audit:

> **¶74.** At or near the completion of fieldwork, the auditor should evaluate whether the accumulated results of auditing procedures and other observations … affect the assessment of the risks of material misstatement due to fraud made earlier in the audit.

Although FASB casts these group discussions in terms of assessing the likelihood of fraud in the financial statements, interacting groups are occurring more frequently throughout the audit process. For example, Rich et al. (1997) note that interacting groups are increasingly replacing the traditional iterative reviews of audit working papers, while Bell et al. (2002) document a decision aid employed by KPMG which aids interacting groups of auditors who are making client acceptance and continuous risk assessments.

Research on interacting groups of auditors in the accounting literature remains scarce despite calls for more such research (Solomon, 1987; Rich et al., 1997; Bamber et al., 1998). While group process gains have been found in the auditing research (Rich et al., 1997), with only a few exceptions the extant group research in auditing tends to examine the sequential review process (e.g. Libby and Trotman, 1993). The sequential review process involves different decision processes than an interacting group of auditors and therefore may have different consequences on group performance; for example, interacting groups better reduce random bias while sequential reviews better reduce systematic bias (Trotman, 1985). This dissertation intends to examine the effects of different decision scales on the group outcomes (which offer insight into the group processes) of interacting groups[1] in both accounting and non-accounting settings. It also intends to examine the effects of scale on information processing when information is not equally distributed to all members.

The current group decision-making literature is essentially divided into three camps. One stream of literature, the "Social Decision Scheme" literature, argues that group decisions are based primarily on the initial opinions of the members who engage in some combinatorial processes to resolve differences. Groups may decide by majority or plurality rules, or by some other rules which will be discussed below. Such an approach allows specific predictions regarding the final group decision. Another stream, which is referred to as the "choice shift" literature, examines the shift between the average of the individuals' pre-discussion opinions and the final group decision. This branch of research relies on two theories – social comparison and persuasive arguments – to explain the direction of the shift, but it cannot offer a specific prediction as to the magnitude of the shift. A final stream of research called the "weighted opinion" stream is just beginning to develop which argues that the final group decision will be a weighted average of the initial opinions, and researchers are beginning to investigate the weighting scheme which best explains group decisions.

Study 1 of the dissertation attempts to identify when group decisions are consistent with certain group processes. First, group outcomes are predicted to be determined by the scale of the decision outcome. Although this is not a novel idea (see Hinsz, 1999), this has not been tested in the literature. This study makes the general prediction that majority-wins rules will dominate discrete scales except in risky tasks, in which case decision schemes which add influence to the more extreme members will be the relevant group processes. In continuous scales, groups could potentially employ either weighted opinion models or SDS-type models (note that in discrete scales, many weighted opinion models are not fine enough to distinguish them from an arithmetic mean model). Here, I predict that a weighted opinion model

[1] Solomon (1987) and Rich et al. (1997) refer to interacting groups of auditors as "audit teams".

will best fit the data, though I do not predict specifically which weighted opinion model will outperform the others. However, I do predict that weighted opinion models will outperform the discrete-type models from the social decision scheme literature which focus on normative influences. No predictions are made regarding an interaction of task type (i.e., risky versus no risk) and the weighted opinion model employed for continuous scales.

I also predict that groups will prefer to employ discrete models first, and then refine the decisions to respond on a continuous scale, in settings where the group must make both discrete and continuous decisions. Finally, I attempt to predict polarization that is consistent with the choice shift literature, the final stream of decision-making research, based on the application of social decision schemes versus weighted opinion models and the initial distribution of opinions.

The second study addresses the effect of scale when information is not evenly distributed; this is examined in an auditing decision-making environment. The auditing setting is predicted to change the processes employed because the tasks are more informational in nature. Further, the distribution of information may result in decision processes which may decrease the effectiveness of the audit. This may result because information is not used effectively by the group members, and the decision scale may enhance or mitigate this effect. Thus, the research question asked in the second study is "How does scale affect group decision-making processes and outcomes in an accounting setting, and what effect does the distribution of information among group members have on group processes?"

According to the first study, auditors would be predicted to employ majority-wins SDS models in discrete environments and weighted opinion models in continuous environments. However, the nature of the accounting decision-making setting might cause an increased focus on the information content of the discussions, and this focus on information rather than normative pressures to achieve consensus will likely change group decision processes in auditors, and particularly in discrete decision scales. Additionally, scale has also been found to affect the way that groups *evaluate* information. Research has shown that groups in discrete scales tend to focus on fewer pieces of information and assess higher importance to the remaining cues that are considered (Gigone and Hastie, 1997), weakening the influence of informational cues.

In an auditing environment, group members will collect information regarding a question, and they are likely to formulate an opinion based on the information they collect. Different team members will collect different pieces of information, but they will also share some common pieces of information among them. This study investigates the effect that scale has on the processing of unique information in an audit setting.

The second study is intended to contribute to our knowledge of interacting group decision-making in auditing settings, especially considering the dearth of group research in the accounting literature (Solomon, 1987; Rich et al., 1997). Recent auditing standards have called for group discussions throughout the audit process (see SAS 99, AICPA, 2002) and therefore highlight the need to understand the effects of group decision-making on audit judgments. The importance of understanding the effects of scale is important as auditor decision aids are employing different types of decision scales and representations in decision aids (Dilla and Stone, 1997a, 1997b; Bell et al., 2002). The implications of different information sets among the audit team members are also important as the current litigious environment places high stakes on conducting an effective audit.

CHAPTER 2

REVIEW OF THE INTERACTING GROUP DECISION-MAKING LITERATURE

Discrete judgments

Social decision schemes

Decisions faced by groups are often discrete in nature; that is, the group is asked to determine whether it should render a guilty or not guilty vote, provide a yes-no decision, or a select from a set of mutually exclusive alternatives. When transforming the distribution of initial opinions (the "decision schemes") into a final group opinion, Davis (1973) described how groups would rely on some combinatorial process to arrive at a group decision if the ultimate decision were discrete. Davis's (1973) general social decision scheme (SDS) model describes how groups may combine their opinions into a final decision. Under SDS, all possible constellations of initial opinions are considered. The number of possible constellations increases dramatically with the number of alternatives from which the group may choose:

$$m = \binom{n+r-1}{r} = \frac{(n+r-1)!}{r!(n-1)!}$$

where *m* equals the number of possible constellations, *n* equals the number of outcome alternatives, and *r* equals the number of group members.

Once all group members have established an initial opinion and the group's constellation is known, the group begins to deliberate. Laughlin and Hollingshead (1995) review some of the processes by which groups will combine their opinions. The process will determine the probability that the group will select a certain alternative, or D_j. If groups randomly select an alternative, then the SDS model is an equiprobability model in which all proposed alternatives have an equal chance of becoming the final group decision. If the group employs a voting mechanism, then the SDS models assigns a probability of 1.00 to the alternative favored by a majority or plurality. If one alternative is demonstrably correct, then a truth-wins model assigns a probability of 1.00 to the correct alternative if at least one person advocates that alternative. Table 1 illustrates how some decision schemes may change D_j, or the probability that a group selects a certain alternative, for each possible constellation given (r = 4, n = 2) and (r = 4, n = 3).

{Insert Table 1 here}

The particular decision scheme employed by a group often depends on the task or the nature of the problem that the group is trying to solve. Most often, groups employ a majority-wins model where the group decision is most likely to be the alternative that is favored by a majority of group members (Laughlin and Hollingshead, 1995; Hastie and Kameda, 2005). If a majority is not available, then groups tend to rely on a plurality-wins model or a median-wins model. However, in certain tasks, groups deviate from these general rules. For example, intellective tasks – tasks with demonstrably correct answers (McGrath, 1984) – usually exhibit a truth-wins pattern in which the group will select the correct alternative if at least one member arrives at the solution (Lorge and Solomon, 1955; Laughlin and Hollingshead, 1995). In SDS research, the accuracy of various decision rules is evaluated by testing the observed distribution of group decisions versus the expected distribution of group decisions using a goodness-of-fit statistic.

In this dissertation, decision-making will be investigated in three-person groups. Therefore, the following social decision schemes will be examined. Each is described below, and Table 2 provides a summary of these social decision schemes.

- *Majority wins / median wins* – If two or more members agree on an alternative, that alternative will be the final group decision. If each member has a unique pre-discussion preference, the group decision will be the median preferred alternative.
- *Risk wins* – The riskiest alternative favored by one of the group members will become the final group decision.
- *Conservative wins* – The most conservative alternative favored by one of the group members will become the final group decision.
- *Polarization wins* – If the group is predisposed towards a risky alternative for the task at hand (i.e., the average of the group members' pre-discussion preferences is risky), then the riskiest alternative favored by a group member will be chosen. If the group is predisposed towards a conservative alternative, then the most conservative alternative will be chosen. The theory of group polarization is discussed in detail in the following section.

Continuous and ordinal judgments

Choice shift and group polarization

In 1961, James Stoner found that surprisingly, groups were willing to take more risks in a chess tournament scenario than were the individual members of the group (Stoner 1961, unpublished thesis; cited by Wallach et al. 1962). Researchers now use an omnibus term, "choice shift", to refer to any difference – either more risky or more cautious – in the mathematical composite of the pre-discussion individual members' judgments and the final group judgment (Butler and Crino 1992). "Polarization" is a specific type of choice shift in which the group shifts away from a neutral position to a more extreme position in the same direction as the initial tendency of the individual members' pre-discussion preferences (Myers and Lamm 1976). In other words, group polarization is an intensification of the beliefs already held by the group members. For example, if a group's judgment is riskier than the average individual member's judgment, then that choice shift is an example of polarization if the initial tendency of the individuals was already risky. Polarization is contradictory to the notion of mean regression that extreme observations will regress towards the middle on repeat testings.

Two theories are generally used to explain group polarization: social comparison and persuasive arguments. *Social comparison theory* is a normative explanation in which individuals try to perceive and portray themselves in a socially desirable light consistent with their perceptions of the group's values (Isenberg 1986). Individuals are constantly processing information about how other people present themselves, and they adjust their own presentation or opinions accordingly. When group members engage in this social comparison process, an average shift in the direction of greater perceived social value occurs. *Persuasive arguments theory* is an informational explanation for the group polarization phenomenon. It posits that an individual's choice is a function of the persuasiveness and quantity of arguments that can be recalled from memory when formulating that position (Isenberg 1986). Group discussion will influence an individual by exposing that person to persuasive arguments which advocate a shift in a given direction (Vinokur and Burnstein 1978; Burnstein 1982). According to persuasive arguments theory, choice shift and therefore polarization are contingent upon the argument pool within the group.

It is important to note that the methodology employed by much of the choice shift research may not actually test for group polarization. A vast amount of research actually tests for sample polarization, not group polarization. Group polarization is the extremization that occurs *within* groups, while sample polarization is the extremization that occurs *across* groups (McGarty et al., 1992). A test of the shift between the entire sample's pre-discussion mean and the mean of all of the groups' decisions is a test of sample polarization, not group polarization, and no theoretical predictions exist for sample polarization with continuous dependent variables under either social comparison theory or persuasive arguments theory

(Rodrigo and Ato, 2002). Despite this critical distinction between individual groups and the entire sample, most research analyzes group polarization at the sample and not group level, so much of our existing knowledge in fact applies to sample polarization.

Weighted opinion models

The two major theories of group decision making discussed thus far have significant shortcomings when attempting to predict group judgments. The traditional conceptualization of group polarization – that is, choice shifts are the result of social comparison and / or persuasive arguments – does not offer a prediction for the magnitude of the choice shifts, only the direction. An SDS approach may be effective in predicting group decisions when the number of possible outcome alternatives is small. However, as the number of possible outcomes increases – as is the case if the decision outcome is located on a continuous number line, either bounded or unbounded – the number of possible decision constellations (*m*) increases or becomes infinite, and the SDS eventually reduces to a median-wins, mean-wins, or equiprobability decision scheme as no majorities or pluralities will exist. The weighted opinion stream of research addresses these problems by examining group decision-making processes on more continuous scales and allowing predictions of the final group outcomes.

Following Friedkin's (1998) notation, the general formulation of a weighted opinion model is

$$\mathbf{y}^{(\infty)} = \mathbf{V}\mathbf{y}^{(1)} \tag{1}$$

where $\mathbf{y}^{(\infty)}$ is a matrix of the actors' final opinions, $\mathbf{V}$ is a matrix of the weights of the opinions, and $\mathbf{y}^{(1)}$ is an $r \times 1$ column vector of the actors' initial opinions. The size of the $\mathbf{y}^{(\infty)}$ and $\mathbf{V}$ matrices may vary depending on the model being used. The challenge to researchers is to find the appropriate values in $\mathbf{V}$, i.e., the best estimate of the weighting of each individual's opinion on the final group decision.

James Davis, who over two decades prior had introduced social decision schemes, presented the social judgment scheme (SJS) in 1996 (Davis, 1996). This model argues that the weight is a function of the distance from the mean of the initial opinions. Under the SJS, the group opinion is a weighted average of each member's initial opinion, and consistent with both the theory of social power and attitude change models, the SJS predicts that the weight of each member's opinion exponentially decays as the distance between the initial opinion and the mean of the initial opinions increases. This model is presented formally in Appendix 1. Davis (1996) finds empirical support for the SJS model in a study of mock juries determining damage awards where the possible awards were bounded within a range of $0 to $1,000,000. Davis et al. (1997) again find support of this model in mock school board budget allocations, with the decision scale bounded between $0 and $2,000,000.

Noah Friedkin and his colleagues have developed another weighted opinion model. Friedkin (1998) develops a structural social influence network theory (hereafter, SINT) which in turn may be applied to group decision making in order to predict choice shifts (Friedkin, 1999). The model accounts for the endogenous influence each actor has on another during the group discussion process, and the entire network of the actors' susceptibilities to each others' opinions serves as the weight applied to each of the initial opinions to arrive at a final predicted group opinion. Friedkin's formulation predicts a final opinion for *each* member. If the matrix of interpersonal influences, which ultimately determines **V** (see Friedkin, 1999, for an explanation), is precisely measured, then each element of $\mathbf{y}^{(\infty)}$ will be equal to each other if the group is forced to reach a consensus decision. If the interpersonal influences are not precisely measured, then the predicted group decision is an average of the elements of $\mathbf{y}^{(\infty)}$ (Friedkin, 1999). To estimate the interpersonal influences, Friedkin's research currently uses a self-reported measure of the influence of another actor on an individual (e.g. Friedkin 1999; Friedkin and Johnsen, 1999).

Hinsz (1999), testing a series of weighted opinion models he labels "SDS-Q" (or Social Decision Schemes for Quantities"), finds evidence that assigning an arbitrary weight to the most valued opinion by the group may accurately predict the group's final decision as it heavily weights the preferred opinion but still considers the influence of the minority opinion. Specifically, he found that when groups were setting a goal in a performance task, a decision scheme was assigns a weight of 0.1 to the lowest goal and a weight of 0.9 to the average of the remaining members' goals best predicts the group's final goal. Such a decision scheme results in a goal that is just below what the group thinks it can actually achieve, and so the likelihood of achieving that goal is high and therefore attractive to the group members. A modification of that decision scheme may also be employed in the present study. If a majority exists, but the group still wants to account for the minority, then a weight of 0.8 to the majority opinion and a weight of 0.2 to the minority opinion may a compromise position. A weight of 0.8 is selected rather than 0.9 because the fineness of the scale may prevent the group decision from straying from the majority opinion (for example, on a seven-point scale, assigning a weight of 0.9 to the majority opinion will likely lead to a prediction that is virtually identical to the majority opinion).

Given the above discussion, and the results of Hinsz (1999) on other continuous-based social decision schemes, the following models will also be tested for group decisions. These are also summarized in Table 2.

- *Arithmetic mean* – The mean of the pre-discussion preferences will be the final group preference.
- *Nearest to the center of gravity of proportionality* – the alternative that lies closest to the arithmetic mean of the pre-discussion preferences will be the final group decision (this accounts for the fineness of the scale).
- *Weighted majority* – the majority opinion will carry a weight of 0.8 and the minority opinion will carry a weight of 0.2.
- *SJS* – the influence of an opinion will decay exponentially as it deviates from the mean of the opinions.
- *SINT* – the group decision will be a function of the weighted influences of each member on each other member.

{Insert Table 2 here}

CHAPTER 3

RESEARCH ON INTERACTING GROUP PROCESSES IN AUDITING

Choice shifts and social decision schemes

Five studies have examined the choice shift phenomenon in the auditing literature, and the results have generally been mixed. Four of the studies found significant choice shifts, but of these, only two papers (Bamber et al., 1996; Carpenter, 2005) have results that are consistent with the theory of group polarization. One audit study did not find any choice shifts. However, similar to most research in the psychology literature, all of the papers tested for sample polarization, not group polarization, and therefore not much is known about the underlying group processes involved. With the exception of Carpenter (2005), this research was conducted in two different communication environments (face-to-face and computer-mediated). While computer-mediated communication is growing in importance due to the internationalization of audit clients and the increasing availability of technology (Bamber et al., 1998), this dissertation is only concerned with results obtained in the face-to-face conditions of those studies as computer-mediated communication may alter decision-making processes (Arunachalam and Dilla, 1992, 1995; Fjermestad and Hiltz, 1998) and is beyond the scope of this study.

Of the five papers investigating choice shift, three papers examine shifts using a continuous (0 to 100 point) scale. One paper found no shifts (Schultz and Reckers, 1981), one paper found small shifts (Reckers and Schultz, 1982), and one paper found significant polarization (Bamber et al., 1996). The striking inconsistency of these results appears to contradict models of decision-making based on the initial distribution of opinions. All three of the studies used virtually identical tasks – specifically, the case developed by Schultz and Reckers (1981) – and the average pre-discussion likelihood of disclosure was approximately 40% in all three studies for the face-to-face groups. However, because the authors in all three studies examined shifts at the sample level and not the group level, the netting effect of the shifts *between groups* went unaccounted, and the within-group shifts may have been more predictable given the initial distribution of opinions within the groups.

