**Dissertations**                                              Theses and Dissertations

2016

# Performance of Individuals and Teams on Cryptographic Tasks: Factors That Affect the Consideration of Alternative Strategies

Joseph Bihary
*Loyola University Chicago*

Follow this and additional works at: https://ecommons.luc.edu/luc_diss

Part of the Experimental Analysis of Behavior Commons

LOYOLA UNIVERSITY CHICAGO

PERFORMANCE OF INDIVIDUALS AND TEAMS ON CRYPTOGRAPHIC TASKS:

FACTORS THAT AFFECT THE CONSIDERATION OF ALTERNATIVE

STRATEGIES

A DISSERTATION SUBMITTED TO

THE FACULTY OF THE GRADUATE SCHOOL

IN CANDIDACY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

PROGRAM IN APPLIED SOCIAL PSYCHOLOGY

BY

JOSEPH G. BIHARY

CHICAGO, ILLINOIS

AUGUST 2016

TABLE OF CONTENTS

iv

# LIST OF TABLES

# LIST OF FIGURES

ABSTRACT

The two studies presented here examined factors that might affect teams' and individuals' tendency to follow outside advice when attempting to solve a complex problem known as letters-to-numbers. Past research on group dynamics suggests that a lack of group consensus or homogeneity reduces group members' confidence in their group's abilities, and may lead members both to seek and accept advice from outside the group. Study 1 experimentally manipulated group diversity in task performance strategies in order to investigate whether dyads whose members have divergent perspectives are more likely than homogeneous dyads to consider and use a problem-solving strategy presented from a source outside the group when trying to solve a letters-to-numbers problem. Experimental sessions were videotaped to allow for observational analysis. Mixed results suggested that diverse-strategy dyads may have been better at processing task-relevant information and had more productive discussions than same-strategy dyads. Study 2 sought to examine the role of confidence in taking outside advice at the individual level. Individuals attempted to solve two letters-to-numbers problems. On the first, the problem was made easier or more difficult in order to experimentally manipulate the participants' confidence in the strategy that they used when trying to solve it. Participants were given an alternative strategy to consider using on the second problem. Mixed results suggested that individuals with low strategy confidence were more likely to consider the alternative strategy and may have performed better on the second problem

than high strategy confidence individuals. Limitations of the studies and

recommendations for future research are discussed.

CHAPTER ONE

INTRODUCTION AND LITERATURE REVIEW

When facing an important problem to be solved or decision to be made, people often prefer to seek the help or advice of others before taking action. The traditions of "strength in numbers" and "two heads are better than one" persist as ingrained tendencies of human behavior and were evident very early in our evolutionary history (Caporael, 1997; Kenrick, Maner, & Li, 2005; Sedikides & Skowronski, 1997). The tendency to affiliate in groups has also been observed in non-human primates (Packer, 1977), suggesting a strong evolutionary drive to affiliate with others. While our more primitive ancestors may have banded together in groups in response to external threats (Baer & McEachron, 1982) or for other simple reasons of mere survival, modern humans continue to believe that cooperating with others, or at least seeking their help or advice, allows us to solve problems more effectively or more efficiently than what we could accomplish alone. Despite this seemingly obvious motivation of human nature, the scientific research community has questioned whether benefits of collaboration with others are absolute. Does relying on others for assistance grant us advantages over working alone in all situations, under all conditions? Are there disadvantages to cooperating with others, and do these outweigh the benefits? These simple questions do not have simple answers, and continue to inspire research scientists from a variety of disciplines that study group behavior. The field of social psychology and its related disciplines (such as

organizational behavior) offer some insights to aid in approaching this problem.

Studying group behavior requires understanding not only the group itself, but the nature of the task or problem the group faces. Groups—both human and otherwise—engage in a large and potentially indefinite variety of activities. Even focusing exclusively on modern humans doesn't do much to narrow this scope. Tasks can be physical or mental, easy or difficult, simple or complex, requiring cooperative group action or performed by a single individual within the group. Within psychology, one of the best recognized systems for organizing group tasks is McGrath's (1984) "task circumplex," which consists of both a two-dimensional structure and eight-category method for classifying tasks performed by groups. The two dimensions account for (1) the tasks involving mental/conceptual vs. physical/behavioral activity and (2) conflict vs. cooperation of group members while performing them. The eight task categories arrayed within these dimensions are intended to be inclusive of all group activities while also keeping the activities mutually exclusive of each other. The labels for these categories are as follows: planning, creativity, intellective, decision-making, cognitive conflict, mixed-motive, contests/battles/competitive, and performance/psycho-motor.

Although the model is admirable for its comprehensiveness, some theorists highlight critical weaknesses of the underlying concepts. Larson (2010) argues that the model still leaves considerable ambiguity in task classification unresolved, as well as clustering tasks together that other researchers have demonstrated as having important differences in their structures, casting doubt on how similar some tasks in the same category actually are. He argues that the simplicity of McGrath's categorization scheme

distorts more fine-grained distinctions between group tasks, and thus may be of limited

utility for understanding group behavior in the context of different tasks.