Two papers examine choice shifts using more discrete scales. Both of these studies find significant choice shifts, though only one is in the direction predicted by polarization. Karan et al. (1996)

asked student subjects to provide acceptable audit risk judgments on a scale of 1 (indicating low acceptable audit risk) to 10 (high acceptable audit risk). The average individual ratings were 6.03, and contrary to the theory of group polarization, the group ratings shifted down to 5.60 which was a significant shift.

However, the subjects in Carpenter's (2005) study exhibited striking group polarization, even by standards observed in purely theoretical psychology laboratories. In her study, audit groups were given a set of financial statements and asked to judge the likelihood that fraud existed in the financial statements on a scale of 1 to 11. In one condition, auditors were given financial statements that actually contained fraud, and in the other condition, the financial statements did not. In both conditions auditors thought that the financial statements contained fraudulent numbers, and the auditors' opinions strengthened after group discussion. Even more interesting, though, is the fact that the average group judgment exceeded the average opinion of the most extreme group member. Carpenter's (2005) groups consisted of an audit staff, senior, and manager, and across the sample, the managers consistently had the highest likelihood ratings for fraud in both sets of financial statements. However, their average opinions were only 7.35 and 6.40 for fraud and no-fraud statements, respectively. The group opinions were 7.45 and 6.90, respectively. Thus, a sample-wide analysis suggests that the group judgment was not bounded by the most extreme opinion in the group, and currently, only the choice shift literature can account for such a large shift.

It appears that subjects in both the Karan et al. (1996) and Carpenter (2005) studies applied continuous decision models to the ten-point and eleven-point scales, respectively, which is unlike the treatment of probability scales as discrete scales that appears to occur in traditional CDQ tasks (Kerr et al., 1996). The shift towards the center of the scale in Karan et al.'s (1996) study suggests that the minority opinion was incorporated into the group decision. In Carpenter's study, social comparison processes clearly appear to be at work as the groups on average shifted past the most extreme – but not most accurate – member who also happened to be the highest ranking member in terms of organizational hierarchy. Differences in member status are beyond the scope of this dissertation, though that subject appears to be a promising area for future research (for example, see Bonner et al., 2002, for a social decision scheme model that assigns higher proportionalities to expert members' opinions). However, Carpenter's (2005) results do shed insight into the group processes studied in this dissertation as they suggest a continuous decision-making model was employed.

Only one study in the accounting literature investigates social decision schemes. Chalos (1985) asked groups of four practicing bank loan officers to issue likelihood assessments that a company would default on a loan using a fifty point scale (0 through 10, in 0.2 increments), and subjects were also asked to choose a course of action (call the loan, work out the loan, or continue as normal). Each group was to assess eight companies, and these eight cases were based on actual companies allowing for a *post hoc* test of accuracy for truth models for which course of action the bank committee should have taken.

Chalos (1985) found that the truth-wins model best fit the data. The truth-wins model predicts that if at least one member favors the correct alternative, the group will ultimately choose that alternative. Chalos's table 3 (page 536) shows that in two cases – case numbers three and six – the group judgment reversed the initial tendency of the pre-discussion opinions on the likelihood of default if the scale were dichotomized. In those two cases, the average pre-discussion assessment was less than 5 (suggesting no default), but the group decision was greater than 5 (suggesting default). This contradicts the theory of polarization and compensatory strategies which merely impound minority opinions, but instead suggests that the groups did in fact employ a discrete decision process (most likely truth-wins given the model fits) in those cases which in turn affected the likelihood assessments. In other words, this suggests that groups first rely on a discrete decision rule to identify the proper choice of action, and then subsequently apply a continuous decision rule to refine the specificity of the decision. In all of the remaining six cases, though, polarization occurred as the groups' opinions were more extreme than the initial opinions as opposed to simply regressing towards the middle of the scale, once again seemingly rejecting compensatory strategies.

Although Chalos's (1985) task involved loan officers making a loan call decision, the subjects were presented with a payoff matrix which imposed a cost-benefit tradeoff when pursuing one course of action or another, and the payoff depended on the accuracy of their choice. However, this is different than an audit environment in which auditors must balance litigation risk against client retention and audit cost, so Chalos's subjects were able to polarize with fewer consequences.

In sum, despite the inconsistent results, the accounting literature appears to suggest that given a continuous or discrete probability scale, groups will tend to employ continuous decision-making processes. However, when given both a discrete set of alternatives and a continuous assessment scale, the group will first make a discrete choice and then a continuous judgment. With the exception of Chalos (1985), none of these papers have examined processes at the individual group level; this research currently analyzes sample-wide data.

Scale effects

Dan Stone and his colleagues authored a series of papers which generally found that accuracy in individual judgments, at least in terms of consistency, increases in numeric scales, while speed increases in discrete scales (Stone and Schkade, 1991, 1994; Stone and Dilla, 1994; Dilla and Stone, 1997a, 1997b). Reimers et al. (1993) found that linguistic scales had higher risk assessments in individual judgments than numeric judgments did, and when mapped into numeric responses, the higher linguistic assessments translated into statistically larger sample sizes using the decision aid employed by the auditing firm that provided the subjects. Piercey (2006) also found that linguistic scales lead to more extreme judgments than numeric scales.

Currently, no research on scale effects *in groups* has been conducted in an accounting setting, and Gigone and Hastie (1997) is the only study in any discipline of which I am aware that explicitly examines scale effects in groups. This study will be discussed more extensively in the hypothesis development section of study 1, but the general results of the study were that groups making continuous judgments impounded more information into the decision and weighted pieces of information more evenly, whereas groups making discrete choices tended to focus on information which supported the final group choice. Also, group discussions were shorter in choice settings if the members' initial choice preferences were unanimous (Gigone and Hastie, 1997).

Effects of information distribution

Throughout the course of an audit, individual audit team members will independently discover or have access to different pieces of information which other team members will not possess. When group members with different information sets begin discussing a common problem, a lengthy stream of psychology research indicates that the group discussion will tend to focus on those pieces of information which are shared across all members (Stasser and Titus, 1985, 1987, 2003; Gigone and Hastie, 1993; Winquist and Larson, 1998; Chernyshenko et al., 2003). However, as the length of the discussion increases, the probability that unique information is entered into the group discussion increases. This is because the groups initially discuss the shared information, but eventually the members run out of shared information to discuss, and they proceed to discuss unique information (Larson et al., 1994).

Surprisingly, though, once unique information is introduced into the group discussion, it is not repeated in the group discussion as much as shared information; this is particularly true in hierarchical groups where lower-status members are reluctant to discuss uniquely-held information in the presence of a superior or expert. This was demonstrated in analyzing group discussions of medical students making diagnoses jointly with medical residents (Larson et al., 1996). Further, when a group member hears *another* group member's unique information, the first member does not value that piece of information as much as the owner of that information (Chernyshenko et al., 2003). Therefore, the effects of novel arguments on group discussions are not as influential as shared information. This may be because unique information does not facilitate reaching group consensus (Postmes et al., 2001; Stasser and Titus, 2003) or due to egocentric biases in which people recall their own contributions to group discussions more than others' (Ross and Sicoly, 1979; Chernyshenko et al., 2003).

Only one study in the accounting literature examines the effect of information distribution in interacting groups. Murthy and Kerr (2004) present auditing students with a hidden-profile task regarding a discrepancy between a client's accounts payable ledger and its supplier's records. This was an intellective task in which a correct answer existed and was demonstrable; however, the information was distributed in such a way that no one person could solve the problem with only his or her information set. The findings revealed that face-to-face groups only correctly answered both parts of the problem 44 percent of the time. Murthy and Kerr (2004) conclude that the inability of interacting groups to solve the problem was due to difficulty in organizing the information in group memory. Other groups which were solving the same problem using an internet bulletin board tool, a group decision support system which allows people to better organize and post information into group memory, were able to solve both parts of the problem 77 percent of the time.

CHAPTER 4

STUDY 1

Background

To date, all of the studies of which I am aware try to examine which model best explains behavior by examining the fits of the various models to a single set of group decisions. In other words, the scale is not manipulated between conditions, but instead, all groups are given one task and the same scale, and the models are tested within that one sample. Further, these papers usually remain within the same research stream; that is, SDS models are tested against other SDS models (e.g., Laughlin and Hollingshead, 1995), polarization is tested to see whether or not it exists (e.g., Sia et al, 2002), and the weighted opinion models are tested against other weighted opinion models (e.g., Davis et al., 1997; Hinsz, 1999). Two papers that do cross research streams are Zuber et al. (1992) and Crott et al. (1991); these papers are discussed further in the hypothesis development section.

The first study of the dissertation takes a different approach to the question of which model of group decision processes best explains group outcomes by posing same judgment to different groups, and manipulating the scale between two conditions (discrete and continuous). In a third condition, groups are given both types of scales, and the consistency of the models across the conditions will be examined. This research design allows for a test of the best-fitting decision-making model for the same decision under two different scales, and it also allows for a direct examination of the effects of scale on decision outcomes. Different group processes are predicted to be used for the same decision under different scales. Also, under a conservation of cognitive effort approach, groups are predicted to prefer to use SDS processes rather than attempting to combine continuous decisions if both options are available.

In addition, the first study of the dissertation addresses the effects of risk in decision making under different scales. In some decisions under uncertainty, social norms may dictate the course of action that the group should take, and therefore the group will follow a truth-wins SDS model under discrete scales in risky decision settings (Laughlin and Earley, 1982). Such conditions should be reflected in the distribution of initial group opinions. However, under continuous scales, some group members may counter the norm, and the fineness of the continuous scale allows the group the incorporate that minority opinion. Therefore, under conditions of risk, discrete alternative choices are predicted to result in more extreme opinions than continuous judgments.

Finally, the first study offers the only explicit test of sample polarization of which I am aware. Group polarization is the intensification of beliefs *within* groups, while sample polarization is the intensification of beliefs *across* groups. One implication of SDS is that the entire sample will polarize if the sample is skewed towards one belief. While the notion that sample polarization is a statistical artifact of SDS is controversial (see Myers and Lamm, 1976, for a critical assessment of this view; see Kerr et al., 1996a, 1996b, for support), the author has not seen any papers that explicitly test for sample polarization, although many papers in fact test for sample polarization due to methodological choices despite motivating the papers with theories of group polarization (McGarty et al., 1992; Rodrigo and Ato, 2002). In this paper, conditions are hypothesized when sample polarization is expected to occur; specifically, it is predicted under discrete and risky decision environments but not continuous environments. These predictions link the weighted opinion models and social decision schemes to the choice shift literature.

Theory and Hypothesis Development

The theories discussed in the literature review section apply mainly to two types of decision settings: those in which the decision scales are discrete (SDS), which is often called "group choice", or continuous (SJS, SDS-Q, SINT, group polarization) which is labeled "group judgment". Research has shown that the scale of the decision response will affect the decisions made by people, though almost all of this research is done at the individual and not group level. At the individual level, Dan Stone and his colleagues find that individual judgments are more accurate in terms of consistency when people are given numeric scales rather than discrete scales, though they are quicker at evaluating information when given discrete scales (Stone and Schkade, 1991, 1994; Stone and Dilla, 1994; Dilla and Stone, 1997a, 1997b). However, only one paper examines scale differences in a group setting (Gigone and Hastie, 1997). The authors found that groups given discrete scales discussed less information than groups making continuous judgments. This may be due to an elimination-by-aspects (Tversky, 1972) or similar strategy. Gigone and Hastie (1997) also found that when the pre-discussion opinions were unanimous, group members tended to end the discussion immediately, and this was more likely to occur in group choice settings because the number of possible constellations *m* is much smaller.

Given the findings that individuals more quickly evaluate information in group choice settings, and the findings that groups discuss less information in group choice settings, groups should be more inclined to engage in choice behaviors rather than judgment behaviors. This is consistent with the notion that effort is likely the most important factor in cognitive cost-benefit calculations (Stone and Schkade, 1994; see also Kleinmuntz and Schkade, 1993). Social decision schemes consume less cognitive effort than continuous decision schemes (Hastie and Kameda, 2005), suggesting that groups should prefer to employ those combination processes. Additionally, such a prediction is consistent with Laughlin and Ellis (1986) who argue that in the absence of demonstrable correctness, people tend to find "safety in numbers" and rely on social consensus as a valid response (Kerr and Tindale, 2004; Festinger, 1954). Thus, the first hypotheses stated in alternative form are:

> **H1:** When making decisions in neutral settings, groups will employ majority-wins social decision schemes in discrete scales and a weighted opinion model in continuous scales.
>
> **H2:** When both a discrete scale and continuous scale are available, groups will utilize the discrete scale to arrive at a decision.

As discussed in the literature review, most of the research on group polarization in judgment and decision-making utilizes the CDQ which presents subjects with several vignettes of risky decisions. The subjects are usually asked to respond on some discrete probability scale, and the average of the pre-discussion responses across the entire sample is compared to the average group response across the entire sample. This statistical method treats the probability scales as continuous, and if groups operate as if the probability scales are continuous, then a continuous judgment scheme should better account for the final group decisions. However, given the (lack of) fineness of the probability scales, and the relatively small number of unique constellations *m* that may exist within the sample, groups may employ discrete choice schemes rather than continuous judgment schemes when making group decisions with discrete probability scales.

The argument that *m* may be small in such situations is, on the surface, tenuous given the large number of decision outcomes that are available even on a relatively small probability scale such a seven-point scale. However, certain decision settings have decision outcomes which are clearly favored over other outcomes (Laughlin and Earley, 1982). In such settings, ceiling effects of the scale should create individual judgments which are clustered towards one end of the scale, reducing the number of unique constellations that will be formed given a limited sample size. The formation of groups should create groups in which some responses are duplicated, and majority decision schemes may emerge.

Additionally, Kerr et al. (1996a) note that in risky decisions, some tasks may have alternatives that are appealing enough that they serve as a "demonstrably correct" response and therefore initiate a truth-wins mode of thinking for the group. Here, risky tasks refer to the Choice Dilemma Questionnaire presented by Wallach et al. (1962), where the decisions that are made have some (hypothetical) consequence to the decision maker. Kerr et al. (1996a) cite the example that if a football captain is given the choice of attempting a one-point play at the end of the game which results in a tie, or a riskier play which results in a win, the riskier play may be "correct" because the point of competitive sports is to win. An early influential study on choice shifts and social decision schemes, Laughlin and Earley (1982), provides support for this notion: 73% of groups had constellations which supported "going for the win". They found that a risk-supported decision scheme best accounted for the group decision in a dichotomous decision scale on some tasks, even more so than a majority-wins scheme.

Two important studies have approached examined risk-taking in both an SDS perspective and choice shift perspective. Consistent with SDS, Crott et al. (1991) find that median social decision schemes predict group decisions better than an intensification of the beliefs already held by the population. Also, Zuber et al. (1992) gave subjects lists of persuasive arguments before group discussion, but still found significant choice shifts after the group discussion. According to persuasive arguments theory, once group members have all of the information, shifts should not occur. Instead, Zuber et al. (1992) found that a reduced-pairwise comparison model, a median model which is consistent with Crott et al.'s findings, best predicts group decisions. These results contrast with Laughlin and Earley (1982) who find risk-wins decision schemes may best fit risky tasks. Considering the potentially discrete scales used by Zuber and Crott (5 and 9 point scales, respectively), the failure of persuasive arguments in Zuber et al.'s study to account for choice shifts, and the lack of polarization in Crott et al.'s study, these studies indicate that SDS decision processes were likely at work in CDQ tasks.

Given Kerr et al.'s (1996a) discussion and the findings of Laughlin and Earley (1982), Crott et al. (1991), and Zuber et al. (1992), it appears that traditional probability scales in risky settings function as discrete scales rather than continuous scales. In severely risky or conservative settings, a risk-wins or conservative-wins model should emerge if the preferred alternative is consistent with a strong social norm. In more neutral settings, a majority-wins model should emerge, and when no majority exists, the group should resort to median-wins following prior research. However, in continuous scales, the processes will still focus on the combining all opinions into a group decision rather than normative pressures. This leads to the following hypothesis, stated in alternative form:

> **H3:** In risky decision settings, groups will employ risk-wins or conservative-wins decision schemes in discrete scales and weighted opinion models in continuous scales.

The implications of H3 are that probability scale environments will result in more extreme judgments than continuous-based scales in risky settings. In neutral settings, no difference is expected if the distribution of initial opinions is symmetrically distributed around the median. Davis (1973) discusses the reasons for this in his initial presentation of group decision schemes; this is an alternative explanation to group polarization instead of the traditional social comparison and persuasive arguments theories (which at the time were still being developed). This application of SDS assumes that groups consider ordinal alternatives as unique, discrete choices.

If groups employ a majority-wins SDS and if the distribution of constellations is symmetric around the median alternative, then no sample polarization should occur. Under these conditions, for every group in which the majority supporting alternative A outnumbers a minority supporting alternative B (or C, or D, etc.), there will be a group in which the minority alternative becomes the majority alternative and is able to suppress the other alternatives. Ultimately, the average of the group discussions will equal the average of the pre-discussion preferences. However, if the initial preferences of group members are not symmetrically distributed around the median but instead are skewed in either direction, sample polarization may result from a majority-wins SDS for discrete but ordinal decision alternatives.

In skewed environments, the probability that a group contains a majority is increased, and a majority-wins decision scheme will induce polarization within each group. Further, because a disproportionate number of groups will have majorities pulling the minority opinion in the same direction, sample polarization should be observed. Note that this also applies to median-wins decision rules even when no majorities exist. If the pre-discussion opinions are skewed towards one "pole", the probability that the median alternative in any group is nearer the preferred pole is greater than the probability that the median is nearer a minority opinion. Thus, sample polarization should also occur. If the groups' tendencies are already risky, then a risk-wins rule will obviously lead to sample polarization, as will a conservative-wins rule if the groups' predispositions are conservative.