Other task organization systems focus on a single, specific characteristic to

categorize tasks. One such method is the intellective—judgmental continuum (Laughlin,

1980). Anchored at one end of this continuum are tasks that have objective, demonstrably

correct answers or solutions (intellective tasks), while tasks with subjective answers that

cannot be demonstrated as objectively correct (judgmental tasks) are anchored at the

other end. One example of a task at or very near the intellective end is a mathematical

problem (e.g., $2x + 3 = 7$, $x = ?$). Assuming that one understands the basic rules of

arithmetic and algebraic functions, it can be objectively demonstrated that $x = 2$ in that

equation. In contrast, an example of a judgmental task is deciding the "better" of two

pieces of artwork. While one might have a strong opinion that one is prettier or more

interesting, there is no objective method of demonstrating that this is objectively true, or

that either is "better" by these or any other means of comparison. In defining the nature

of this demonstrability concept in the context of group behavior, Laughlin & Ellis (1986)

list four criteria that must be met to consider tasks as demonstrable among group

members: (1) the group agrees on the underlying conceptual system applied to the task;

(2) the group possesses sufficient information to be able to complete the task; (3) group

members who are unable to solve the problem themselves can recognize a correct

solution if it is proposed to them; (4) group members who are able to solve the problem

themselves have sufficient ability, motivation, and time to demonstrate the correct

solution to other group members.

An understanding of task structures is incredibly important for informing

methodological approaches to study groups (Larson, 2010). Group behavior cannot be understood without a keen understanding of what behavior the group is engaging in, or put another way, the task it performs. It is for this reason that I next turn to a thorough review of the task being employed in the present study.

### Letters-to-Numbers

The current study utilized a task known as letters-to-numbers, a highly intellective (demonstrable) task notable for its featured role in group problem-solving literature (first published by Laughlin, Bonner, & Miner, 2002). Letters-to-numbers is a cryptographic task that involves solving a code of the numbers 0-9 randomly assigned to the letters A-J, with no repetitions (each letter A-J represents one number 0-9). The goal is to solve the full mapping of letters to numbers (determining which number each letter represents). In performing the task, participants propose algebraic equations in letters and are given the answers to those equations, also in letters. An example of an equation is "A+B-H = ?". The answer provided for this equation might be "A+B-H = DF" (the numerical value of "A" plus "B" minus "H" is equal to "DF," a 2-digit number). The task only involves the mathematical operations of addition and subtraction; multiplication and division are not permitted in proposed equations. Thus, multi-digit expressions such as "DF" do not imply "D multiplied by F," as would normally be the case in algebraic notation. "DF" is simply representing a 2-digit number.

The task is organized as a sequence of "trials," the nature of which has varied throughout the short history of the task being used in experimental research studies. However, all variations consist of the following components, in this order: participants (1) propose an equation in letters, (2) receive an answer to that equation (provided by the

experimenter) also in letters, (3) make guesses about the value of the coded letters (e.g., A=5, B=8, etc.), and (4) receive feedback about whether or not their guesses are correct. The task proceeds until participants either solve the full code or reach a predetermined stopping or failure state (e.g., a limited number of trials to use to try to solve it). Performance on the task is generally measured by the number of trials participants use to solve for the numerical value of all letters in the code, with the goal of solving the full code using as few trials as possible (fewer trials indicating better performance).

The letters-to-numbers task is considered to be highly intellective (demonstrable) due to its basis in mathematical logic (Laughlin, 1980; Laughlin & Ellis, 1986) and multi-step structure. Each step (trial) offers intermediate progress towards solving the full letters-to-numbers code, and information about the code can be inferred from the equations proposed, answers received, and guesses made. In the equation example above ("A+B-H=DF"), a logical assumption is that D=1. Recall that the letters each represent one number, 0-9. The largest possible value for the equation "A+B-H" would be 17 (if A or B was equal to 9, the other letter was equal to 8, and H was equal to 0), so D cannot possibly be greater than 1. D cannot be zero either, since it would be omitted from the answer (e.g., if F was equal to 4, the answer would be displayed as "F" (4), not "DF" (04)—zero is never included in the answer if it is the first digit).

Additionally, the highly logical structure of the letters-to-numbers problem lends itself to understanding progress through the task in two cyclical stages (Larson, 2010). The first stage, referred to here as "strategy development" for simplicity, involves developing a strategy (a method of developing equations that can be used to obtain information about the code). This is followed by the second stage, referred to here as

"strategy implementation," that involves actually using that strategy (submitting the equation to obtain its answer). These two stages are often cyclical and can be implemented numerous times as participants attempt to solve the code. Larson (2010) argues that groups of participants can perform both stages simultaneously (and even explore multiple options at each stage), with different members focusing on different stages. Thus, groups should be more efficient than individuals performing the task, since individuals are less adept at performing both stages simultaneously. The demonstrable nature of the task also allows group members to demonstrate to one another the relative value of different strategies during the strategy development stage. This is consistent with Laughlin & Ellis' (1986) criteria for demonstrability.

Group members being able to demonstrate their strategies to their fellow members is critical for the group to solve the code efficiently (or at all). A variety of strategies can be used to solve the code, but some are more efficient than others, allowing for the code to be solved using fewer trials. Larson (2010) discusses two aspects of letters-to-numbers strategy. One aspect of strategy is the types of equations used to obtain information about the code. Participants can propose relatively simple equations (such as A+B) or highly complex equations with many letters and mathematical operations. Laughlin (2011) discusses how more complex equations—defined as using more letters in the equations— are generally better to use to obtain information about the code, as they tend to help identify more letters per trial used.

An example of a particularly effective strategy to use is to add all unknown letters together. Initially this would involve adding all ten letters together on the first trial, since none of the letters are known (A+B+C+D+E+F+G+H+I+J). Given that the ten letters

always represent the same ten numbers (0-9) albeit in an unknown order, the answer to this equation will always be 45 (0+1+2+3+4+5+6+7+8+9), immediately providing the answer to two letters (the letters representing 4 and 5) without needing to know the value of any other letter. A participant could proceed with this strategy by removing the letters representing 4 and 5 from the equation, and adding the remaining eight unknown letters. The answer to this equation, regardless what the letters are, will always be 36 (0+1+2+3+6+7+8+9), providing answers for another two letters (3 and 6). One could continue with this method, submitting subsequent equations that are guaranteed to produce 27 and 18. Such a method is objectively more efficient than simpler strategies, such as those adding two letters. Although simpler equations can be informative in certain circumstances (e.g., if A+B equals a two-digit answer, the first digit of the answer must be 1, since adding two numbers cannot be greater than 17), they are not nearly as effective as adding all of the unknown letters together, and may not be informative at all. For example, if A+B equals a single-digit answer, this is relatively unhelpful early on in the problem with few or no letters known, due to the number of possible combinations of numbers that could produce this result.