However, in continuous environments, polarization may or may not occur because the groups engage in compensatory strategies which incorporate the minority opinion. If the groups employ SJS, then the exponential decay of the minority member will cause very small group polarization, whereas if the groups employ an influence model, the amount of polarization is unknown. Thus, the fourth hypothesis stated in alternative form is:

> **H4:** Sample polarization will occur in discrete decision scales but not in continuous decision scales.

The SDS approach to group polarization illustrates why polarization is generally considered a decision bias. While group decision making can improve decision accuracy, it may also attenuate or strengthen individual biases. Generally, a consensus in initial opinions will exaggerate individual biases, while an even distribution of opinions may or may not attenuate them (Kerr et al., 1996b). As Kameda et al. (2003) discuss, group decision making will only overcome individual biases if the group collectively identifies and corrects the individual biases. However, if a majority of the group shares the same bias, then the individual bias will likely be accentuated through group discussion (Einhorn et al., 1977; Kerr et al., 1996a; Sunstein, 2002; see Rutledge and Harrell, 1994, for an example of this in a budgeting setting). Biases have been well-documented in judgment tasks (e.g. Kahneman et al., 1982) and therefore group polarization may reduce decision accuracy. Kameda et al. (2003) illustrate that larger group sizes will increase the magnitude of polarization under majority-wins decision schemes, and therefore larger groups sharing the same biases will compound their errors more severely than smaller groups.

Method

Two sets of within-subjects tasks were developed for this study. The first set of tasks asked subjects to assume the role of a teacher's assistant who has been asked to grade responses to two questions: for the first question, they were asked to grade a hypothetical student's definition of capitalism, and in the second, they were asked to grade a student's haiku (see Appendix 2). These two tasks are used to test H1 and H2. The second set of tasks consisted of three traditional CDQ items (chess, surgery, career change) which have been found in past research to induce a risky shift, no shift, and a conservative shift, respectively. Therefore, each group completed five questions: two questions from the first set of tasks, and three questions from the second.

Scale was a between-subjects manipulation. In the two grading tasks, there were three scale conditions: continuous only (assign a percentage grade), discrete only (assign a letter grade), and both continuous and discrete. Subjects in all conditions were instructed that the course followed a standard university curve of 90% to 100% is an A, and so on, and that the hypothetical student was promised to receive at least 50% for turning in the assignment. Thus, a letter grade could be converted to a number grade (an F became a 55%, a D became a 65%, etc.), and *vice versa*, for comparison across conditions where each letter grade consisted of a 10-point range. In the three CDQ tasks, there were two conditions: continuous (assign a probability from 10% to 90%) and discrete (select one of the following five probabilities: 10%, 30%, … , 90%).

Two-hundred eighty-five students from an introductory computer class at a large Midwestern university participated in the experiment and received extra credit worth approximately one percent of their total grade in the course. Students are appropriate for this study because decision-making theory is being examined directly and independent of an accounting setting, and the use of professional accounting subjects would be inefficient in such a situation (Libby et al., 2002). Subjects arrived at the experiment and were given instructions. They were then randomly assigned to groups of three and began working on the first set of tasks. Each subject worked through each question individually. Then, the subjects worked together as a group to answer the questions together. Finally, the subjects answered the questions individually again. After finishing the questions, subjects then allocated poker chips which represented influence for the SINT model (see Appendix 1). This process was repeated for the second set of tasks, and after allocating the poker chips, they completed a post-experimental questionnaire, were thanked for their participation, and were dismissed.

Dependent variables and model testing

Studies which test the effectiveness of various models employ different methods of evaluating the models. Usually, tests involve three dependent variables: model bias, model precision, and hitrate. Occasionally, papers employ other tests (Wasserman and Davis, 1991; Zuber et al., 1992; Hinsz, 1999), though most papers include examinations of bias, precision, and hitrate. However, to date, all studies of which I am aware involve model tests of one sample with one condition, whereas this study has multiple conditions. Examining multiple dependent variables across multiple conditions may lead to confusing results, and therefore I will use model precision as the primary dependent variable. Model precision, which is the absolute value of the difference between the model's prediction and the actual group decision, best reflects accuracy because overestimates or underestimates of the group decisions are both counted as errors. Model bias, which is the simply the difference between the model's prediction and the group prediction, allows overestimates and underestimates to net against each other, and therefore a highly inaccurate model may appear to correctly to predict group decisions if the overestimates and underestimates are symmetrical (consider, for example, a state where polarization wins is the true model of group decisions: a mean wins model may have a very small bias if groups are polarizing in opposite directions even though groups are engaging in exactly the opposite behavior of mean wins). Hitrate is the percentage of groups for which the model exactly predicts the group decision (precision < 0.5), but this measure is heavily biased against weighted opinion models in discrete scale situations as the opinion weights must lead to a group prediction which is equal to one of the possible scale responses. In this study, model precision (where a lower precision indicates a better model fit) is calculated and presented in the tables, and the models are also ranked by precision for ease of comparison across conditions. Hitrate is also presented for informational purposes, though the primary analyses and discussions will focus on precision. Model bias will be discussed in the situations where it sheds insight into the group processes.

The statistical test which is frequently employed in the literature is a test of the differences in model precisions. In these tests, the best performing model's precision (i.e., the model with the lowest precision score) is compared to each of the other models' precisions by conducting a *t*-test on the difference scores. However, testing the difference scores violates the independence of observations assumption because the same actual group decision is used to calculate each of the models' precisions (Wasserman and Davis, 1991). Further, most of the models are derived from the same set of initial opinions which also can lead to non-independence. Wasserman and Davis (1991) demonstrate that using *t*-tests in these situations leads to a conservative bias in the tests. To date, no other method of comparing model precisions has been developed, and given the conservative bias in the tests, the tests of the differences in model precision are used in this study. Because the tests must be directional by construct, one-tailed tests are employed in model tests.

The fourth hypothesis addresses polarization. Polarization is defined as the intensification of already-held beliefs by the group members, and in general, it is calculated by changing the sign on the choice shift (which is the absolute value of the difference between the mean of the pre-discussion opinions of the group members and the actual group decision) to positive or negative, depending on if the choice shift occurred in the direction that the group already favored. This is done for each group to calculate *group* polarization; to calculate *sample* polarization, the overall difference in the means of all of the pre-discussion opinions for the whole sample are compared to the overall means of all of the group opinions (though the specific calculation of sample polarization is done to account for the non-independence of observations). Landis et al. (2006) offer a detailed discussion of methodological considerations and challenges in measuring polarization, and in this study, the following calculations for polarization are employed. In the risk-free tasks, because of the extreme skew in the data relative to the midpoint of the scale (median grade for capitalism = 91.0; median grade for haiku = 81.5; midpoint of scale for both capitalism and haiku = 75.0), the median score is used as the fulcrum from which groups are thought to polarize rather than the midpoint (Myers and Lamm, 1976). Therefore, groups whose members hold an average pre-discussion opinion higher than 91.0 (81.5) in the capitalism (haiku) task are coded as exhibiting positive polarization if the group decision is higher than the average of the initial opinions, and that shift is coded as negative polarization if the group decision is lower than the average of the initial opinions. If the groups have an average initial opinion lower than 91.0 (81.5) for the capitalism (haiku) task, then a shift down from the average initial opinion is coded as positive polarization and a shift up is coded as negative polarization (Landis et al., 2006). In the risky tasks, the pre-discussion averages are close to 50, or the midpoint of the scale, so 50 is retained as the fulcrum from which polarization is thought to occur. Calculations are identical to those described for the risk-free tasks.

Results

The first hypothesis predicted that majority wins would be the dominant model in discrete scales in risk-free settings, while a weighted opinion model would dominate the continuous scales. Because majority wins models can, by definition, only be employed by groups if a majority exists, the sample is first pared down to those groups in which two or three members agree on an alternative before discussion. Further, if all three members agree before discussion, then all of the decision models predict that the unanimous initial opinion will be the group opinion, and those groups are also eliminated from this analysis. In the continuous scale, this leaves 12 groups in the capitalism task (19 groups were eliminated because all members had unique initial opinions, and no groups had unanimous initial opinions) and 12 groups in the haiku task (again, 19 groups had all unique initial opinions and none had unanimous initial opinions) for a total of 24 groups. In the discrete scale, 21 groups were left in the capitalism task (after eliminating 5 unique and 6 unanimous groups) and 21 groups in the haiku (after eliminating 9 unique and 2 unanimous groups).

Results of the analysis on the remaining groups are presented in Table 3. In the continuous scale, the weighted opinion models fared quite well as the five continuous models better fit the group decisions than the four discrete models, thus supporting the second part of H1. SINT was the best model with an average miss of 1.56, and it marginally outperformed the other weighted opinion models, as well as best discrete model, majority wins (*p* = 0.05). The *p*-value for the difference between SINT and majority wins is higher than the *p*-values between SINT and the other weighted opinion models because of the variation in the discrete model precision scores. In general, the standard deviations for the discrete models (not tabled; range from 3.34 to 5.57) are higher than for the continuous models (range from 2.06 to 2.47). Consistent with Hinsz (1999), the hitrates for the discrete-based models were generally higher than for the continuous models (with the exception of the low wins model with a 4% hitrate). However, the discrete models had higher precision errors indicating that when they missed, they missed by a large amount. Across both tasks, SINT also significantly outperformed the second-best model, mean wins (*p*-value = 0.040). The top three continuous models all systematically underestimated the group decision (not tabled). For those models, the average bias was -1.071 for SINT (the t-test for a significant difference from zero yielded a *p*-value of 0.037), -1.359 for SJS (*p*-value = 0.044), and -1.417 for mean wins (*p*-value = 0.037)[2,3].

{Insert Table 3 here}

In sum, it appears that the influence model (SINT) better predicted group decisions than other weighted opinion models (SJS, weighted majority, and mean wins). However, given the similarities in precision, and the same pattern of underestimating the group decision, it appears that SJS is essentially mimicking the mean wins model. This is explored further in Appendix 1. In general, and as expected, weighted opinion models outperformed SDS models in the continuous scale as predicted by H1.

---

[2] Bias is the signed precision of a model and indicates whether a model systematically overestimates or underestimates the group decision. In most other models, a systematic bias was not present.

[3] The bias of a mean wins model is equivalent to sample polarization times negative one if the sample is predicted to shift up; otherwise, if the sample is predicted to shift down, the bias is equal to sample polarization. In this case, the negative bias of the mean wins model indicates that the mean wins model systematically underestimated the group decision (definition of bias of a mean wins model). Stated differently, groups on average intensified their beliefs away from the mean wins model towards the high end of the scale (definition of sample polarization). These two statements are mathematically equivalent. Table 6 shows the sample polarization for the remaining tasks and initial constellation types, and therefore the bias of the mean wins model for the remaining tasks may be calculated by multiplying the sample polarization by negative one. The *p*-values will remain the same for tests of significance from zero. The grading tasks with two members equal was the only task and constellation type with a significant mean wins bias.

In the discrete scale, a majority wins model was predicted to be the best model. However, once again, SINT was the top model with an average precision score of 2.15 across both tasks, significantly outperforming majority wins (majority wins precision = 3.57; rank = third best model; *p*-value = 0.011). In the discrete scale, hitrates were much higher for the discrete models (NTCGP, majority wins, and the "extremization" models … high wins, low wins, and polarization wins), but SINT tended to be significantly closer *on average* to the group decision than all models except NTCGP. Therefore, unlike the continuous scale, the first hypothesis is rejected in discrete scales. Instead of behaviors based on social decision schemes and choice-type behaviors, groups seemed to arrive at decisions which were independent of the constellation of initial opinions and instead based their decisions more on the persistent interpersonal influences of the actors (i.e., SINT wins). A weighted opinion model (NTCGP) was also the second-best model in the discrete scale, indicating that many groups were trying to equally weight each members' opinions and settling on the nearest answer. In sum, H1 is partially supported: the continuous condition results were consistent with expectations, but the discrete condition results were not.

Because the continuous models consistently outperform the discrete models in both continuous and discrete scales, Table 3 provides initial evidence that groups prefer to employ continuous decision models as opposed to discrete models. This is contrary to H2, which posits that due to the added cognitive effort, groups should prefer to employ discrete models. This surprising finding is further supported by the results presented in Table 4. Because majority wins is a model of particular interest given its prevalence in the literature, once again only those groups with two members who have identical pre-discussion preferences are analyzed. Here, the performance of the various decision models are presented for the condition in which subjects were asked to provide both a continuous and discrete response. The results indicate that in general, the model fits for the continuous responses (best fit is weighted majority, precision = 1.41) are better than the model fits for the discrete responses (best fit is NTCGP, precision = 2.58[4]). Therefore, H2 is rejected. Results indicate that subjects prefer to incorporate models which utilize

[4] A t-test of the difference between 1.41 and 2.58 is inappropriate because the data more severely violates the non-independence assumption, and t-tests are particularly not robust to non-independence. In this instance, one the model precisions of 1.41 and 2.58 are derived from each group providing different answers to the same question, so independence is violated twice (once for each group, and once for the same question). No method of testing this difference currently exists (Wasserman and Davis, 1991).

continuous responses, and let they let the continuous responses dictate the discrete response of the group. Cognitive effort does not appear to deter groups from weighing and incorporating each others' opinions in a systematic fashion, though this may be because groups only contain three members and combining three opinions is a relatively simple procedure compared to larger groups.

{Insert Table 4 here}

The third hypothesis predicts that for risky tasks, groups will employ risky or conservative decision models in discrete scales, and weighted opinion models in continuous scales. Table 5 presents the results of the analysis, and once again, the prediction is supported in the continuous scale but not the discrete scale. In the continuous scale, SINT was once again the leading model, though its performance was not statistically better than the other weighted opinion models if two members had identical initial opinions, though it was statistically better than the weighted majority model (p-value = 0.003). In the discrete scale, NTCGP was the best model if two members started with the same opinions. However, median wins, a discrete model, had the best precision score if all members were unique. This model did not statistically outperform the NTCGP, SJS, or mean wins models, however. Further, the polarization, risky ("high wins"), and conservative ("low wins") models were consistently statistically worse than the weighted opinion model, contrary to H3. Thus, H3 is partially supported, where the continuous condition predictions were supported but the discrete condition predictions were not.

{Insert Table 5 here}

The fourth hypothesis was dependent on the discrete scales exhibiting discrete decision models. The results of the formal tests of this hypothesis are presented in Panel A of Table 6, and neither group polarization nor sample polarization is consistently present throughout the sample, which once again is consistent with the continuous but not the discrete condition. Groups with unanimous initial opinions were excluded from this analysis; with the exception of one group which exhibited slight polarization, the remaining groups selected the unanimous pre-discussion opinion as the group opinion.

{Insert Table 6 here}

The fourth hypothesis is conditional on the following two assumptions: first, groups would employ majority-wins behaviors in discrete scales, and second, the distribution of initial opinions would be relatively skewed, particularly in the risky settings. Tables 3 and 5 show that the majority did not pervasively "win" in the discrete scales, thus failing the first assumption. To examine if the second assumption would have been sufficient to induce polarization if the first assumption held, Panel B of Table 6 isolates those groups where a majority won to see if sample polarization was present in that subset of groups. In Panel B, groups who had two members with the same initial opinions were split into groups where the majority won and groups where the majority did not win.

Note that in groups of three, the median opinion will also be the majority opinion because two out of the three members agree on the same opinion. Therefore, a majority wins rule will simply correct the skew in the initial opinions as each group will settle on its median opinion, and overall, the average pre-discussion median opinion of all of the groups becomes the average opinion of all the groups. For example, consider two groups, each of whose members' pre-discussion opinions are (10, 10, 40). The median opinion is 10, and the average opinion is 20 for each group, and thus the entire sample. A majority-wins rule will yield a group decision of 10 for each group, and now the average of the groups' opinions (group A is 10 and group B is 10 = average opinion of 10) equals the pre-discussion median (10). A test of sample polarization in this instance is merely a test of the significance of the skew: if the correction of the skew was significant, then skew was big enough to cause sample polarization. Panel B indicates that, for those groups who exhibited majority wins rules, the skew was only significant in one condition (continuous, risk-free: sample polarization = 2.56, *p*-value = 0.007). This is contrary to the results of Laughlin and Earley (1984) who argue that CDQ tasks should cause a skew towards one end of the scale and thus induce sample polarization. In sum, H4 failed because neither assumption of majority wins nor the skew of the sample was met.

The relative failure of majority wins to explain group decisions throughout this study was surprising given the amount of research which finds support for such a decision rule (see, for example, Hastie and Kameda, 2005). However, in many studies, the decision alternatives are mutually exclusive (e.g., choose between candidate A, B, C, etc.) and/or dichotomous in nature (e.g., guilty or not guilty). One possible explanation for the ubiquity of majority wins in those situations is that the group cannot find "middle ground" and must therefore engage in conformity as the next best alternative to making decisions.

Conversely, if an alternative that lies between the majority opinion and the minority exists, groups appear likely to choose that alternative and engage in more central tendency behaviors which incorporate the opinions of all group members. If this is the case, then the standard deviation of the initial opinions should be inversely related to the likelihood that the group will choose the majority opinion.

To examine this possibility, a logistic regression was tested with a successful majority wins hit coded as 1 (successful) or 0 (not successful) as the dependent variable, and the standard deviation of the initial opinions, task, and scale condition (discrete or continuous) as independent variables. All five tasks were included in the model, and only those groups with two equal initial opinions were analyzed. All interactions were insignificant and were dropped from the model[5]. The main effect of standard deviation was negative and highly significant (coefficient = -0.1620; Wald chi-square p-value < 0.0001), as was condition with the continuous condition coded as 1 (coefficient = -0.8526; Wald chi-square p-value < 0.0001). The intercept of the model was 1.7458, and the model was able to correctly classify 74.5% of observations. The results are consistent with the hitrates in Tables 3 and 5 which show that groups in the continuous condition were less likely to settle on the majority opinion. Additionally, and consistent with the above prediction, the farther away the outlying opinion was from the two equal opinions, the less likely the group was to settle on the majority opinion.