A second aspect of strategy that is demonstrable is the method of making guesses about the code, described as "hypothesis strategies" by Larson (2010). In previous studies, one step in attempting to solve the code is proposing the value of one letter (e.g., A=1) and receiving feedback as to whether or not that guess is correct before guessing the full code (all letters and all numbers). Consider the example above where an equation "A+B" produces a 2-digit answer such as "CE," revealing that C=1. When it comes time to make a guess, a participant could guess C=1. However, such a guess would not

provide any new information about the code, since this part of the code is already known. A more efficient alternative would be to make a guess about any other letter, even a letter not involved in the equation (e.g., H=8). There is a chance (if only a small one; a 10% chance if no letters are known) that this single guess is correct, and reveals new information about the code. Thus, as participants work through the code, it is more efficient to guess about letters that they don't know than those they do.

Although the letters-to-numbers task was used in a series of studies by Laughlin and colleagues (Laughlin et al., 2002; Laughlin, Hatch, Silver, & Boh, 2006; Laughlin, Zander, Knievel, & Tan, 2003), these studies will not be reviewed in detail here as the results are only tangent to the current topic. However, the task has been useful in learning about group problem-solving performance, yet certain aspects of it remain relatively unexplored. What was not examined much in previous studies is the group process that led to performance outputs. The studies cited above generally follow an input-output method of understanding group problem-solving, with only brief insights into group interaction processes. For example, Laughlin et al. (2006) compared performance of individuals and groups with two, three, four, and five members. Performance and steps to solution (equations and guesses) on the letters-to-numbers task were examined, but the authors could only speculate as to what social processes were at work within the groups, or why they would have performed better or worse than expected due to a number of factors. The results demonstrated that groups—specifically those with at least three members—outperformed groups with two members (dyads) and individuals, however it is not clear what produced this pattern of results. Further, results suggested that groups outperform even what their best individual members would be capable of—an extremely

rare finding in problem-solving research, which typically supports the idea that groups perform better than their average member would, but not better than their "best" members, who often outperform interactive groups.

Previous studies suggest that there is something about the letters-to-numbers task that proves advantageous to interactive problem-solving groups, and allows them to perform the task very efficiently. One possibility offered by Larson (2010) involves diversity of member strategies: as problem-solving groups increase in size, there is a higher probability that the groups' members would propose different strategies to use to solve the letters-to-numbers code. The highly demonstrable nature of the task allows for an individual member with a strategy idea (such as an effective type of equation to use) to demonstrate its effectiveness to their fellow group members. The group can then debate the merit of the strategy and choose whether or not to use it. For this reason, a key component of understanding why groups are effective letters-to-numbers solvers may rest in an understanding of how intragroup diversity affects group process and performance.

<div align="center">**Diversity in Groups and Teams**</div>

As is the case with categorizing group tasks, various systems of defining and measuring intragroup diversity are available (see Mannix & Neale, 2005, for a broad review). Many methods utilize separate subtypes of diversity as qualitatively distinct and impacting groups in different ways, similar to the McGrath (1984) model of group tasks. One recent system is Carton & Cummings' (2012) Typology of Subgroups in Work Teams. Although the authors focus on subgroups within teams (that is, diverse groups within a group), their criteria for measuring diversity is useful in highlighting different varieties of diversity within a group or team, and how these different types of diversity

may affect group process and performance, even if diversity is measured as differences between individuals within a group.

The first type of diversity addressed by Carton & Cummings (2012) is based on identity. Drawing heavily from social identity theory (Tajfel & Turner, 1986), identity diversity follows the logic that individuals view themselves and their group members on the basis of social categories, such as age, gender, occupation, etc. Social identification as a member of a social category involves a feeling of "oneness" with that category as an actual member or symbolic prototype of that category (Ashforth & Mael, 1989). Within a workgroup, judgments that oneself or one's fellow group members belong to a social category are made relatively quickly, since the criteria for doing so are often readily apparent if one has even basic demographic knowledge and visual confirmation of their group members' characteristics. Carton & Cummings (2012) argue that the presence of multiple social identities for members within a workgroup can potentially undermine the sense of a single workgroup identity, which may breed intragroup competition in defense of individual members' social identities at the cost of maintaining a singular group or team identity. This tension may impact team dynamics as the competitive drive for maintaining self-concept counters the desire to achieve a unified team identity.

The second type of diversity considered by Carton & Cummings (2012) is based on distribution (or ownership) of resources within the team and is set in the framework of social dominance theory (Levin, Federico, Sidanius, & Rabinowitz, 2002). Disparate control over finite resources or differences in authority, status, or power within a group may cause tension between group members if such differences are perceived as unfair or unjustified (Mannix, 1993; Sidanius & Pratto, 1999), particularly in cases where lower-

power members are at the mercy of high-power members. However, there may also be advantages to groups with unequal distribution of resources. Some research suggests that if power is centralized within an individual or subgroup within a team, there can be a marked benefit to overall efficiency and work output of the team on some tasks, compared to groups with decentralized structures or no formal authority structure (Bunderson & Boumgarden, 2010).

Similar divergent outcomes of diversity are also apparent in Carton & Cummings' (2012) third type of diversity, which centers on differences in knowledge and information processing styles. Citing early organizational systems theories (Ashby, 1958), the authors suggest that development of subgroups with different sources of knowledge, and ways of using that knowledge, is a natural occurrence in organizational workgroups as a consequence of evolving demands placed on the organization. The need for specialization prevents completely overlapping knowledge, perspectives, and practices within an organization, and potentially even within small interacting workgroups. Put simply, it is not practical for every person in an organization to be capable of addressing every problem or need of their organization or even their local workgroup. This need for specialization often manifests in the form of workgroups with members who do not have the same knowledge, skills, or general work styles. A workgroup will likely have a single manager, who has different responsibilities from the employees that he/she supervises, that creates not only a power differential but the manager may also have unique skills such as administrative knowledge. The employees may have specialized technical skills or personal working styles that no other member of the work group possesses. Carton & Cummings (2012) discuss several potential outcomes for these scenarios where

workgroups consist of knowledge-diverse (or heterogeneous) vs. homogeneous

membership.

Such knowledge diversity may potentially lead to either positive or negative

outcomes. In terms of negative consequences, the main issue discussed by Carton &

Cummings (2012) is the impact of knowledge diversity on shared mental models for

group tasks. Mental models, a concept borrowed from cognitive research (Brauner, 1996;

Cannon-Bowers, Salas, & Converse, 1993; Hinsz, 1996), are mental representations of

how a system or mechanism operates (Tindale, Meisenhelder, Dykema-Engblade, &

Hogg, 2001). Groups generally function better when their members share mental models,

and have similar understandings of their task. As an example, sports teams rely on their

players understanding not only the rules of their sport, but also on the players

understanding each other's role on the team and what everyone should do in a specific

situation. When a play is called in a football game, each player must know not only what

he should do, but should also have a sense of where his teammates will be and what

they'll be doing. If a team does not have a shared mental model, this may impair the

team's ability to function as a cohesive unit (such as football players on the same team

crashing into each other).

However, consequences of knowledge diversity are not wholly negative in Carton

& Cummings' (2012) model. The authors argue that diversity in knowledge also prompts

a workgroup to consider alternative sources of knowledge beyond their personal

knowledge base. Even in seemingly productive groups, a lack of debate or disagreement

between group members may stifle innovation and ultimately inhibit the group's ability

to perform at an optimal level. On the other hand, too much disagreement also limits the

group's ability to function if member's disagreements prevent progress. Thus, the

relationship between inter-member disagreement and group performance can be

described as curvilinear—moderate disagreement that is enough to encourage thoughtful

discussion, but doesn't prevent the group members from working together, seems to be

the ideal scenario.

Despite this logic, there is no reason to assume that groups actually behave in this

way. Like-minded group members failing to debate unpopular, but ultimately more

productive, solutions to problems the group faces is an interesting concept, but does it

actually occur "in nature"? Do real-world groups engaged in very low or very high levels

of disagreement demonstrate predicted shortcomings in process and performance?

Examples of both basic and applied research studies suggest that such outcomes do occur.

Within the group decision-making literature, a recurring finding is that groups whose

members share common knowledge or information will too heavily rely on that shared

information to aid in making a decision, even at the cost of neglecting "unshared"

information (known to one or more, but not all, group members) that might have led to a

better decision (Larson, 2010). To demonstrate this behavior, researchers commonly

utilize what is often referred to as a "hidden profile" paradigm, (Stasser & Titus, 1985).

In these studies, researchers will task a group of research participants with making a

decision as a group (such as identifying the likely suspect in a hypothetical murder

mystery), and distribute information packets among them to help make that decision

(such as the "evidence" about the murder). Unbeknownst to the participants, they receive

different information in their packets—some details are common among all or the

majority of group members (shared information), while some information is given to only

one or a minority of members (unshared). The task is designed so that the unshared

information must be used to make the correct decision (identifying the correct suspect). If

shared information is relied on too heavily, it will lead the group to make the wrong

decision, one that is not supported by the full set of information, including the unshared

details. Results of such studies consistently demonstrate a tendency of groups to spend

more time discussing shared information, relying more heavily on shared information

when making their decision, and neglecting or disregarding the unshared information.

This effect has been demonstrated with a variety of tasks (Hollingshead, 1996; Kelly &

Karau, 1999; Larson, 2010; McLeod, Baron, Marti, & Yoon, 1997; Stasser & Stewart,

1992; Stasser, Stewart, & Wittenbaum, 1995; Stewart & Stasser, 1998; Scholten, van

Knippenberg, Nijstad, & De Dreu, 2007; Schulz-Hardt, Brodbeck, Mojzisch,

Kerscheriter, & Frey, 2006) and with different participant populations (Christensen,

Larson, Abbott, Ardolino, Franz, & Pfeiffer, 2000). This finding maps nicely onto the

theories that groups who spend too much time agreeing on what they have in common,

and not enough time disagreeing and debating, may actually perform worse as a result.

Applied researchers have found evidence of similar effects with real-world

workgroups. Franz (2012) tells the story of Jeffrey Jolton, PhD, an organizational

consultant, who studied a "team that got along too well" at a manufacturing firm. The

president of the firm was concerned that his engineering teams were not realizing their

full potential, and lacked innovative insights into production issues. Jolton conducted

assessments and interviews with the engineers, and the results demonstrated a familiar

pattern. It was apparent that the engineers not only had similar backgrounds and

knowledge, but also reported being averse to conflict, preferring to agree with colleagues

instead of expressing opposing viewpoints and encouraging debate. Jolton implemented an intervention plan designed to "stir up" the engineering teams, which included appointing individuals to the roles of "devil's advocate" (tasked with arguing against colleagues' ideas), "master of ceremony" (tasked with encouraging discussion from all members), "librarian" (tasked with finding novel innovations and techniques outside the engineers' experience) and "judge" (tasked with monitoring the behaviors of those assigned to other roles to ensure active input). Within a few weeks, teams reported better energy and more positive team dynamics, and the business successfully grew. It would seem that "getting along too well" potentially inhibits teams in a variety of contexts.