Discussion

Results of the tests indicate that groups of three members do not employ discrete decision models based on social decision schemes, but instead employ continuous decision models which involve the systematic weighting of group members' opinions. This decision setting is different than previous papers testing social decision schemes because in this study, the decision alternatives were ranked, and therefore groups were not forced into using social decision schemes because no other alternative exists, such as in the case of many categorical decisions. The results of this study suggest that social decision schemes are in

---

[5] The variance of the initial opinions was also included in the model to test for nonlinear effects of distance between alternatives. The variance is equal to the square of the standard deviation; this is analogous to including, for example, assets and assets squared in a logistic regression model. The coefficient on the variance was not significant and was dropped from the model.

fact used by groups because they have no other choice in other studies, as opposed to groups using social decision schemes because they are an attractive cognitive shortcut for decision-making. The results of the study also indicate that the scale of the decision outcome does not change group behaviors as suggested by the group decision making literature (e.g., Hinsz, 1999).

The first hypothesis predicted that groups would employ majority-wins rules in discrete scales and some weighted opinion model in continuous scales. In both scales, groups employed an influence-based weighted opinion model. The second hypothesis predicted that when both a discrete scale and continuous scale were available, groups would use the discrete scale to arrive at a decision because discrete scales require cognitive effort. Instead, groups used weighted opinion models which are more consistent with continuous scales. When making decisions involving risk, the third hypothesis predicted that groups would use a risky or conservative model in discrete scales but a weighted opinion model in continuous. Once again, groups used weighted opinion models in both scales. The fourth hypothesis predicted that a majority-wins rule would lead to sample polarization in discrete scales, but the weighted opinion models that were actually used in the discrete scales were partially responsible for the lack of sample polarization.

Perhaps the most surprising result from the study was the poor performance of majority wins in the discrete scales. Majority wins has enjoyed much empirical success in the literature, but two reasons may exist for the lack of support for majority wins in this study. First, this study used groups of three people which tend to be smaller than many of the studies examining social decision schemes. As group size increases, conformity pressures also increase, and this may lead to more majority-wins type decisions. Second, in many cases, the majority wins model is tested by dichotomizing a scale with multiple decision alternatives (for example, Laughlin and Earley (1982) collapsed a 10-point probability scale to a yes / no response). The loss of information from such an approach may be significant as the distance between the majority opinion and the dissenter was found to be inversely related to the likelihood that the majority opinion would be the final group opinion. If a substantial difference between the two alternatives exists, then groups are likely to pick an alternative that lies between the two opinions. The accuracy of central tendency models suggests the groups are likely to settle on the alternative that equals the mean of the group members' initial opinions. This inability of a majority wins model to overcome the distance between initial opinions warrants further attention by researchers. For example, the Norm-Information-Distance model (Crott et al., 1996) may provide some insights into such results.

Classifying decision models based on their theoretical origins (e.g., SDS, SDS-Q, SJS, SINT) leads to an unclear picture of group decision making. This is evident throughout the tables presented in this study because consistent patterns among performances in the decision models are difficult to identify, with the exception that SINT appears to be more successful than the others. However, a much clearer pattern emerges when the models are classified by the group behavior that is exhibited if the groups employ the decision models. In other words, by identifying possible group behaviors and then identifying which models are consistent with those behaviors, the results of this study shed substantial insight into group decisions. These group behaviors, and the corresponding models, are discussed below.

Extremization

Some evidence exists that groups may choose the most extreme judgment or alternative that is favored by one of the group members. For example, a small decision-making group may distinguish itself from a larger population of decision-makers because of feelings of social identity. Therefore, other group members side with the member that is most prototypical of the small group's position relative to the remaining population, and this prototypical person may be the most extreme group member (McGarty et al., 1992). Sometimes an opinion assumed by a group member may have a significant informational advantage over other opinions, and this "rhetorical asymmetry" may cause an extremization of the group opinion, perhaps even past the most extreme member's initial opinion (Schkade et al., 2000; 2004). The groupthink literature documents many cases of groups pursuing extreme alternatives (Janis, 1982; Turner and Pratkanis, 1998). Models which are consistent with extremization are risk wins, conservative wins, and polarization wins.

Factional influence

Faction models assume that group members align themselves with other group members who share the same or similar opinions, and together those group members are able to pull the group decision in their preferred direction. From an informational standpoint, their ability to influence the group may arise because groups tend to search for information that supports group consensus (Schulz-Hardt et al., 2000), and because dominant factions contribute most to group discussions (Parks and Nelson, 1999). Therefore, the group conversation is likely to be driven by those members who already agree discussing why they agree, and the information pool will favor their initial opinions. Or, from a social comparison standpoint, once group members see that several members agree on a course of action, the group preference becomes known, and they are likely to shift their opinions in that direction to make themselves appear more favorable to the group (Baron and Roper, 1976). Models which are consistent with factional behaviors are majority wins, SJS, and weighted majority.

Central tendency

Central tendency models are intuitive models that suggest that the group decision will gravitate towards the center of the group members' initial opinions. These models accommodate all members' opinions and may potentially avoid conflict within the group by equally valuing each members' opinion (although not necessarily equally *weighting* the opinions, i.e., as in the median model). Models which are consistent with central tendency behaviors are mean wins, NTCGP, and median wins.

Idiosyncratic influence

Idiosyncratic influence models assume that the group decision is determined by the interpersonal influences of the actors, and therefore the constellation of initial group opinions is not helpful in predicting the ultimate group decision. These influences may be derived from members observing the expertise of other members and adjusting the importance of the experts' comments when contributing their own comments. Because these models allow for the identification and incorporation of expert opinions, these are sometimes referred to as "rational" models of decision-making (e.g., Lehrer and Wagner, 1981). Several such models exist (for discussions, see Davis et al., 1997b; Friedkin, 2003). This study examines the structural influence network model (SINT) which is related to many of those models (Friedkin, 2003).

Table 7 presents the results in terms of group behaviors. Here, a clear pattern can be seen. In continuous scales, idiosyncratic influence behaviors dominate (SINT), followed by central tendency behaviors (mean wins) and faction-based models (SJS). However, as discussed in Appendix 1, SJS essentially functions as a central tendency model when measured nominally as opposed to percentages as was done in this study. Therefore, influence models dominate followed by central tendency models. This means that if the network of interpersonal influences is not known, then the average of the group members' opinions is the next best estimate of the group's decision.

{Insert Table 7 here}

In the discrete scale, the pattern is less pronounced, but some behaviors do clearly emerge. Surprisingly, influence models are not as dominant if everyone holds a unique pre-discussion opinion, but they do dominate if two people agree on a course of action before discussion. Once again, though, if the network of interpersonal influences is not known, then central tendency behaviors will result in the next best estimate of the group's decision. Faction behaviors do begin to emerge in the discrete scale, but they do not dominate the central tendency behaviors.

The above analysis supports the principal finding that the scale of the decision outcome does not affect group outcomes as expected, despite prior research that shows scales may change the actual discussions and exchange of information that occurs (Gigone and Hastie, 1997). Further research may attempt to identify when, or if, faction behaviors do exert more influence on group decision outcomes than central tendency behaviors. Manipulating group size appears to be the next step in isolating these behaviors.

CHAPTER 5

STUDY 2

Background

The second study incorporates the importance and distribution of information in an auditing decision-making environment in addition to the scale effects. The research question asked in the second study is "How does scale affect group decision-making processes and outcomes in an accounting setting, and what effect does the distribution of information among group members have on these processes and outcomes?" According to the first study, auditors are predicted to employ SDS models in discrete environments and continuous models in continuous environments. However, the nature of the accounting decision-making setting might cause an increased focus on the information content of the discussions, causing auditors to employ SDS models which more closely resemble the compensatory models used in continuous scales. These models are predicted to be different than the SDS models used in tasks in which normative pressures are more prevalent.

Scale has been found to affect the way that groups incorporate information into the decision-making process. Research has shown that groups in discrete scales tend to focus on fewer pieces of information and assess higher importance to the remaining cues that are considered (Gigone and Hastie, 1997), weakening the influence of informational cues. In a group choice setting (i.e., under discrete scales), auditors are expected to have a more defensive posture while justifying their choices during the group discussions. This is expected to focus the discussion on confirmatory information. However, in continuous scales, the discussion is expected to examine the weights of confirming and disconfirming information. Therefore, in continuous scales, auditors are expected to discuss more information when making judgments.

The distribution of information has also been found to affect group decision-making. Specifically, information which is shared by all group members tends to be discussed more and more heavily weighted by groups (Gigone and Hastie, 1993; Winquist and Larson, 1998; Chernyshenko et al., 2003). Thus, shared information will dominate the discussion, though the weighting process of a continuous environment should cause auditors to more closely examine both shared and unique information. Thus, auditors are predicted to more completely discuss and weigh information in continuous environments compared to auditors in discrete environments.

This information distribution can create conflicts in combinatorial processes such as SDS if the unshared information suggests members should change their initial assessment. Groups may prematurely end discussion in discrete scales if members agree on a decision; this is more likely to occur in discrete alternatives because the probability that members have unanimous initial preferences increases as the number of alternatives decreases. If the information is distributed in such a way that conflicts arise between the initial preferences and the "correct" alternative based on the entire information pool (that is, if auditors are given different pieces of information which individually appear to support the one alternative, but when all information is pooled, the information in actuality supports another alternative), then groups in discrete alternative choices are predicted to be less likely to change their opinions to the correct alternative than groups in the continuous scales.

Theory and Hypothesis Development

As discussed in the first study, the task type is likely to have an effect on the decision processes that are employed by interacting groups. Different influences will affect groups in different situations. Groups solving intellective tasks – tasks with demonstrably correct answers (McGrath, 1984) – tend to focus on informational cues rather than normative pressures (Kaplan and Miller, 1987), because the group can identify the correct response given the information set. Consensus-based SDS rules are less important in such situations despite their appeal as a more cognitively efficient decision process, and instead groups rely on truth-wins or truth-supported rules in discrete scales and polarize in continuous scales once the correct answer is identified.

However, judgment tasks – tasks without demonstrably correct answers – tend to focus on normative pressures because the group has no logical criteria for establishing a correct answer (Kaplan and Miller, 1987). Here, consensus-based SDS models are anticipated to be effective models of group decision-making. The importance of information in judgment tasks is secondary to normative influences because groups have difficulty finding a correct decision to compare their favored decision against. Instead, groups adopt a "safety in numbers" approach (i.e., judgment consistency) to make group decisions.

The traditional characteristics of intellective and judgment tasks suggest that a group of auditors trying to assess the riskiness of a client are performing a judgment task, and therefore, the psychology literature would predict that normative influences should dominate the decision-making processes. However, although there is no *demonstrably* correct risk or probability assessment which the auditors should arrive at during the discussion, a group of auditors assessing the likelihood of possible alternatives should focus on the facts of the situation when making a group judgment because a correct solution does exist, even if it is not demonstrable *ex ante*. For example, when assessing the cause of a wayward ratio calculated during the analytical review, a true cause for the unexpected ratio does exist. Likewise, auditors debating whether to issue a going-concern report should consider and debate the facts of a client more rather than trying to focus on consensus, because ultimately the going-concern status of the company will be known. Therefore, while the decisions being made by auditors exhibit the characteristics of judgment tasks, the nature of the audit setting will cause groups to exhibit processes that are more characteristic of intellective tasks and will therefore focus on the information exchanged.

Such an informational focus should lead to the incorporation of the minority opinion into the group decision, where the minority opinion is the opinion farthest from the median opinion in a three-person group. Therefore, under a discrete decision outcome scale where groups much choose a decision outcome, an SDS model that incorporates the minority decision into the final group outcome is likely to be selected by the group. The two SDS models which do that are the mean-wins, or the nearest to the center of gravity of proportionality wins. In continuous decision outcome scales, the group will employ a continuous decision model which accounts for all opinions. However, whether SINT, SJS, or an arithmetic mean-wins SDS-Q is employed is unknown. This leads to the following hypothesis stated in alternative form:

> **H5a:** In an audit task with fully shared information and a discrete decision alternative, groups will employ either a mean-wins or nearest to the center of gravity of proportionality SDS model.

H5b: In an audit task with fully shared information and a continuous decision alternative, groups will employ a weighted opinion model.

In study 2, the predicted SDS models imply that no sample polarization should occur if the responses are symmetrically distributed around the median alternative in the discrete scale. Likewise, sample polarization is not predicted in the continuous scale *ex ante*. This is consistent with the lack of sample polarization found in Schultz and Reckers (1981) and Karan et al. (1996). As discussed above, the sample polarization observed in Chalos's (1985) study was in a bank loan setting which did not include the competing risk factors that an audit setting faces (i.e., balancing litigation risk with client retention and audit costs), and the task setting may have contributed to Chalos's (1985) subjects' polarization. Carpenter's (2005) sample polarization appears to be driven by the differing statuses of the group members; this study will be conducted with homogenous groups of auditors. An alternative prediction, however, is that the materiality of the task may cause conservative shift similar to Bamber et al. (1996) and somewhat consistent with Reckers and Schultz (1982), in which case a conservative-wins model may or a conservatively-weighted SDS-Q model may better explain the results.

The decision scale will determine whether the group approaches the problem as a group choice (discrete) or a group judgment (continuous). Gigone and Hastie (1997) find that in group choice settings in intellective tasks, group members discuss less information; this may be due to an elimination-by-aspects strategy (Tversky, 1972) in which pieces of information which do not meet a certain relevance threshold are discarded by the group members. However, given the informational (rather than normative) focus of audit tasks, groups should consider more information regardless of the decision scale before making a decision. While H5a predicts that group members will more fully incorporate the minority opinion into the final decision, groups will still be discussing the decision alternatives as group choices rather than group judgments in discrete scale environments. To arrive at a choice, an individual must consider evidence which supports his or her position, and therefore the information set that each group member brings to the discussion will tend to be supportive of his or her choice. In this sense, the group discussion serves as a forum in which group members must justify their opinion because justifiability of one's opinion is a means by which group members may judge the expertise of each other (Peecher, 1996; Kennedy et al., 1997).

If a person makes a judgment on a continuous scale, the justification is likely to be less in defense of a choice and will instead focus on how the group member weighed the various pieces information against each other. Therefore, the information set relevant to each individual's opinion is larger, and the amount information discussed among group members will be greater because each individual will consider both confirming and disconfirming information. Thus, more information is predicted to be exchanged in continuous as opposed to discrete scales. Additionally, when all group members have the same initial opinions, groups tend to truncate the discussion (Gigone and Hastie, 1997). This is more likely to occur in discrete scales where the number of constellations *m* is small, particularly if the evidence suggests that some alternatives are more appropriate than others (e.g., a conservative option is more appropriate than a risky option). In such cases group discussion is predicted to be truncated due to the lack of dissenting opinions to debate. Therefore, hypothesis six, stated in alternative form, is:

> **H6:** Groups given discrete decision alternatives will discuss less information than groups given continuous outcome alternatives.

H5a states that groups will engage in more compensatory SDS models which would reflect the incorporation of the minority group member's opinion, suggesting that informational influences trump normative influences in accounting settings. This appears inconsistent with H6 which predicts that groups in discrete scales will share less information than they would if they were in a continuous scale. However, the two predictions will be consistent with each other if the group discussion focuses on the justification of the minority opinion as opposed to a justification of the majority opinion. Because the majority opinion already has "consensus support", the discussion is likely to be limited to attempts by the minority opinion to persuade the majority opinion to accept, or at least incorporate, the minority choice.

When information is distributed among group members so that group members have both common and unique information, the shared information is likely to dominate the group's discussion and influence the group members more than unique information (Chernyshenko et al., 2003). In continuous scales, because the decision processes for individuals will likely involve a compensatory strategy between information cues, the subsequent group discussions regarding the individual judgments will likely focus on the balance between pieces of information. Thus, the discussion should reveal more pieces of information whether the information was common or unique to the individuals. However, in discrete scales, because the group members may be justifying their choices more, the discussion is likely to focus on only those pieces of information which support their opinions.

If the information is distributed in such a way that the shared information supports one end of the decision spectrum (e.g., low risk versus high risk), then auditors in discrete scales are less likely to discuss unique information which contradicts the initial tendency. This is because the minority opinion, if it exists, will be the only party which needs to justify his or her opinion, and therefore the unique information discussed will be that held by the minority opinion, as that information was persuasive enough to turn the individual's decision against the majority opinion. The remaining unique information is less likely to be discussed. Such a distribution of information may be encountered in practice because unique information may be discovered through individual audit testing or discovery, while common information may be found from more public sources or through discussions with management who have an incentive to attain a favorable opinion from the auditor. Therefore, the following hypothesis is put forth in alternative form:

> **H7a:** Unique information will be discussed less by groups making decisions in discrete scales than by groups making decisions in continuous scales.

If the unique information contradicts the shared information, then the added discussion of unique information in continuous scales will cause groups to revise their decisions more to incorporate the opposite effect of the unique information. Further, unique information, even when discussed, is less influential than shared information (Chernyshenko et al., 2003). Therefore, more unique information must be discussed to have an effect on the group decision, and this is more likely to occur in continuous scales. This leads to the following hypothesis in alternative form:

> **H7b:** Unique information will be incorporated into the group decision less by groups making decisions in continuous scales than by groups making decisions in discrete scales.

Method and Participants

The experimental design was a 2 × 2 setup utilizing both between-subjects and within-subject conditions. Subjects were given a hypothetical company – the company developed by Asare (1992) – and were asked to complete two tasks (within-subjects; see Appendix 3). The first task involved subjects assessing the likelihood of a material misstatement in the accounts receivable balance which was based on the Brown and Solomon (1991) task; this was supplanted into Asare's (1992) company. Subjects were given the results of some audit procedures which have already been completed and the results (for example, analytical ratios), and a list of those not yet completed (for example, verification of individual transactions), and were asked to make a likelihood assessment that the balance of the account could be misstated. In this task, all subjects had all of the same information. The between-subjects manipulation was again scale. In one scale condition, subjects were given a discrete probability scale with five outcomes (0.10, 0.30, … , 0.90), and in the other, subjects were asked to make a probability assessment, though the maximum and minimum probabilities they may assign will be 90% and 10% to maintain consistency (and comparability) with the first task[6].

In the second task, subjects were asked to give a recommendation to the audit partner regarding whether a going-concern opinion should be issued in the audit opinion. The task utilized the twelve information cues developed by Ashton and Kennedy (2002) for Asare's (1992) case. However, information was distributed in such a way that each person receives eight cues, six of which indicate the company should not receive a going-concern opinion, and two cues which might suggest such an opinion is warranted. The two 'mitigating' cues were uniquely held, so of the total information set, six shared cues suggest the company is not a going concern risk, while six unique cues suggest it is a going concern risk. Subjects were asked to assess the likelihood that they would recommend a going concern opinion to the engagement partner on a discrete probability scale with five outcomes, or on a continuous scale bounded by 10% and 90%, depending on the scale condition in which they were placed.

Subjects were 93 volunteers from public accounting firms in Missouri, Kansas, and Nebraska giving 31 groups of three. Big 4, regional, and local firms all participated in the study, and the study was administered at each of the participating firms' offices[7]. Whenever possible, groups were composed of

---

[6] Subjects were also asked to estimate how many staff hours should be budgeted to complete the accounts receivable audit. This information is not analyzed to investigate the current hypotheses.
[7] The study was administered at the University of Missouri for two groups who were visiting the campus for other events.

accountants with identical organizational levels (e.g., partners, managers, etc.)[8]; however, due to subject requirements this was not always possible. Descriptive statistics of the study participants are presented in Table 8[9]. The table shows that roughly two-thirds of the sample consists of staff and seniors who have on average 4.2 years and 1.4 years of experience respectively. This level of experience is appropriate for the accounts receivable task, but decisions regarding going concern opinions are usually made at the partner level. This experience level was anticipated, and so the question for the going concern task asked for the likelihood that the individual or group would recommend a going concern qualification to the engagement partner, rather than asking for the likelihood that the individual or group would actually issue a going concern opinion. Table 8 also indicates are somewhat high level of "other" audit opinions issued by managers and partners; this was driven by partners in the local firms (who represent 33% or (6+2)/(16+8) of the combined managers and partners in the sample) who work on a substantial number of government audits. Finally, four out of the nine "unequal" groups – that is, groups that did not consist of three individuals at the same organizational level – occurred in local firms because of the availability of participants at those firms.

{Insert Table 8 here}

Results

H5a (H5b) predicts that a mean-wins or NTCGP (weighted opinion) model will be employed by groups making decisions in a discrete (continuous) scale. To test these hypotheses, the various models are tested as was done in Study 1 where the primary test of model performance is the difference in model precision, but hitrate and bias are also presented. Results are presented for the accounts receivable task in Table 9. Consistent with the strong results of Study 1, SINT is again most accurate at predicting group decisions in both scale settings, indicating that the persistent network of interpersonal influences is most important in group decisions. However, surprisingly, the next best model in both scales was the majority

[8] Two senior managers participated in this study. Their descriptive statistics more closely matched the descriptive statistics of the partners, and therefore, they were classified as partners for these analyses.

[9] Due to additions to the post-experimental questionnaire which occurred after the first four groups had completed the study, some descriptive data was not collected for the first four groups.

wins model, indicating precisely the opposite predictions of the hypothesis. Rather than focusing the information context of the task as predicted, groups chose the opinion favored by a majority of the group members (or, if none agreed, the middle group members' opinion was selected. This occurred 40% of the time in the continuous scale setting, and 75% of the time in the discrete scale setting. It appears that groups believed that a consensus of group members' opinions prior discussion was the best judgment in the absence of a demonstrably correct answer, consistent with Festinger's (1954) original theory of social comparison. Also noteworthy is the lack of a significant bias of the mean wins model in both scales; this indicates that no sample polarization occurred.

{Insert Table 9 here}

One possible explanation for the above results is that groups are settling on the median opinion when all group members begin the discussion holding unique initial opinions. However, when groups with unique initial constellations of opinions are excluded (not tabled), 5 groups remain in the continuous condition, and majority wins predicted the group decision 60% of the time (3 of 5 groups), which was not unexpectedly tied with polarization wins. Given the small sample size, statistical tests are not appropriate on the relative precision, but majority wins is still the second best model with a precision of 5.0 (also tied with polarization wins) compared to SINT's precision of 3.2. In the discrete scale, 11 groups remain, and 82% of them used a majority wins rule, which was the best decision model. Again, SINT had the best precision of 1.3, but majority wins was second with 3.6. The next closest model was weighted majority with a precision of 5.1. While the groups with unique constellations did contribute some of the accuracy of the majority (or in this case, median) wins model, the majority wins model still performed well relative to other models when groups with unique constellations were excluded.

In both Study 1 and Study 2, scale has been shown to not significantly alter decision outcomes for groups of three, contrary to predictions drawn from the psychological literature. Given these results, and particularly the success of decision models based on social influences as opposed to informational influences, it is less likely that differences will be observed in the amount of discussion of individual pieces of information, or the importance of the cues on the group decision, between scale conditions. After completing the task, each group member was asked (individually) to rate both how much each of fourteen cues was discussed, and how important each cue was in the group decision, on a scale of 0 (not discussed at all or not important at all) to 4 (discussed extensively or extensively affected the group's decision). Each cue's importance and amount discussed was calculated as the *average* evaluation of that cue for each group, and therefore there are 15 observations in the continuous condition and 16 observations in the discrete condition. Tests of H6 are presented in Table 10, and as expected, few differences are observed between conditions for individual cues. The remaining tests refer to the going concern task.

{Insert Table 10 here}

Panel A shows that generally, cues are equally discussed between conditions, with a few but unremarkable marginal differences. The lack of significant differences between the conditions for the cues in general rejects H6 (two of the three significant differences are in the wrong direction). A more formal test of H6 is presented in Panel B, where the total discussion for all cues is presented by continuous and discrete scales (average cue discussion for continuous = 1.38; average cue discussion for discrete = 1.34; *p*-value of difference = 0.82). Therefore, H6 is rejected. Further, H7a predicts that when the cues are separated by common and unique cues, the unique cues should be discussed more in the continuous scale. Panel B indicates that on average, cues were discussed less in the continuous scale (average discussion = 0.80) than in the discrete scale (average discussion = 1.03), but this difference is insignificant (*p*-value = 0.24). Thus, H7a is also rejected.

When subjects' perceptions of the information cues' importance is analyzed, once again differences in individual cues are not systematically observed (Table 11, Panel A). Some marginal differences emerge between conditions when cues are pooled by common and unique cues, as participants in the continuous condition believed the common cues were more important to the group decision relative to subjects in the discrete condition ($p = 0.08$, Panel B). However, H7b hypothesizes that the differences will emerge in the conditions regarding the unique cues, and subjects rated the importance of the unique cues nearly identically in both conditions ($p = 0.99$) rejecting H7b.

{Insert Table 11 here}

Many of the findings in prior research regarding cue importance are based on factorial designs which make it much easier for researchers to isolate the effects of individual cues on decisions. To validate that the self-reports of the cues in this study are adequate measures of the incorporation of cues by the groups, this data set is tested to examine two well-known biases regarding individuals' contributions to group discussions when unique information is held. First, prior research finds that subjects tend to remember their unique contribution to group discussions more than other group members remember those contributions. If that occurred in this study, then individual group members should believe that their unique cues were discussed more, relative to the assessment of the same cues by other group members. Second, subjects tend to think their uniquely held cues are incorporated into the group discussion more, and if so, they should think their unique cues were more important than other members' thought those cues were.

Subjects' own relative assessments are calculated by averaging the subjects' assessments of the two cues that are unique to that subject. This is done for each of the dependent variables: the amount discussed and the importance on the group decision. Then, the other group members' assessments of those same cues are averaged, and the difference between the owner's assessment and the other members' assessments are subtracted. A positive number indicates that the owner thought the two cues were discussed more (or were more important to the group discussion), and this would be consistent with both biases. A negative number indicates that the other members thought the cues were discussed more or was more important. Results of tests of the predicted biases are presented in Table 12.

{Insert Table 12 here}

Results show that the self-contribution bias is robust in this sample, as individuals thought that their uniquely-held cues were discussed more than their other group members did. Regarding the perception of the importance of uniquely-held cues, in the continuous scale, subjects did not believe their uniquely-held cues were more important to the group decision than the other group members did, although in the discrete scale subjects did believe that[10]. However, given these results, it appears that subjects do

---

[10] Note that these two biases are analogous to the "endowment effect" in which a good (in this case, information) attains more value to the individual once the good is obtained by the individual (Thaler, 1980). However, the endowment effect is thought to be one manifestation of prospect theory in which the loss of that good will cause an asymmetric loss to the subject than the value gained by the subject when the good is obtained (Kahneman et al., 1990). In this case, the sharing of unique information reduces its value as an asset to the individual, and therefore the individual has incentives not to reveal her unique

believe their unique cues are discussed more, and are somewhat more important to the group than other group members do, and therefore this experimental design yields results which are relatively consistent with the literature on the contributions of uniquely-held information. Thus, the self-reports appear to be valid, and the overall results obtained regarding cue usage in this study appear to be driven by the consistent finding that decision outcomes are more consistent with group processes in which social influences are more important than informational influences.

Discussion

In Study 2, SINT was the best-fitting model followed by faction-based models as described in the discussion section of Study 1. This indicates that persistence interpersonal influences dominate group decision making by experienced accountants, even in information-rich settings in which the discussion is expected to focus on the information exchanged rather than social influences. Had information played a more important role, then groups would have been expected to employ more central tendency models. If the network of interpersonal influences is unknown, then a majority / median wins decision rule is the next best estimate of the group decision. Contrary to theory, the scale manipulation did not substantially alter model performance or group outcomes, and this was verified by the analysis of cue usage between the two scale groups.

The results of this study are most consistent with Carpenter (2005) who finds that organizational level is most important in group decision-making by accountants. In her study, the managers had the most influence, and they were able to pull the group decision in their favored direction which is consistent with the tenets of SINT. One alternative explanation (and the explanation that Carpenter gives) is that group polarization occurred; however, the results of both Study 1 and Study 2 of this dissertation find that group

---

information. The results in Table 12 indicate that the individuals thought their private information was in fact discussed more and more important to the group, suggesting a willingness of the individuals to share that information. Investigating when an individual would share that information is a potential direction for further research, likely from an experimental economics standpoint. If the value of the unique information to the group outweighs the value loss to the individual for revealing the private information, the individual may still not reveal the private information if the individual does not adequately realize her portion of the benefits obtained by the group. The goal of the group (in this study, answering questions as if trying to conduct an effective audit) and the strength of the members’ sense of group identity may moderate these effects.

polarization is not ubiquitous in groups of three people. Further, given the results of other papers in accounting which were not able to find group polarization (Schultz and Reckers, 1981; Reckers and Schultz, 1982; Karan et al., 1996), this study fits with other accounting studies which find that group polarization may not be reliable in accounting settings.

This study appears most inconsistent with Chalos (1985) who finds that information is the most important factor in group decisions. However, it is possible that Chalos is interpreting the results of his experiment incorrectly. First, Chalos argues that groups exhibit a truth-wins model, but a requirement for a truth-wins model is that the correct answer is demonstrable during the group discussion (McGrath, 1984; Laughlin and Hollingshead, 1995). In Chalos's study, a correct answer existed because the bank loan tasks were based on real situations (in which some of the loans did actually default), but the answer was only known *ex post* and was not demonstrable during the group discussion. Second, Chalos dichotomized his scale similar to Laughlin and Earley (1982), and as demonstrated in Study 1 the loss of information from such a procedure may lead to incorrect conclusions because groups may actually compromise (particularly if an outlying member's opinion is far away from the other group members' opinions), but this effect is hidden in the dichotomization of the scale.

Recently, FASB has required group discussions to occur throughout audits. Presumably, FASB issued these requirements in an attempt to increase audit effectiveness by enlarging the size of the information pool which is salient to auditors during the decision-making process. However, a large body of psychological research regarding the sharing of information suggests that groups will not incorporate information optimally into the group decision optimally, particularly if some of the information is uniquely held (Stasser and Titus, 2003). Therefore, group discussions may not have the effect that FASB presumably intended. The results of this study are consistent with the notion that groups are not pure information processors as group decisions in this study were guided more by social influences than by informational influences, and people with outlying opinions were drawn into the majority's viewpoint. The lack of sample polarization (not tabled; sample polarization was insignificant in all tasks) suggests that as a whole, groups did not identify a correct answer from the scenarios presented to them, and therefore the groups were shifting in random directions rather than towards a universal "correct" answer.

Because these groups were making decisions which had no right-or-wrong answers, the accuracy of the group decisions can not be evaluated in this study. Two important studies with contradictory results suggest that social influences should be further examined in terms of decision accuracy. As discussed already, Carpenter (2005) finds that social influences guide decisions, but in her study the social influences were correct only half the time, and therefore the groups were not able to correctly overcome those influences. On the other hand, Hastie and Kameda (2005) demonstrate that majority-wins rules can lead to optimal decisions in simulations of resource harvesting.

Other potential areas of further research include investigating how group members attain influence. Given the success of SINT, this appears a worthwhile pursuit. Clearly, in Carpenter's (2005) study, organizational level bestowed influence on group members. However, organizational behavior research shows that a variety of factors such as gender, organizational status, experience, SES, race, etc. may affect group dynamics, and research may attempt to map these to the endogenous network of influence in groups. Further, accounting-specific characteristics such as experience or perceived expertise (Libby and Luft, 1993) may be traced to influence in groups.

In this study, the tasks were presented in a specific order. Subjects were first asked to complete the accounts receivable task which contained full information. Then, subjects were asked to complete the going concern task in which the relevant cues for that task were distributed differently among group members. This order was chosen because groups would most likely not encounter the experimental manipulation of unique cues until the last question that was asked. If groups discovered that each person held unique information in the first task, then groups completing the second task might spend time searching for another experimental manipulation that did not exist rather than discussing the case at hand.

The purpose of the unique cues in the experimental design was to investigate which information was mentioned during the group discussion without physically monitoring the group discussion (which may significantly change group behaviors). Groups that discovered the unique information are assumed to have engaged in a more informational-focused discussion, and therefore, groups were not informed *a priori* that some members held unique information. However, in practice, auditors would know that each person holds some pieces of unique information such as the results of some audit tests when entering group discussions, and it is possible that groups may deliberately seek certain pieces of unique information when making their decisions. Note, though, that the holder of the unique information may not be aware that some of her information is unique, such as in the current study. Nonetheless, this experimental design may impose a limitation on the generalizability of the results.

Another limitation of the current research is the use of groups which consisted of people with equal organizational status. An analysis of the importance of certain cues on group decisions (not tabled) shows that certain groups of auditors thought certain cues were more important to the going concern task. For example, the fact that shareholders were willing to invest additional capital if needed was rated as very important by partners (importance = 3.56 on a scale of 0 to 4), but that cue's importance decreased as organizational level decreased (managers = 2.83; seniors = 2.14, staff = 2.27). A similar pattern is observed for the importance of the doubling number of customers (cue 14; partners = 2.23, managers = 1.56, seniors = 0.95, staff = 1.44). These cues refer more to the company's external environment and are therefore likely more important to partners than to audit staff, whose experience is mostly with testing internal accounts. Interestingly, partners were less concerned with current operating losses, a cue whose importance generally increased as organizational levels decreased (partners = 1.13, managers = 1.39, seniors = 2.67, staff = 2.33). Therefore, within groups, it appears that experience dictated which cues were discussed and evaluated more, and having groups consisting of equal organizational levels may further decrease the importance of certain unique cues if those cues would not be considered as important when evaluated by members of that organizational level or intensify the importance of the unique cue if it is valued. Further, and consistent with Larson et al. (1996), if a lower ranking member of an audit team holds a unique cue, the importance of that cue will likely carry more or less weight depending on whether the higher ranking member values that cue. In short, the importance of unique cues may be moderated by the organizational level of the group members who are evaluating that cue.

While some audit decisions may be made by auditors of the same ranks (for example, partners discussing client acceptance or rejection decisions), in many actual auditing situations, and particularly the ones represented by the tasks that the subjects completed, groups are likely to be composed of auditors of different ranks. However, this study provides a baseline examination of the group dynamics in auditing situations, and therefore manipulating ranks may be a subject of future research which can isolate the effects of rank (for example, does majority still win with two audit staff members and a manager if the two staff members agree?) or other variables on social influence.

CHAPTER 6

DISCUSSION AND CONCLUSIONS

The extant body of group decision making literature is both vast and complex. The complicated theory and hypothesis development sections of Studies 1 and 2 illustrate that finding consistent patterns in the research and establishing predictions are difficult, though several studies imply or explicitly suggest that one potential reason for the varied findings is the scale of the decision outcome that is utilized in the research. This dissertation is the first study of which I am aware that experimentally manipulates scale, and the results of the experiments show that scale is likely not the cause of substantially different group outcomes. Results of the hypotheses tested in this dissertation are presented in Table 13. Rather, the results of Study 1 and Study 2 show that a remarkably simple pattern in group decision making emerges for groups of three making decisions under uncertainty, regardless of scale.