Carton & Cummings (2012) argue that for groups to take full advantage of knowledge diversity, they must balance the two potential outcomes stemming from this form of diversity: considering alternative sources of knowledge, but also finding common ground to synthesize that knowledge. Too much emphasis on either leads to suboptimal group performance. If the focus is on considering alternative sources of knowledge, this may prevent a shared mental model from emerging. Alternatively, a group too focused on developing and adhering to a shared mental model may discourage consideration of alternative sources of knowledge. One such source of knowledge that may affect groups is advice or counsel from someone outside the group. The extent to which groups and individuals consider such advice has been investigated in recent research studies, and is discussed in the next section.

### Advice-Taking by Groups

The degree to which groups will consider alternative perspectives before making a consensus decision dates back to at least Janis' (1982) groupthink theory. A component

of this theory is that groups often fail to consider information or "advice" from sources outside the group when making a decision. This notion largely evaded empirical scrutiny until very recently. Minson & Mueller (2012) demonstrated that groups (dyads), relative to individuals, have a tendency to underutilize advice when making a judgment. In their study, both dyads and individual participants were asked a series of general knowledge questions (e.g., What percentage of Americans own pets?). After making initial estimates, participants were given the opportunity to view estimates of the same quantities as given by a previous participant in the study, and could consider that information before rendering a final judgment (with the opportunity to change their own answer if they so desired). Dyadic teams were significantly less influenced by these previous participants' answers then were individuals. This behavior appeared to inhibit performance, since estimates would have been more accurate if dyads had given the previous participants' answers more weight when making their final judgments. Simple probability dictates that this should be the case—when forming such judgments, an average of several individuals' judgments will likely be more accurate than any individual's judgment. These results, along with previous findings, suggest that both individuals and groups are likely to benefit when considering outside advice.

The authors attributed this effect to groups' higher levels of confidence in their ability to make accurate judgments. Working with others has been shown to promote higher confidence in one's decision-making ability (Forsyth, 1999; Park & Hinsz, 2006), as well as a reduced willingness to consider advice from others outside the group relative to individuals (Gino & Moore, 2007; Harvey & Fischer, 1997; Soll & Larrick, 2009). The finding that groups may underutilize information presented from outside the group to an

even greater extent than individuals—even if that information is useful—is troubling

given that most people assume groups are generally better than individuals at making

important decisions (Kerr & Tindale, 2004; Larson, 2010). While groups may outperform

individuals in many contexts (Tindale, Kameda, & Hinsz, 2003), they are not universally

superior and are prone to specific flaws in information processing, such as a failure to

take advice as suggested by Janis (1982).

In considering the implications of the Minson & Mueller (2012) findings, it is

important to take note of exactly when outside information (advice) was made available

for groups to consider. In that study, groups had access to outside advice only after they

had reached consensus. This is a critical feature given the importance of consensus as a

part of the group interaction process. Reaching consensus significantly increases

members' confidence in their group's performance (Tindale, 1989). In short, the very

existence of consensus may render the group particularly reluctant to consider outside

advice. One potential reason for this effect is that the group members reached a state of

cognitive closure, which refers to the inclination for individuals to seek a sense of

confirmation once they have made a decision, and a subsequent reluctance to accept

ambiguity regarding the accuracy or merit of that decision (Kruglanski, 1990, 2004;

Kruglanski, Dechesne, Orehek, & Pierro, 2009; Kruglanski & Fishman, 2009; Kruglanski

& Webster, 1996). Two underlying tendencies are thought to promote this process. The

first, a tendency for urgency, involves a desire to "seize" a definitive decision as quickly

as possible. The second, a tendency for permanence, seeks to maintain this definitive

state ("freezing" on knowledge already obtained or decisions already made). In the

context of consulting outside advice when making a decision, groups may adopt their

initial consensus judgments rapidly, followed by an aversion to any advice offered after that decision has been made because the advice threatens the permanence of the initial decision.

Reaching a state of cognitive closure may explain why such groups would be reluctant to utilize any information presented from an outside source. But what of groups that have not reached consensus—is there an opportunity to introduce outside information before the members "close" their minds? A study conducted by Greitemeyer & Schulz-Hardt (2003) found evidence that such an opportunity might occur during group discussions. The authors found that when a decision-making group is offered decision-relevant information to consider before members commit to an initial opinion, the members are better able to recall the information and more likely to notice inconsistencies in information available to different group members. This inconsistency serves as a cue that the information available to any single member may be flawed, and that the combination of members' information may reveal an optimal decision outcome that could only be reached through such a combination (but not by means of any one members' information). In short, there may be specific moments within the group's decision process during which outside information is more likely to be attended to by the group.

This idea was tested in a recent study by Bihary, Larson, & Tindale (2015), who investigated whether making outside advice available to a group before the group reached a consensus decision would increase the chances that the advice would be considered or used. Dyads and individuals were tasked with making quantity estimation judgments as answers to general knowledge questions. After making an initial set of judgments,

participants received advice in the form of a previous participant's answers to the same

questions. The participants then had an opportunity to revise their initial judgments if

they wished. Three between-subjects conditions had participants make judgments either

in 2-person groups for both sets of judgments, as individuals for both sets of judgments,

or individually for the first set and as a dyad for the second set. Additionally, the third

condition was later median-split into two conditions based on differences (large vs.

small) between dyad members' initial judgments. It was predicted that dyads with larger

differences (higher diversity) between members' initial judgments would be more likely

to utilize outside advice, relative to dyads with smaller differences (lower diversity)

between members' initial judgments. Results supported this prediction; high-diversity

dyads were significantly more likely to consider the outside advice. Members' confidence

in their group's ability to perform was measured as a potential mediator of this effect, but

no support for this hypothesis was found.