Using the classification presented in the discussion section of Study 1, where decision models are classified by the underlying group behavior that is exhibited (extremization, faction, central tendency, or idiosyncratic influence), groups are shown to exhibit consistent behaviors. Most importantly, of the factors examined in this dissertation, the *persistent* network of interpersonal influences is the most determinant factor in group decision making. Because subjects were asked to identify the network of influence (i.e., distribute the poker chips) after completing several tasks, the network of influences obtained reflect the influence over those tasks rather than just one single task. If influence data was collected after each task, the SINT model would simply reflect subjects' abilities to identify the network of interpersonal influences that occurred in that task, and the network data would not necessarily have any predictive value. However, prior research has shown the network to be persistent (Friedkin and Johnsen, 1999), and the data collected in this study reflects the persistency of the network. The success of SINT in both Studies 1 and 2 indicates that knowledge of the network will allow an observer to more effectively predict the group decision than if the observer utilizes one of the other decision models examined.

If the network of interpersonal influences is not known, then the two studies diverge on which model is the second-best predictor of group outcomes. The first study utilized *ad hoc* groups (or "laboratory groups") to make simple decisions which were scarce in information, and in these situations, a central tendency model tended to be the best model. The precise model utilized (mean wins, NTCGP, or median wins) varied depending on the decision scale, but the ultimate outcome was that the groups gravitated towards the center of the constellation of initial opinions. The most likely explanation is that group members wished to avoid conflict and therefore chose a model which equally valued each group members' opinions.

Surprisingly, majority wins models did not perform well in the first study. This may be due to the small group size. In groups of three, calculating the center of the group opinions is relatively easy to do compared to larger groups, and therefore groups could easily calculate the answer which best incorporated the initial opinions of all group members. In larger groups, this may be more difficult to do, and a majority wins rule may become the group decision process because it saves cognitive effort. Or, conformity pressures may be stronger in larger groups, and so factions may be more prevalent and have more power in larger groups. Both explanations suggest that group size appears to be an important avenue in future research to examine the ubiquity of majority wins decision rules. Kameda et al. (2003) show in simulations that majority wins rules are more likely to occur in larger groups.

In the second study, groups of professional accountants who had experience with each other were making decisions in information-rich environments. Here, faction behaviors tended to dominate, and majority wins once again emerged as an effective predictive model. The reason for this may lie in social comparison theory: given the complex information environment but the lack of a demonstrably correct answer, subjects may have assessed their expertise by comparing their responses to the responses of the other group members. The lack of either group or sample polarization suggests that persuasive arguments did not dictate the decision process; if the pool of information possessed and valued by the group was the determining factor, then systematic shifts within groups and / or between groups should have been observed in the direction that was already favored by the group. However, the shifts occurred in the direction of the majority, regardless of where the majority's opinion was relative to the midpoint of the scale (which is reflective of initial tendency).

In summary, the surprising results found in this dissertation provide several insights into group decision-making, and they suggest that a different view of group decision-making may be warranted. First, a behavioral classification of group decision models such as the one presented in Table 7 is surprisingly successful at linking the various decision models together and identifying a consistent pattern in group outcomes. Second, the results of Study 1 appear to suggest that cognitive effort is not the primary driver of "laboratory" group decision outcomes, but instead groups first try to equally weight the opinions of each member. If the cognitive effort to equally weight the members is too great, then faction-based models may ensue. Third, the information-rich setting of the auditing tasks apparently, and paradoxically, caused group members to seek social influences as a means of judging the best answer to the audit questions. This result has practical implications to policy makers and audit firms because if a bias can be instilled in auditors before the group meetings begins (for example, if a conservative bias is induced by reminding auditors of the principle of professional skepticism), then it is more likely the bias will be reflected in an audit group's decision. Thus, the psychological notion of priming may have a strong effect on audit teams if members are primed before group discussions.

Given the success in succinctly describing group behaviors in Studies 1 and 2, the classification scheme presented in Table 7 of Study 1 appears to be a promising area for future research. Subsequent studies may begin refining the classification scheme and identifying which group behaviors are most likely to occur in which task settings. For example, the classification scheme currently does not account for intellective tasks for which a demonstrably correct answer exists, in which case the truth-wins model enjoys a great deal of success. Additional models may also be classified as they are identified (for example, a dictator-wins model may be included in the idiosyncratic influence classifications). This suggests a two-tier classification system: first, the group behavior is identified for a task (e.g., judgment, intellective), and second, the decision model which is most appropriate for the specific task or group characteristics (e.g., continuous or discrete scale; mixed or similar organizational structures) is identified.

In this study, the discrete scale response was a five-point scale using probability assessments between 10% and 90% (or, in the grading task, between 55% and 95%). Russell et al. (1991) demonstrated that using a five-point Likert scale with points of 10% through 90% forced subjects to adjust their responses so that significant information loss resulted in the subjects' assessments of a problem. Russell and Bobko (1992) show that subjects using that same five-point scale are not only forced into choosing a response which does not adequately represent their viewpoints, but that subject often employ a nonlinear response function in the five-point scale condition when compared to responses on a continuous (150-point) scale. That study was a between-subjects design in which subjects were essentially given five scenarios[11], and in one condition subjects were asked to provide a continuous response by marking their likelihood on a 150mm line (their response was measured as the number of millimeters from the endpoint of the line), and in the other condition, subjects were asked to provide a percentage response given choices of 10% through 90% in 20% increments (as is done in this study). When the five scenarios are placed on a graph along the X-axis, the responses in the continuous scale condition resembled an elongated 'S' curve. However, in the discrete scale condition, responses resembled a rotated 'U' curve. Therefore, those two studies demonstrated that a five-point scale does induce responses which are different than those which are observed using a much finer continuous scale.

While a five-point probability scale has been shown to induce different behaviors than a continuous scale, this particular scale may not necessarily induce purely "discrete" decision behaviors that are documented in the SDS stream of research. In many of those studies, the decision alternatives are mutually exclusive (such as guilty or not guilty; hiring Candidate A, B, or C; etc.) whereas in this study the decision alternatives are ordinal. While some SDS research uses ordinal scales (for example, Laughlin and Earley, 1982), the use of ordinal alternatives may create some differences in behaviors than may be observed if the decision alternatives were mutually exclusive.

Several limitations exist to these studies. In the first study, some of the analyses required a substantial paring of the sample size to isolate those conditions under which the hypotheses could be tested (for example, to conduct the analysis in Table 3, of the 66 total group observations in the continuous scale condition, 42 had to be dropped to isolate only those 24 groups which had two members with identical pre-discussion opinions). This loss of data may reduce the generalizability of some of the specific hypothesis tests. However, almost all groups in the continuous and discrete scale conditions are included in Table 7

[11] The actual design used by Russell and Bobko (1992) was a more complicated 5 × 5 design, but for simplicity of exposition, the scenario in which Violence = 5 is described because the results in that condition are easiest to explain in words. Their reported results are relatively robust across most of their conditions.

which summarizes the main findings of Study 1 that central tendency behaviors dominate the group decisions[12]. Therefore, while the loss of information may affect some specific tests, the primary analysis and findings of Study 1 (which is reported in Table 7) retains most of the information that was collected.

Limitations which apply to both studies include the use of a relatively small group (three members), and the use of only one type of task (judgment). Some of the limitations imposed by the small group size have been discussed in Study 1; namely, that conformity effects tend to be smaller in groups of three and therefore some of the behaviors which are observed in other group research utilizing larger groups may not be observed in this study. The smaller group size may also facilitate the computation of the average of the members' pre-discussion opinions which may also affect the group decisions. However, three-person groups were chosen for Study 1 because their behaviors may be compared to the three-person groups in Study 2 which more closely reflect actual audit teams found in practice (see the discussion section of Study 2 for limitations regarding the composition of those groups). One potential avenue for further research is to specifically manipulate group size and/or the potential decision outcomes (which may be easier or harder to calculate group means) to isolate whether cognitive ease or the relative lack of faction-based conformity pressures are driving the central tendency behaviors observed in Study 1.

Second, both studies may suffer from a potential lack of interest by the subjects because no incentives for effort or performance were offered. This is more likely a potential limitation of Study 1 rather than Study 2 because in Study 2, participants were volunteers who presumably had an interest in the study when they agreed to participate. In Study 1, the lack of performance- or effort-based incentives may have contributed to the central tendency behaviors as group members may have chosen the decision outcome that avoided conflict – namely, mean wins or another decision model which equally incorporated group members' opinions. However, much psychological research on group decision making does not offer incentives to their participants, and because of that, the results of Study 1 are comparable to that research.

[12] Groups with unanimous pre-discussion opinions are excluded from Table 7 because all decision models predict the same group outcome (that the group will decide on the unanimous initial opinion), and therefore those groups may not be classified by behavior.

Third, the use of student subjects in Study 1 limits both the generalizability of the results as well as the comparability of results between Study 1 and Study 2 which uses professional auditors. The costs and benefits of using student subjects have been well documented in both the psychological and accounting literatures. The use of professional auditors for Study 1 would be inappropriate because decision making theory is being examined independent of the task setting, and using professional subjects who are donating their time would be inefficient (Libby et al., 2002), and therefore, student subjects are a better pool of subjects. Further, using students in Study 1 offers the advantage of comparability to other psychological studies of group decision making which frequently employ student subjects. The use of student subjects is inappropriate for Study 2 because of the lack of professional knowledge and experience which is needed to replicate behaviors that may occur in practice, and therefore, auditors are a better pool of subjects for Study 2. Therefore, student subjects were chosen for Study 1 and professional auditors were chosen as subjects for Study 2.

SINT, the most successful decision model, has its own set of limitations. First, SINT is based on the actors' perceptions of influence of each other, and therefore it is subject to biases by the actors such as egocentric biases or stereotypes. Although these studies demonstrate that subjects are able to identify interpersonal influences rather well, in some situations the assessment of the network may be less accurate and therefore reduce the effectiveness of SINT. Second, SINT is based on data collected after the tasks have been completed, and therefore the predictive ability of SINT is limited to similar tasks being completed by the same group members. Future research may seek to establish *a priori* predictions for the network of influences. Third, the sustainability of the network of interpersonal influences is unknown. In other words, how long does the network hold? Many research opportunities exist regarding both the persistence and formation of the network (Do group members become fixated on the network, even when the relative expertise of group members shifts? Do task order effects change the formation of the network?).

The finding that SINT is the most successful model implies a consistent network of influences across tasks. This may in fact represent reality because one person may exhibit a strong leadership style which is carried forward through many tasks and therefore consistently influence the group decision, or the group may take on an identity which is to avoid conflict with group members which may lead to a network of approximately influences, and so on. It is entirely possible that the group establishes a norm for behaviors and decision making which is persistent. However, as discussed above, unless that norm is known or may be estimated *a priori*, the usefulness of SINT as a predictive model is limited.

Because the network of interpersonal influences is identified after the group discussion, SINT may appear to be a circular decision model. However, in both studies, several tasks were completed before the network was identified, and therefore the persistent model of influence was presumably obtained from the poker chip distribution. If the true network of interpersonal influences varied from task to task, then the network of influences that is derived from the poker chip distribution completed at the end of all of the tasks should only apply to one task if all subjects allocated poker chips based on the network of influences for the same task, or none of the tasks if each group member allocated poker chips based on the network of influences for different tasks. Most likely, if the network did truly vary, then subjects would allocate poker chips based on the last task completed because the salience of the network would be strongest. The results do not bear this out. For example, in Study 2, the accounts receivable task was the first task completed, followed by the budget allocation, followed by the going concern likelihood. Subjects allocated poker chips after the going concern task (the last task), but the network of influences derived from the poker chip distribution provided the best model fit for the accounts receivable task (the first task). Further, correlation tests (not reported) conducted on the data in Study 1 in which SINT data was collected at two points (after the grading tasks and after the CDQ tasks) show a high correlation in the number of poker chips that a subject received between the first poker chip allocation and the second poker chip allocation. In other words, if a person received a large number of poker chips after completing the grading tasks, that person likely received a large number of poker chips after the CDQ tasks. Therefore, it appears that the poker chip distribution does in fact describe a persistent network of influences.

The results of Study 2 indicate that information may not be shared optimally in group discussions, and especially unique information. At least two solutions have been identified in the group literature to overcome this phenomenon. First, a group agenda which facilitates the identification and evaluation of information (such as requiring decision makers to justify their initial opinions in writing) can overcome the group process loss (Nunamaker et al., 1991). Second, employing GDSS technology which supports the group agenda or provides tools which allow the organization of information may facilitate the sharing and evaluation of information (Fjermestad and Hiltz, 1998; Murthy and Kerr, 2004).

One important question that is raised by the results of this dissertation, and specifically Study 2, is how an information generation (i.e., brainstorming) session affects the subsequent evaluation of that information by the group. In this study, groups were not instructed to brainstorm ideas before making their decisions, so that question can not be answered by this study. However, these results suggest that the brainstorming session required by SAS 99 at the beginning of audits may not increase audit effectiveness. Even though more information is generated as a result of the brainstorming session (Carpenter, 2005), audit teams may focus on social influences rather than on the information that was generated in the brainstorming session. While Carpenter (2005) finds that groups which brainstormed had higher fraud risk assessments than the individual members of the groups before discussion (which she labels "nominal" groups; in other words, she finds polarization of fraud risk assessments in her groups), the experimental design of her study does not allow her to test whether the higher assessments were caused by the information that was generated, or if the assessments were caused by the differing levels of the organization status of the group members. In her study, the group members at the highest organizational level (in her case, managers) also had the highest fraud risk assessments, so the polarization that is reported in her study may be caused by either informational influences (the additional ideas from the brainstorming session) or by social influences (the differing organizational levels). Therefore, the effect of the brainstorming session on group decisions remains an open question; the results of this study suggest that the brainstorming session may not function as intended by FASB.

This study is a response to the many calls for more group research by accounting researchers, particularly given the passage of SAS 99. As mentioned in the introduction to Study 2, research on group decision making in the accounting, and especially the audit, literature is scarce. However, accounting researchers may have a difficult task in studying group decision processes for two reasons: first, the cost of group decision research imposes a practical constraint on its pervasiveness, and second, the psychological literature on group decision making is so complex that accounting researchers may have difficulty in identifying the relevant theories upon which to draw. Using groups of three, which is a common size of audit teams, this dissertation simultaneously examines group decision theories in an information-scarce and information-rich setting using students (as is typical in psychology research) and accounting professionals (as is preferred in accounting research) to highlight the differences in behaviors between the two settings and provide a baseline for future audit research. The primary difference between the two settings – that laboratory groups employ central tendency models and auditors employ faction-based models – was surprising, and exactly opposite prior expectations where laboratory groups were predicted to be more subject to factional influences and auditors would incorporate the opinions all of group members. However, given this starting point, audit research may now begin examining under which circumstances information will play a greater role and how the interactions of social and informational influences map into audit effectiveness.

REFERENCES

AICPA. 2002. *Consideration of Fraud in a Financial Statement Audit*, Statement on Auditing Standards No. 99. New York, NY: AICPA.

Arunachalam, V., & Dilla, W. 1992. Computer-mediated communication and structured interaction in transfer pricing negotiation. *Journal of Information Systems* 6 (2):149-170.

Arunachalam, V., & Dilla, W.N. 1995. Judgment accuracy and outcomes in negotiation: A causal modeling analysis of decision-aiding effects. *Organizational Behavior and Human Decision Processes* 61 (3):289-305.

Asare, S.K. 1992. The auditor's going-concern decision: interaction of task variables and the sequential processing of evidence. *Accounting Review* 67 (2):379-393.

Ashton, R.H., & Kennedy, J. 2002. Eliminating recency with self-review: the case of auditors' 'going concern' judgments. *Journal of Behavioral Decision Making* 15 (3):221-231.

Bamber, E.M., Hill, M.C., & Watson, R.T. 1998. Audit groups and group support systems: A framework and propositions for future research. *Journal of Information Systems* 12 (2):45-73.

Bamber, E.M., Watson, R.T., & Hill, M.C. 1996. The effects of group support system technology on audit group decision making. *Auditing: A Journal of Practice & Theory* 15 (1):122-134.

Bell, T.B., Bedard, J.C., Johnstone, K.M., & Smith, E.F. 2002. KRisk(SM): A computerized decision aid for client acceptance and continuance risk assessments. *Auditing: A Journal of Practice & Theory* 21 (2):97-113.

Bonner, B.L., Baumann, M.R., & Dalal, R.S. 2002. The effects of member expertise on group decision-making and performance. *Organizational Behavior & Human Decision Processes* 88 (2):719-736.

Brown, C.E., & Solomon, I. 1991. Configural information processing in auditing: The role of domain-specific knowledge. *Accounting Review* 66 (1):100-119.

Burnstein, E. 1982. Persuasion as argument processing. In *Contemporary Problems in Group Decision-Making*, edited by G. Stocher-Kreichgauer. New York, NY: Academic Press.

Butler, J.K., and M.D. Crino. 1992. Effects of initial tendency and real risk on choice shift. *Organizational Behavior and Human Decision Processes* 53 (October):14-34.

Carpenter, T.D. 2005. Partner influence, team brainstorming, and fraud risk assessment: Some implications of SAS No. 99. Working paper, University of Georgia.

Chalos, P. 1985. Financial distress: A comparative study of individual, model, and committee assessments. *Journal of Accounting Research* 23 (2):527-543.

Chernyshenko, O.S., Miner, A.G., Baumann, M.R., & Sniezek, J.A. 2003. The impact of information distribution, ownership, and discussion on group member judgment: The differential cue weighting model. *Organizational Behavior & Human Decision Processes* 91 (1):12-25.

Crott, H.W., Szilvas, K., and Zuber, J.A. 1991. Group Decision, Choice Shift, and Polarization in Consulting, Political, and Local Political Scenarios: An Experimental Investigation and Theoretical Analysis. *Organizational Behavior and Human Decision Processes* 49 (1):22.

Davis, J.H. 1973. Group decision and social interaction: A theory of social decision schemes. *Psychological Review* 80 (2):97-125.

Davis, J.H. 1996. Group decision making and quantitative judgments: A consensus model. In *Understanding Group Behavior*, edited by E. H. Witte and J. H. Davis. Mahwah, New Jersey: Lawrence Erlbaum.

Davis, J.H., Au, W.T., Hulbert, L., Chen, X.P., & Zarnoth, P. 1997. Effects of group size and procedural influence on consensual judgments of quantity: The example of damage awards and mock civil juries. *Journal of Personality and Social Psychology* 73 (4):703-718.

Dilla, W.N., & Stone, D.N. 1997a. Representations as decision aids: The asymmetric effects of words and numbers on auditors' inherent risk judgments. *Decision Sciences* 28 (3):709-743.

Dilla, W.N., & Stone, D.N. 1997b. Response scales in risk judgments: The effects of representation, fineness and user choice. *Journal of Information Systems* 11 (2):75-96.

Einhorn, H.J., Hogarth, R.M., & Klemper, E. 1977. Quality of group judgment. *Psychological Bulletin* 84 (1):158-172.

Festinger, L. 1954. A theory of social comparison processes. *Human Relations* 7:117-140.

Fjermestad, J., & Hiltz, S.R. 1998. An assessment of group support systems experimental research: Methodology and results. *Journal of Management Information Systems* 15 (3):7-149.

Friedkin, N.E. 1998. *A structural theory of social influence*. Cambridge, New York: Cambridge University Press.

Friedkin, N.E. 1999. Choice shift and group polarization. *American Sociological Review* 64 (6):856-875.

Friedkin, N.E., & Johnsen, E.C. 1999. Social influence networks and opinion change. *Advances in Group Processes* 16:1-29.

Gigone, D., & Hastie, R. 1993. The common knowledge effect: Information sharing and group judgment. *Journal of Personality and Social Psychology* 65 (5):959-974.

Gigone, D., & Hastie, R. 1997. The impact of information on small group choice. *Journal of Personality and Social Psychology* 72 (1):132-140.

Hastie, R., & Kameda, T. 2005. The robust beauty of majority rules in group decisions. *Psychological Review* 112 (2):494-508.

Hinsz, V.B. 1999. Group decision making with responses of a quantitative nature: The theory of social decision schemes for quantities. *Organizational Behavior and Human Decision Processes* 80 (1):28-49.

Isenberg, D.J. 1986. Group polarization: A critical review and meta-analysis. *Journal of Personality and Social Psychology* 50 (6):1141-1151.

Kahneman, D., Slovic, P., & Tversky, A. 1982. *Judgment under uncertainty: heuristics and biases*. Cambridge, New York: Cambridge University Press.

Kahneman, D., Knetsch, J.L., & Thaler, R.H. 1990. Experimental tests of the endowment effect and the Coase theorem. *The Journal of Political Economy* 98 (6):1325-1348.

Kameda, T., Tindale, R.S., & Davis, J.H. 2003. Cognitions, preferences, and social sharedness: Past, present, and future directions in group decision-making. In *Emerging perspectives on judgment and decision research*, edited by S. L. Schneider and J. Shanteau. Cambridge, New York: Cambridge University Press.

Kaplan, M.F., & Miller, C.E. 1987. Group decision making and normative versus informational influence: Effects of type of issue and assigned decision rule. *Journal of Personality and Social Psychology* 53 (2):306-313.

Karan, V., Kerr, D.S., Murthy, U.S., & Vinze, A.S. 1996. Information technology support for collaborative decision making in auditing: An experimental investigation. *Decision Support Systems* 16:181-194.

Kennedy, J., Kleinmuntz, D.N., & Peecher, M.E. 1997. Determinants of the justifiability of performance in ill-structured audit tasks. *Journal of Accounting Research* 35 (3):105-123.

Kerr, N.L., MacCoun, R.J., & Kramer, G.P. 1996a. Bias in judgment: Comparing individuals and groups. *Psychological Review* 103 (4):687-719.

Kerr, N.L., MacCoun, R.J., & Kramer, G.P. 1996b. "When are N heads better (or worse) than one?": Bias judgment in individuals versus groups. In *Understanding Group Behavior*, edited by E. H. Witte and J. H. Davis. Mahwah, New Jersey: Lawrence Erlbaum.

Kerr, N.L., & Tindale, R.S. 2004. Group performance and decision making. *Annual Review of Psychology* 55:623-655.

Kleinmuntz, D.N., & Schkade, D.A. 1993. Information displays and decision-processes. *Psychological Science* 4 (4):221-227.

Landis, M., Arunachalam, V., & Wheeler, P. 2006. Choice shifts and group polarization in dyads. Working paper, University of Missouri-Columbia.

Larson Jr., J.R., Foster-Fishman, P.G., & Keyes, C.B. 1994. Discussion of shared and unshared information in decision-making groups. *Journal of Personality & Social Psychology* 67 (3):446-461.

Larson Jr., J.R., Christensen, C., Abbott, A.S., & Franz, T.M. 1996. Diagnosing groups: Charting the flow of information in medical decision-making teams. *Journal of Personality & Social Psychology* 71 (2):315-330.

Laughlin, P.R., & Earley, P.C. 1982. Social combination models, persuasive arguments theory, social comparison theory, and choice shift. *Journal of Personality and Social Psychology* 42 (2):273-280.

Laughlin, P.R., & Ellis, A.L. 1986. Demonstrability and social combination processes on mathematical intellective tasks. *Journal of Experimental Social Psychology* 22 (3):177-189.

Laughlin, P.R., & Hollingshead, A.B. 1995. A theory of collective induction. *Organizational Behavior and Human Decision Processes* 61 (1):94-107.

Libby, R., & Luft, J. 1993. Determinants of judgment performance in accounting settings: Ability, knowledge, motivation, and environment. *Accounting, Organizations and Society* 18 (5):425-450.

Libby, R., & Trotman, K.T. 1993. The review process as a control for differential recall of evidence in auditor judgments. *Accounting, Organizations and Society* 18 (6):559-574.

Libby, R., Bloomfield, R., & Nelson, M.W. 2002. Experimental research in financial accounting. *Accounting, Organizations and Society* 27 (8):775-812.

Lorge, I., & Solomon, H. 1955. Two models of group behavior in the solution of Eureka-type problems. *Psychometrika* 29:139-148.

McGarty, C., Turner, J.C., Hogg, M.A., David, B., & Wetherell, M.S. 1992. Group polarization as conformity to the prototypical group member. *British Journal of Social Psychology* 31 (1):1-19.

McGrath, J. 1984. *Groups: Interaction and Performance*. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Murthy, U.S., & Kerr, D.S. 2004. Comparing audit team effectiveness via alternative modes of computer-mediated communication. *Auditing: A Journal of Practice & Theory* 23 (1):141-152.

Myers, D.G., and H. Lamm. 1976. The group polarization phenomenon. *Psychological Bulletin* 83 (4):602-627.

Nunamaker, J.F., Dennis, A.R., Valacich, J.S., Vogel, D.R., & George, J.F. (1991). Electronic meeting systems to support group work. *Communications of the ACM* 34 (7):40-61.

Peecher, M.E. 1996. The influence of auditors' justification processes on their decisions: A cognitive model and experimental evidence. *Journal of Accounting Research* 34 (1):125-140.

Piercey, M.D. 2006. The biases, strategic skewness, and justifiability of verbal (vs. numerical) audit risk assessments. Working paper, University of Illinois.

Postmes, T., Spears, R., & Cihangir, S. 2001. Quality of decision making and group norms. *Journal of Personality & Social Psychology* 80 (6):918-930.

Reckers, P.M.J., & Schultz, J.J. 1982. Individual versus group assisted audit evaluations. *Auditing: A Journal of Practice & Theory* 2 (1):64-74.

Reimers, J., Wheeler, S., & Dusenbury, R. 1993. The effect of response-mode on auditors' control risk assessments. *Auditing: A Journal of Practice & Theory* 12 (2):62-78.

Rich, J.S., Solomon, I., & Trotman, K.T. 1997. Multi-auditor judgment / decision making research: A decade later. *Journal of Accounting Literature* 16:86-126.

Rodrigo, M.F., and M. Ato. 2002. Testing the group polarization hypothesis by using logit models. *European Journal of Social Psychology* 32 (1):3-18.

Ross, M., & Sicoly, F. 1979. Egocentric biases in availability and attribution. *Journal of Personality and Social Psychology* 37 (3):322-336.

Russell, C. J., Pinto, J.K., & Bobko, P. 1991. Appropriate moderated regression and inappropriate research strategy: A demonstration of information loss due to scale coarseness. *Applied Psychological Measurement* 15:125-135.

Russell, C. J., & Bobko, P. 1992. Moderated regression analysis and Likert scales: Too coarse for comfort. *Journal of Applied Psychology* 77 (3):336-342.

Rutledge, R.W., & Harrell, A.M. 1994. The impact of responsibility and framing of budgetary information on group-shifts. *Behavioral Research in Accounting* 6:92-109

Schultz, J.J., Jr., & Reckers, P.M.J. 1981. The impact of group processing on selected audit disclosure decisions. *Journal of Accounting Research* 19 (2):482-501.

Sia, C.-L., Tan, B.C.Y., & Wei, K.-K. 2002. Group polarization and computer-mediated communication: Effects of communication cues, social presence, and anonymity. *Information Systems Research* 13 (1):70-90.

Solomon, I. 1987. Multi-auditor judgment / decision making research. *Journal of Accounting Literature* 6:1-25.

Stasser, G., & Titus, W. 1985. Pooling of unique information in group decision making: Biased information sampling during discussion. *Journal of Personality and Social Psychology* 48 (6):1467-1478.

Stasser, G., & Titus, W. 1987. Effects of information load and percentage of shared information on the dissemination of unshared information during group discussion. *Journal of Personality and Social Psychology* 53 (1):81-93.

Stasser, G., & Titus, W. 2003. Hidden profiles: A brief history. *Psychological Inquiry* 14 (3/4):304-313.

Stone, D.N., & Dilla, W.N. 1994. When numbers are better than words: The joint effects of response representation and experience on inherent risk judgments. *Auditing: A Journal of Practice & Theory* 13 (Supplement):1-19.

Stone, D.N., & Schkade, D.A. 1991. Numeric and linguistic information representation in multiattribute choice. *Organizational Behavior & Human Decision Processes* 49 (1):42-59.

Stone, D.N., & Schkade, D.A. 1994. Effects of attribute scales on process and performance in multiattribute choice. *Organizational Behavior & Human Decision Processes* 59 (2):261-287.

Stoner, J.A.F. 1961. A comparison of individual and group decisions involving risk, unpublished master's thesis, Massachusetts Institute of Technology.

Sunstein, C.R. 2002. The law of group polarization. *Journal of Political Philosophy* 10 (2):175.

Thaler, R.H. 1980. Toward a positive theory of consumer choice. *Journal of Economic Behavior and Organization* 1 (1):39-60.

Trotman, K.T. 1985. The review process and the accuracy of auditor judgments. *Journal of Accounting Research* 23 (2):740-752.

Tversky, A. 1972. Elimination by aspects: A theory of choice. *Psychological Review* 79 (3):281-299.

Vinokur, A., and E. Burnstein. 1978. Novel argumentation and attitude change: The case of polarization following group discussion. *European Journal of Social Psychology* 8:335-348.

Wallach, M., Kogan, N., & Bem, D. 1962. Group influence on individual risk taking. *Journal of Abnormal and Social Psychology* 65:75-86.

Wasserman, S., & Davis, J.H. 1991. A methodology for comparing predictions from many models to few data. *Quality and Quantity* 25 (2):189-209.

Winquist, J.R., & Larson, J.R. 1998. Information pooling: When it impacts group decision making. *Journal of Personality and Social Psychology* 74 (2):371-377.

Zuber, J.A., H.W. Crott, and J. Werner. 1992. Choice shift and group polarization: An analysis of the status of arguments and social decision schemes. *Journal of Personality and Social Psychology* 62 (1):50-61.

TABLE 1
Illustration of selected social decision scheme (SDS) models

Panel A: Four person group with two alternatives (yes or no): $r = 4$ and $n = 2$

	Equiprobability		Majority / Plurality		Proportionality		Truth-wins*	
(r_Y, r_N)	Yes	No	Yes	No	Yes	No	Yes	No
(4, 0)	1.00		1.00		1.00		1.00	
(3, 1)	0.50	0.50	1.00		0.75	0.25	1.00	
(2, 2)	0.50	0.50	0.50	0.50	0.50	0.50	1.00	
(1, 3)	0.50	0.50		1.00	0.25	0.75	1.00	
(0, 4)		1.00		1.00		1.00		1.00

* Assumes "Yes" is the demonstrably correct answer

Panel B: Four person group with three alternatives (A, B, or C): $r = 4$ and $n = 3$

	Majority / Plurality			Proportionality		
(r_A, r_B, r_C)	A	B	C	A	B	C
(4, 0, 0)	1.00			1.00		
(3, 1, 0)	1.00			0.75	0.25	
(3, 0, 1)	1.00			0.75		0.25
(2, 2, 0)	0.50	0.50		0.50	0.50	
(2, 0, 2)	0.50		0.50	0.50		0.50
(2, 1, 1)	1.00			0.50	0.25	0.25
(1, 3, 0)		1.00		0.25	0.75	
(1, 2, 1)		1.00		0.25	0.50	0.25
(1, 1, 2)			1.00	0.25	0.25	0.50
(1, 0, 3)			1.00	0.25		0.75
(0, 4, 0)		1.00			1.00	
(0, 3, 1)		1.00			0.75	0.25
(0, 2, 2)		0.50	0.50		0.50	0.50
(0, 1, 3)			1.00		0.25	0.75
(0, 0, 4)			1.00			1.00

The first column of each panel lists the possible constellations of members' initial opinions. In the other columns, the various alternatives are listed as well as possible decision schemes that might be employed by the group (e.g., equiprobability, majority, etc.). The probability that the group will decide on an alternative is shown in the table given the decision scheme and the constellation of initial group opinions. For example, as shown in the second row of Panel B, if the group employs a proportionality decision scheme and three people initially favor alternative A and one person favors alternative B, then the probability that the group will choose Alternative A (B) is 0.75 (0.25).

TABLE 2
Group decision models

Underlying Theory	Decision Rule	Individual cognitive effort*	Social effort*
Rules based on discrete decision alternatives			
SDS (Davis, 1973)	*Majority wins / Median wins* - If two or more members agree on an alternative, that alternative will be the final group decision. If each member has a unique pre-discussion preference, the group decision will be the median preferred alternative.	low	low
	Risk wins – The riskiest alternative favored by one of the group members will become the final group decision.	low	low
	Conservative wins – The most conservative alternative favored by one of the group members will become the final group decision.	low	low
	Polarization wins – If the group is predisposed towards a risky alternative for the task at hand (i.e., the average of the group members' pre-discussion preferences is risky), then the riskiest alternative favored by a group member will be chosen. If the group is predisposed towards a conservative alternative, then the most conservative alternative will be chosen.	low	low
Rules based on continuous decision alternatives (weighted average models)			
SDS-Q (Hinsz, 1999)	*Arithmetic mean* – The mean of the pre-discussion preferences will be the final group preference.	high	high
	Nearest to the center of gravity of proportionality – the alternative that lies closest to the arithmetic mean of the pre-discussion preferences will be the final group decision (this accounts for the fineness of the scale).	high	high
	Weighted majority – the majority opinion will carry a weight of 0.8 and the minority opinion will carry a weight of 0.2.	high	high
SJS (Davis, 1996)	*SJS* – the influence of an opinion will decay exponentially as it deviates from the mean of the opinions.	high	high
SINT (Friedkin, 1999)	*SINT* – the group decision will be a function of the weighted influences of each member on each other member.	high	high

* Discrete effort levels are based on the findings of Dan Stone and William Dilla (Dilla and Stone, 1997a, 1997b; Stone and Dilla, 1994) and Stone and Schkade (1994) while continuous effort levels are based on Hastie and Kameda's (2005) assessments.