**Current Studies**

The current studies attempted to extend these results and expand upon the

previously observed effects of diversity on group process and performance. Particular

interest is on the extent to which relatively heterogeneous groups (dyads whose members

have different perspectives) will consider and use advice presented from outside the

group when performing a task collectively, compared to relatively homogeneous groups

(dyads whose members have the same or very similar perspectives). It was predicted that

the heterogeneous groups would be more likely to consider and use advice from outside

the group that offers an alternative perspective. Building from Carton & Cummings'

(2012) category of knowledge diversity, the current studies examined how diversity in

problem-solving strategies would affect group problem-solving performance. Other relevant factors, such as confidence in one's abilities or performance, were also examined in the interest of advancing our understanding of both group performance and group interaction processes.

Most studies on diversity, particularly those in applied organizational settings, rely on pre-existing states of diversity. One advantage of the current experimental approach is that the diversity factor of interest (problem-solving strategies) could be experimentally manipulated in a controlled fashion. Letters-to-numbers is a highly customizable, computerized task. This allows for incredible flexibility in modifying the task as needed. Specific to this study, it allowed for direct manipulation of strategy diversity. When participants submit an algebraic equation to the program in order to get information about the code, they may have a reason for selecting a certain equation to use. A participant's method of selecting equations can be directly manipulated the same way a training program might train an employee to perform a work task in a specific way. Participants can thus be "trained" to perform the letters-to-numbers task in a specific way, and then paired with a teammate who was trained to use either a similar or different approach. Further, the program can be manipulated to ensure that a participant's strategy training leads them to solve the problem with reasonable (though suboptimal) effectiveness, and the program can require that the participants adhere to their given strategies.

Additionally, the use of the letters-to-numbers task in the present design allowed for multiple measures of group process and performance. The program tracks the entire solution (or attempted solution) method, including equations proposed and guesses made

about the code. Drawing from previous studies using the task, performance on the task

can be measured by the number of trials used to solve the code, the complexity of

equations proposed (with more complex equations suggesting heightened motivation to

process the more complex information), the guessing strategy (whether guesses are

merely confirming part of the code that participants should already know, or used to

explore unknown parts of the code), and the efficiency of progress toward solving the

code.

      The demonstrable nature of the task was also hypothesized to encourage higher

levels of group interaction than what might be observed in tasks that are less

demonstrable. Certain predictions depended on observational measures of group

interaction to advance our understanding of group processes.

CHAPTER TWO

STUDY 1

Study 1 investigated the role of diversity in group problem-solving process on a group's willingness to consider outside advice when trying to solve a letters-to-numbers problem. Participants attempted to solve two letters-to-numbers problems. For the first problem, individual participants were given one of two simple strategies (methods of creating input equations) to use, and were required to follow their given strategy to solve the letters-to-numbers code. For the second problem, participants were given a more complex, but potentially more effective, strategy to consider. They then attempted to solve the second code paired with a teammate. In contrast with the first problem, the team was not restricted to using any specific strategy on this second problem. Participants were randomly assigned to either a Same- or a Diverse-strategy condition, so that dyad members' first strategy either matched their partner's or was different. The letters-to-numbers program was designed to make all participants' given strategies seem effective for solving the problem (described further in the Method section below). All sessions were video-recorded to allow for subsequent analysis of group behaviors. The following predictions were tested:

Hypothesis 1: Diverse-strategy dyads will outperform Same-strategy dyads on the letters-to-numbers task. Better performance was defined as solving more letters per trial used, and using fewer trials to solve the full code.

Hypothesis 2: Diverse-strategy dyads will demonstrate superior letters-to-numbers information processing than Same-strategy dyads. Superior information processing was defined as proposing more complex equations (Hypothesis 2A) and more exploratory (vs. confirmatory) guesses (Hypothesis 2B).

Hypothesis 3: Diverse-strategy dyads will be more likely to use outside advice (i.e., use the more complex but potentially more efficient Add All Unknown strategy), than Same-strategy dyads.

Hypothesis 4: The tendency to use the more complex Add All Unknown strategy will mediate the relationship between strategy diversity condition and letters-to-numbers performance, with Diverse-strategy dyads being more likely to use the Add All Unknown strategy.

Hypothesis 5: Diverse-strategy dyads will self-report more productive discussions (Hypothesis 5A) and better perceived performance (Hypothesis 5B) than Same-strategy dyads on post-session questionnaires.

Hypothesis 6: Diverse-strategy dyads will be rated by observers as having more productive discussions than Same-strategy dyads.

## Study 1 Method

### Participants

One hundred thirty-two participants were recruited for Study 1. A power analysis, based on results of previous studies conducted by the investigator, suggested that this number should be sufficient to produce reasonably reliable results (80% power) with a small to moderate effect size. Undergraduate students from introductory psychology

classes were recruited for the study. Participants were randomly assigned to one of three

strategy diversity conditions, and attempted to solve two letters-to-numbers problems

(described below). Participant assignment was equal across conditions, with 44

participants (22 dyads) in each condition. For their participation, students earned credits

toward fulfilling a research participation course requirement. Participants also had the

opportunity to enter a raffle to win a prize.