TABLE 3
Decision model performance for groups with two members with identical pre-discussion preferences (risk-free tasks)

		Overall				Capitalism			Haiku		
Model	Classification	precision	*p* -value	rank	hitrate	precision	rank	hitrate	precision	rank	hitrate
Continuous setting											
SINT	continuous (weighted opin.)	1.56	--	1	38%	1.24	1	50%	1.87	1	25%
Mean wins	continuous (weighted opin.)	2.50	0.040	2	17%	2.14	3.5	17%	2.86	2	17%
SJS	continuous (weighted opin.)	2.51	0.036	3	17%	2.14	3.5	17%	2.88	3	17%
Weighted majority	continuous (weighted opin.)	2.58	0.028	4	8%	2.22	5	0%	2.95	4	17%
NTCGP	continuous (weighted opin.)	2.63	0.027	5	17%	2.25	6	17%	3.00	5	17%
Majority wins	discrete	2.92	0.050	6	50%	2.42	7.5	50%	3.42	6	50%
Polarization wins	discrete	3.25	0.011	7.5	42%	2.42	7.5	50%	4.08	7	33%
High wins	discrete	3.25	0.034	7.5	50%	2.00	2	50%	4.50	8	50%
Low wins	discrete	6.92	0.001	9	4%	5.67	9	8%	8.17	9	0%
Number of groups			24				12			12	
Discrete setting											
SINT	continuous (weighted opin.)	2.15	--	1	38%	1.09	1	43%	3.20	1	33%
NTCGP	continuous (weighted opin.)	3.10	0.130	2	69%	2.38	3	76%	3.81	2	62%
Majority wins	discrete	3.57	0.011	3	64%	2.38	3	76%	4.76	6	52%
Weighted majority	continuous (weighted opin.)	4.05	0.001	4	0%	3.43	5	0%	4.67	5	0%
SJS	continuous (weighted opin.)	4.35	0.001	5	0%	4.09	6	0%	4.63	4	0%
Mean wins	continuous (weighted opin.)	4.37	0.001	6	0%	4.13	7	0%	4.60	3	0%
Low wins	discrete	5.00	0.001	7	52%	4.29	8	57%	5.71	7	48%
Polarization wins	discrete	5.24	0.001	8	52%	2.38	3	76%	8.10	9	29%
High wins	discrete	6.43	0.001	9	38%	5.71	9	43%	7.14	8	33%
Number of groups			42				21			21	

Precision is the absolute value of the difference between the model's prediction and the actual group decision.
p -value is the significance level of the t-test between the precision of the most accurate model and the given model (one-tailed)
Rank is the model's rank, relative to the other models, in terms of precision
Hitrate is the percentage of predictions that are within +/- 0.5 scale units of the actual group decision

TABLE 4

Decision model performance for groups which are asked to provide both a continuous and discrete response (groups have two members with identical pre-discussion responses)

		Overall			**Capitalism**		**Haiku**	
Model	Classification	precision	*p* -value	rank	precision	rank	precision	rank
Continuous responses								
Weighted majority	continuous (weighted opin.)	1.41	--	1	1.26	2	1.68	4
Majority wins	discrete	1.65	0.179	2	1.05	1	2.75	6
SJS	continuous (weighted opin.)	1.74	0.048	3	1.86	3	1.52	1
NTCGP	continuous (weighted opin.)	1.82	0.062	4	1.95	5	1.58	3
Mean wins	continuous (weighted opin.)	1.84	0.043	5	2.02	6	1.53	2
SINT	continuous (weighted opin.)	2.05	0.021	6	1.95	4	2.24	5
Polarization wins	discrete	3.59	0.035	7	3.68	7	3.42	7
Low wins	discrete	3.71	0.028	8	3.77	8	3.58	8
High wins	discrete	4.11	0.001	9	4.05	9	4.25	9
Number of groups			34			22		12
Discrete responses								
NTCGP	continuous (weighted opin.)	2.58	--	1	2.35	1.5	2.86	1
SINT	continuous (weighted opin.)	3.49	0.085	2	3.24	3	3.79	2
Majority wins	discrete	3.55	0.080	3	2.35	1.5	5.00	6
Weighted majority	continuous (weighted opin.)	4.06	0.001	4	3.53	4.5	4.71	5
SJS	continuous (weighted opin.)	4.40	0.001	5	4.22	7	4.62	4
Mean wins	continuous (weighted opin.)	4.41	0.001	6	4.31	8	4.52	3
Polarization wins	discrete	5.81	0.001	7	4.12	6	7.86	8
High wins	discrete	6.13	0.001	8	3.53	4.5	9.29	9
Low wins	discrete	7.42	0.001	9	8.24	9	6.43	7
Number of groups			62			34		28

Precision is the absolute value of the difference between the model's prediction and the actual group decision.
p -value is the significance level of the t-test between the precision of the most accurate model and the given model (one-tailed)
Rank is the model's rank, relative to the other models, in terms of precision
Hitrate is the percentage of predictions that are within +/- 0.5 scale units of the actual group decision

TABLE 5
Decision model performance for risky tasks

	Model	Model type	precision	*p* -value	rank	hitrate
Panel A: Continuous, two members equal						
	SINT	continuous (weighted opin.)	3.74	--	1	19%
	SJS	continuous (weighted opin.)	4.12	0.245	2	6%
	Mean wins / NTCGP	continuous (weighted opin.)	4.13	0.216	3	13%
	Weighted majority	continuous (weighted opin.)	4.47	0.160	4	3%
	Majority wins	discrete	6.64	0.003	5	31%
	Polarization wins	discrete	8.30	0.001	6	25%
	High wins	discrete	10.05	0.001	7	19%
	Low wins	discrete	11.17	0.001	8	16%
	Number of groups			32		
Panel B: Continuous, all members unique						
	SINT	continuous (weighted opin.)	5.43	--	1	13%
	Mean wins / NTCGP	continuous (weighted opin.)	5.94	0.163	2	19%
	SJS	continuous (weighted opin.)	6.02	0.113	3	14%
	Weighted majority	continuous (weighted opin.)	7.15	0.006	4	6%
	Median wins	discrete	7.59	0.002	5	28%
	Polarization wins	discrete	17.13	0.001	6	5%
	High wins	discrete	20.26	0.001	7	3%
	Low wins	discrete	20.45	0.001	8	2%
	Number of groups			109		
Panel C: Discrete, two members equal						
	NTCGP	continuous (weighted opin.)	5.37	--	1	74%
	Majority wins	discrete	7.56	0.030	2	65%
	SINT	continuous (weighted opin.)	8.05	0.012	3	24%
	Weighted majority	continuous (weighted opin.)	8.39	0.001	4	0%
	Mean wins	continuous (weighted opin.)	8.94	0.001	5	2%
	SJS	continuous (weighted opin.)	9.06	0.001	6	0%
	Polarization wins	discrete	11.22	0.001	7	52%
	High wins	discrete	12.20	0.001	8	48%
	Low wins	discrete	15.85	0.001	9	30%
	Number of groups			82		
Panel D: Discrete, all members unique						
	Median wins	discrete	4.65	--	1	77%
	NTCGP	continuous (weighted opin.)	5.12	0.330	2	74%
	SJS	continuous (weighted opin.)	5.48	0.114	3	47%
	Mean wins	continuous (weighted opin.)	5.74	0.121	4	47%
	Weighted majority	continuous (weighted opin.)	6.70	0.001	5	23%
	SINT	continuous (weighted opin.)	8.42	0.011	6	9%
	Polarization wins	discrete	20.00	0.001	7	7%
	High wins	discrete	26.05	0.001	8	2%
	Low wins	discrete	27.44	0.001	9	5%
	Number of groups			43		

TABLE 6
Group polarization and sample polarization

Panel A: All groups whose members' initial opinions were not unanimous

	N (groups)	Individuals Median	Individuals Mean	Groups Mean	Sample polarization	*p*-value		Group polarization	*p*-value	
Risk-free tasks (grading)										
Continuous: two members equal	24	90.00	87.50	88.92	1.42	0.037	**	1.56	0.021	**
Continuous: all members unique	38	86.67	86.61	87.14	0.54	0.626		0.70	0.528	
Total continuous	62	87.17	86.95	87.83	0.88	0.223		1.03	0.152	
Discrete: two members equal	42	88.33	86.98	86.42	-0.56	0.443		-0.56	0.443	
Discrete: all members unique	14	80.00	79.76	79.29	-0.48	0.699		-0.95	0.435	
Total continuous	56	88.33	85.18	84.64	-0.54	0.385		-0.66	0.287	
Risky tasks (CDQ)										
Continuous: two members equal	32	53.33	55.77	55.70	-0.07	0.944		-0.02	0.987	
Continuous: all members unique	109	58.33	58.00	57.54	-0.46	0.586		-0.04	0.964	
Total continuous	141	58.33	57.49	57.12	-0.37	0.588		-0.03	0.962	
Discrete: two members equal	82	60.00	59.11	60.24	1.14	0.325		-0.33	0.779	
Discrete: all members unique	43	56.67	56.82	56.98	0.16	0.907		-0.16	0.907	
Total continuous	125	56.67	58.32	59.12	0.80	0.364		-0.27	0.763	

Sample polarization is the average difference between the mean pre-discussion opinions of the individuals of a particular group, and the actual group opinion. Group polarization is calculated by obtaining the choice shift (the absolute value of the difference between pre-discussion opinions within a group and the group's actual decision) and changing the sign of the choice shift depending on the direction of the group decision relative to the pre-discussion average. If the group's opinion was farther away from the midpoint (or median in the risk-free tasks) of the scale than the average pre-discussion opinions, then polarization is equal to choice shift. Otherwise, if the group shifted towards the midpoint (median), the polarization is equal to the choice shift times negative one.

TABLE 6 (continued)
Group polarization and sample polarization

Panel B: All groups with two members with equal opinions

		Individuals		Groups				
	N (groups)	Median	Mean	Mean	Sample polarization	*p* -value	Group polarization	*p* -value
Risk-free tasks (grading)								
Majority wins								
Continuous	12	92.08	89.52	92.08	2.56	0.007 ***	1.72	0.099 *
Discrete	27	86.48	87.10	86.48	-0.62	0.394	1.60	0.021 **
Majority does not win								
Continuous	12	84.58	85.47	85.75	0.28	0.774	1.61	0.073 *
Discrete	15	87.00	86.78	86.33	-0.44	0.784	-4.00	0.004 ***
Risky tasks (CDQ)								
Majority wins								
Continuous	10	64.00	63.30	64.00	0.70	0.661	0.70	0.661
Discrete	53	60.57	58.68	60.57	1.89	0.075 *	2.39	0.023 **
Majority does not win								
Continuous	29	58.99	59.89	59.66	-0.23	0.932	-3.91	0.137
Discrete	22	53.86	52.35	51.93	-0.42	0.733	0.61	0.615

Sample polarization is the average difference between the mean pre-discussion opinions of the individuals of a particular group, and the actual group opinion. Group polarization is calculated by obtaining the choice shift (the absolute value of the difference between pre-discussion opinions within a group and the group's actual decision) and changing the sign of the choice shift depending on the direction of the group decision relative to the pre-discussion average. If the group's opinion was farther away from the midpoint (or median in the risk-free tasks) of the scale than the average pre-discussion opinions, then polarization is equal to choice shift. Otherwise, if the group shifted towards the midpoint (median), the polarization is equal to the choice shift times negative one.

TABLE 7
A behavioral classification of group decision model performance

Continuous scale

	# of groups	Best model			2nd best model			3rd best model		
		Model	Precision	Hitrate	Model	Precision	Hitrate	Model	Precision	Hitrate
Grading tasks (risk-free)										
Two members are equal	24	SINT *influence*	1.56	38%	Mean wins *central tendency*	2.50	17%	SJS *faction*	2.51	17%
All members are unique	38	SINT *influence*	3.57	14%	Mean wins *central tendency*	4.09	21%	SJS *faction*	4.12	18%
CDQ tasks										
Two members are equal	32	SINT *influence*	3.74	19%	SJS *faction*	4.12	6%	Mean wins *central tendency*	4.13	13%
All members are unique	109	SINT *influence*	5.43	13%	Mean wins *central tendency*	5.94	19%	SJS *faction*	6.02	14%

Discrete scale

	# of groups	Best model			2nd best model			3rd best model		
		Model	Precision	Hitrate	Model	Precision	Hitrate	Model	Precision	Hitrate
Grading tasks (risk-free)										
Two members are equal	42	SINT *influence*	2.15	38%	NTCGP *central tendency*	3.10	69%	Maj wins *faction*	3.57	64%
All members are unique	14	Median *central tendency*	1.43	86%	NTCGP *central tendency*	2.14	79%	SJS *faction*	2.27	64%
CDQ tasks										
Two members are equal	82	NTCGP *central tendency*	5.37	74%	Maj wins *faction*	7.56	65%	SINT *influence*	8.05	24%
All members are unique	43	Median *central tendency*	4.65	77%	NTCGP *central tendency*	5.12	74%	SJS *faction*	5.48	47%

TABLE 8
Descriptive statistics for Study 2 participants

Panel A: Firms and organizational levels of study participants

Firm type	# of participants	% of sample	Organizational level: Partners / Sr. Managers	Managers	Seniors	Staff	# of equal groups	# of unequal groups
Big 4	51	54.8%	9	5	9	16	11	2
Regional	27	29.0%	1	1	14	11	6	3
Local	15	16.1%	6	2	4	3	1	4
Total	93	100.0%	16	8	27	30	18	9

Panel B: Experience of study participants

Organizational level	Years experience	Approx. # of audits	% of audit reports that were: Standard reports	Going concern qualifications	Disclaimers of opinion	Other
Partner / Sr. Manager	18.4	262.7	89.8%	1.4%	0.9%	7.8%
Manager	16.6	134.0	86.0%	1.0%	0.2%	12.8%
Senior	4.2	46.1	89.9%	5.3%	2.5%	0.4%
Staff	1.4	18.0	96.0%	3.2%	0.1%	0.7%

TABLE 9
Decision model performance for Accounts Receivable task

Model	Classification	hitrate	precision	Precision *p*-value*	Precision rank	bias	Bias *p*-value†	Bias rank
Continuous setting								
SINT	continuous (weighted avg.)	27%	1.75	--	1	0.67	0.290	1
Majority wins / median wins	discrete	40%	4.27	0.026	2	2.27	0.206	2
Weighted majority	continuous (weighted avg.)	13%	4.40	0.005	3	2.44	0.119	3
SJS	continuous (weighted avg.)	13%	6.29	0.001	4	2.76	0.166	4
Mean wins	continuous (weighted avg.)	20%	6.71	0.001	5	2.93	0.171	6
NTCGP	continuous (weighted avg.)	20%	6.73	0.001	6	2.87	0.180	5
Polarization wins	discrete	27%	9.20	0.006	7	3.60	0.307	7
Low wins	discrete	20%	12.07	0.002	8	-12.07	0.001	8
High wins	discrete	20%	18.60	0.001	9	18.60	0.001	9
Number of groups		15						
Discrete setting								
SINT	continuous (weighted avg.)	44%	2.31	--	1	-0.96	0.365	1
Majority wins / median wins	discrete	75%	6.25	0.067	2	-6.25	0.056	7
Mean wins	continuous (weighted avg.)	31%	7.50	0.002	3	-3.33	0.251	4
SJS	continuous (weighted avg.)	31%	7.53	0.003	4	-3.65	0.217	5
NTCGP	continuous (weighted avg.)	31%	7.63	0.002	5	-3.25	0.264	3
Weighted majority	continuous (weighted avg.)	13%	7.75	0.005	6	-3.75	0.234	6
High wins	discrete	44%	11.25	0.004	7	11.25	0.001	8
Low wins	discrete	44%	15.00	0.002	8	-15.00	0.003	9
Polarization wins	discrete	38%	16.25	0.001	9	-1.25	0.835	2
Number of groups		16						

Precision is the absolute value of the difference between the model's prediction and the actual group decision.
* *p*-value is the significance level of the t-test between the precision of the most accurate model and the given model (one-tailed)
† *p*-value is for a significant difference from zero (two-tailed)
Rank is the model's rank, relative to the other models, in terms of precision
Hitrate is the percentage of predictions that are within +/- 0.5 percentage points of the actual group decision

TABLE 10
Discussion of cues

Panel A: Average group ratings of cue discussion

Cue		Continuous	Discrete	Difference	*t* -stat	*p* -value
1	Willingness of shareholders to invest additional capital	2.78	1.90	0.88	1.74	0.09 *
2	Decreased overhead costs	0.43	0.46	-0.03	-0.09	0.93
3	Collection of overdue accounts receivable	2.23	2.08	0.15	0.42	0.68
4*	Loss of major customer	0.70	1.33	-0.63	-1.80	0.08 *
5	Deferment of payment of accounts payable	1.31	1.47	-0.16	-0.42	0.68
6	Operating losses	1.71	2.21	-0.50	-1.27	0.21
7	Willingness of shareholders to forgo dividends	1.48	0.92	0.56	1.25	0.22
8	Forecast of increasing profits	1.89	1.63	0.26	0.60	0.55
9*	Refusal of bank to renew line of credit	1.19	1.25	-0.06	-0.12	0.90
10*	Competitor's ERP implementation	0.61	0.67	-0.06	-0.17	0.86
11*	Engineer's resignation	0.39	1.00	-0.61	-1.77	0.09 *
12*	Expiration of patent	0.88	1.03	-0.15	-0.41	0.68
13*	Slow inventory turnover	1.00	0.85	0.15	0.40	0.69
14	Increasing customer base (number of users is doubling)	1.06	0.67	0.39	1.58	0.12

* indicates uniquely-held cue

Panel B: Average group ratings of cue discussion by common and unique cues

Cue type	Continuous	Discrete	Difference	*t* -stat	*p* -value
Common cues	1.70	1.46	0.23	1.19	0.42
Unique cues	0.80	1.03	-0.23	-0.82	0.24
All cues	1.38	1.34	0.04	0.23	0.82

TABLE 11
Importance of cues

Panel A: Average group ratings of cue importance

Cue		Continuous	Discrete	Difference	*t* -stat	*p* -value
1	Willingness of shareholders to invest additional capital	2.97	1.94	1.03	2.02	0.05 *
2	Decreased overhead costs	0.78	0.42	0.36	1.33	0.19
3	Collection of overdue accounts receivable	2.16	1.96	0.20	0.80	0.43
4*	Loss of major customer	1.09	1.54	-0.45	-1.17	0.25
5	Deferment of payment of accounts payable	1.51	1.45	0.06	0.20	0.84
6	Operating losses	1.87	2.32	-0.46	-1.19	0.25
7	Willingness of shareholders to forgo dividends	1.70	0.96	0.74	1.65	0.11
8	Forecast of increasing profits	1.88	1.65	0.23	0.47	0.64
9*	Refusal of bank to renew line of credit	1.70	1.48	0.22	0.43	0.67
10*	Competitor's ERP implementation	0.82	0.71	0.11	0.32	0.75
11*	Engineer's resignation	0.59	0.98	-0.39	-1.09	0.29
12*	Expiration of patent	1.20	1.05	0.15	0.38	0.71
13*	Slow inventory turnover	1.27	0.90	0.37	1.06	0.30
14	Increasing customer base (number of users is doubling)	1.80	1.17	0.63	1.39	0.18

* indicates uniquely-held cue

Panel B: Average group ratings of cue discussion by common and unique cues

Cue type	Continuous	Discrete	Difference	*t* -stat	*p* -value
Common cues	1.83	1.48	0.35	1.82	0.08 *
Unique cues	1.11	1.11	0.00	0.01	0.99

TABLE 12
Validation of self-reports regarding information cues

Panel A: Self-contribution bias

Cue type	n	Difference in perceived discussion	*t*-stat	*p*-value
Continuous	45	0.333	2.44	0.019 **
Discrete	48	0.443	3.24	0.002 ***
Overall	93	0.390	4.05	0.001 ***

Panel B: Perceived importance bias

Cue type	n	Difference in perceived importance	*t*-stat	*p*-value
Continuous	45	-0.031	-0.33	0.745
Discrete	48	0.185	2.44	0.018 **
Overall	93	0.081	1.34	0.185