<div align="center">Experimental Task</div>

Participants performed two separate letters-to-numbers problems (Laughlin, 2011;

Laughlin et al., 2002; Laughlin et al., 2003; Laughlin et al., 2006) that were presented via

computer. For each problem, the numbers 0-9 were randomly coded as the letters A-J

with no repetitions (each letter represented a unique number). The participants' goal was

to figure out which number each letter represented, and to accomplish this using as few

"trials" as possible. To get hints about the code, participants submitted algebraic

equations to the program using the letters A-J and the mathematical operations of

addition and subtraction (e.g., A+B=?). The program provided answers to the equation

entered (e.g., A+B=CJ). After receiving the answer to their equations, participants then

had the opportunity to guess at the number value of one letter (e.g., C=1), and were

informed whether or not this guess was correct. Participants then proposed a mapping for

the full code (all letters and all numbers), and were told whether the full code was correct

or not in its entirety. One trial consisted of these six steps (proposing an equation,

receiving an answer to that equation, making a guess about one letter, receiving feedback

about that guess, proposing a full code mapping, and receiving feedback about the full

code), and participants continued through a series of trials until they either solved the full

code, exceeded the trial limit (20), or exceeded the time limit (15 minutes), at which point the program terminated.

Experimental Conditions

In order to experimentally manipulate diversity of equation strategies on the first letters-to-numbers problem, participants were assigned to one of three strategy diversity conditions. Similar to earlier letters-to-numbers studies (Laughlin et al., 2003), participants were forced to use a specific strategy to create equations. One strategy (referred to here as the "Add Two" strategy) involved constructing equations that added two letters (e.g., A+B, E+E). A second strategy (referred to as "Use Three") involved using equations consisting of three letters and either one or two types of mathematical operations (either addition or subtraction or both). This second strategy allows for a larger variety of equations (e.g., A+A+A, A+D+F, C-D+J, D-B-G, BD+F, C-HJ) compared to the first. These two strategies were selected because, while they are different, previous research suggests that participants using equations with either two or three letters do not differ in their overall performance on letters-to-numbers tasks (Laughlin et al., 2003).

Since participants were obligated to use their assigned strategy, an additional feature of the letters-to-numbers program was implemented to suggest that a participant's strategy was useful and helpful. Participants were told that the letters-to-numbers code was randomly generated. While this was true for the second problem, it was only partially true for the first problem. In truth, there was no set code for the first problem until participants submitted their first equation to the program. When this first equation was entered, the program deliberately assigned values to the letters used in the equation so

that the answer was always a two-digit number, and then assigned remaining numbers to the rest of the letters at random. For example, if a participant submitted the equation "A+B=?" the program could fix the value of A to 9 and B to 8, producing an answer of 17, and randomly assign the numbers 0-7 to the remaining letters. The intent of this feature was to make it easier for participants to solve the code then if, for example, A+B was equal to a single-digit answer. Two-digit answers produced from the two equation strategies were limited to including only 1 or 2 as the first digit of the answer, although other possibilities may have been evident from the participant's point of view for participants assigned to use the Use Three strategy (e.g., AB+C could represent 86+5, producing an answer of 91). However, equations of this style (a single-digit number added to or subtracted from a two-digit number, versus some combination of three one-digit numbers) were rare; only about 6% of participants assigned to use the Use Three strategy used such equations on the first trial of the first problem. Conducting analyses with these cases removed did not alter the results.

Each experimental session included two participants, who first individually attempted to solve one letters-to-numbers problem using one of the strategies described above. The three experimental conditions included two conditions in which participants were instructed to use the same strategy (Add Two or Use Three) and one condition in which one participant was instructed to use the Add Two strategy, while the other was instructed to use the Use Three strategy. The letters-to-numbers program restricted the equations that could be entered, so that if a participant attempted to submit an equation that didn't match his/her assigned strategy, the program rejected it and asked for a

different equation. Participants were not informed of each other's given strategies, nor whether they were given the same or different strategies.

Procedure

Participants were escorted from the participant waiting area to the laboratory room by the experimenter. Each session included two participants in the same room, working at computers on different sides of a large table. Upon arrival, the experimenter provided an overview of the experimental session as part of the informed consent process. The overview included a brief explanation that participants would be working on computers to complete a "code-breaking" task. Participants were not expressly told that they would be working together as a team, or other details of experimental procedures, to avoid suspicions of the purpose of the study. They were also notified that the session would be video-recorded, with the camera identified by the experimenter in plain sight on a shelf in the room. The experimenter explained that research participation credits and a chance to enter a raffle were offered as compensation for participating. After the explanation, the experimenter answered any questions participants had, and asked them to sign the informed consent document.

Participants were next introduced to the letters-to-numbers program. An introduction screen briefly described the basic structure and goal of the task. After giving participants a minute to read this screen, the experimenter asked if participants understood the description and if they had any questions about the task at this point. Participants next completed a short tutorial that walked them through a few trials of a letters-to-numbers problem, step-by-step, and explained relevant features of the program along the way. This tutorial was expected to take 5-10 minutes to complete, and

participants were asked to inform the experimenter when they had finished. Since the two participants in the session may have finished the tutorial at different times, the first one to finish was asked to wait until the other participant had finished.

Once both participants finished the tutorial, the experimenter reviewed the basic goal of the task, and emphasized that the participants' goal was to solve the code in as few trials as possible, or at least try to figure out as much of the code as possible (to allow for participants who may not solve the full code). Participants could again ask any questions that they had regarding the task or the letters-to-numbers computer program. After questions were answered, the experimenter introduced the equation strategy manipulation described earlier. The strategies were presented as aids to assist participants in solving the letters-to-numbers code. Participants were told that the program would require them to follow the strategy for the first problem, but that they would attempt to solve a second problem later that did not have this restriction. Strategies were provided on paper sheets distributed to each participant (to prevent participants from knowing each other's strategy), with a brief description of the strategy and several examples of equations that fit that strategy (see Appendices A and B). The experimenter placed a divider on the table so that participants could not see each other, and instructed them to begin the first problem. When finished, a participant would notify the experimenter and wait for the other participant to finish the task.

After solving the first problem (or reaching either the time or trial limit and thus failing to solve it), participants completed a short questionnaire (see Appendix D) that asked for their perceptions of their performance on the first letters-to-numbers problem. Questions asked about perceived difficulty of the task and asked participants to rate their

performance in relation to other students who might participate in the experiment. This was used as a measure of confidence in their problem-solving abilities.

After completing the questionnaire, the two participants in the session were teamed up (as a dyad) to attempt to solve a second letters-to-numbers problem as a team working on one computer. Deciding logistics of the team's setup (which participant's computer would be used, which side of the table they would sit on, etc.) was left up to participants. They were given a third equation strategy (on a paper sheet, similar to strategies for the first problem, see Appendix C) to consider to use for solving this second problem, however, unlike the first problem they were not forced to use it. Participants were told that they could consider this strategy alongside the strategies they used for the first problem, and could decide with their partner how they should ultimately work through the problem and which strategies (if any) they should use. They were able to switch strategies or freely abandon a strategy at any time. The third strategy, referred to here as the "Add All Unknown" strategy, involves adding all of the unknown letters together. This strategy can be particularly effective if used correctly. When all letters are unknown, adding all letters A-J in one equation (A+B+C+D+E+F+G+H+I+J=?) always produces an answer of 45, providing the letter codes for both 4 and 5. This can be followed with a second equation adding all of the letters except for those that equal 4 and 5, producing 36 and two more letter solutions, and so on with subsequent equations producing answers of 27 and 18. Clever guesses will reveal the letters representing the remaining digits (0 and 9) during these trials, allowing the code to be solved in four trials using this strategy. Participants had the same goal of solving the code in as few trials as possible, or at least solving as much of the code as possible. The experimenter

encouraged participants to discuss their strategies used for the first problem, their

performance on the first problem, and the Add All Unknown strategy. However,

participants were asked to begin the second problem (initiating the letters-to-numbers

program and starting the 15-minute countdown) before beginning discussion, to avoid

variations in the amount of time teams discuss before starting the second problem. A

raffle prize (one of three $50 gift cards) was also offered as a performance incentive—

participants all began with 20 "tickets" to enter the raffle, and lost a ticket for each trial

they used to work through the problem.

When participants solved the second problem (or ran out of trials or time), they

each individually completed a second set of questionnaire measures that included

questions about perceived performance similar to the first questionnaire (but addressing

the team's performance) as well as questions regarding the team's general social

behaviors. Participants were unable to view their teammate's answers or to discuss the

questionnaire with their teammate.

After both participants completed the second questionnaire, the experimenter

verified whether participants wanted to be entered into the raffle or not. For those that

did, the experimenter noted the number of raffle tickets participants earned, based on

their team's performance on the second problem. Physical raffle tickets were marked

with participant ID codes and placed in a container. Raffle winners were selected at the

end of the semester, and contacted via email. Prior to selecting winners, any between-

conditions differences in performance were balanced so as not to favor participants in any

one condition. That is, after determining the average number of raffle tickets earned by

participants in each condition, all conditions scoring lower than the best condition had

extra tickets added to even out the between-conditions averages. Lastly, the experimenter

informed participants that their research credits would be posted within the next day, and

that a debriefing email would be sent out to all participants at the end of the semester.

Participants were then thanked and dismissed.

<div align="center">Measures</div>

**Letters-to-Numbers Information Processing**

In order to assess participants' methods of working through the problem, two

measures of letters-to-numbers problem-solving processing were used. The first, *equation

complexity*, is measured as the average number of letters used in equations. More

complex equations (using more letters) are considered to be demonstrably superior to

simpler equations since they potentially reveal more information about the code

(Laughlin, 2011). The second measure of letters-to-numbers process is *guessing strategy*.

When participants make guesses about the number assigned to a single letter (e.g., A=2),

these guesses can be either confirming something participants should already know

(confirmatory) or be a guess about part of the code that they could not yet know

(exploratory). For example, if a participant were to submit the equation "A+B=?" and the

program provided the answer "A", it is clear that B=0 and A is some other number.

Guessing "B=0" would be a confirmatory guess, while guessing at the value of A would

an exploratory guess (assuming the guess is not "A=0"). Exploratory guesses are more

likely to reveal new information about the code (Larson, 2010), and are thus more

efficient than confirmatory guesses.

**Letters-to-Numbers Performance**

Letters-to-numbers performance has traditionally been measured by trials-to-solution (the number of trials participant used to solve the full code), with fewer trials-to-solution indicating better (more efficient) performance (Laughlin, 2011; Laughlin et al., 2002; Laughlin et al., 2003; Laughlin et al., 2006). However, using trials-to-solution as a measure of performance assumes that all participants successfully solve the code. This is not always the case, as participants may reach one of two failure states (reaching either the time limit or maximum number of allowed trials) without solving the code. Although this was accounted for in earlier studies by assigning failures-to-solve a score of *maximum number of trials + 1*, this method potentially distorts the measure, particularly if many participants fail to solve the code. For example, in a study that allowed participants a maximum of 10 trials, participants who failed to solve received a score of 11 trials, regardless of how much of the code they had solved. Thus, participants who performed differently (in terms of how much of the code they had solved) could receive the same score, artificially reducing the variance in the measure.

Thus, an additional performance measure was also used. It is the average number of letters solved per trial used, calculated by dividing the total number of letters solved correctly by the number of trials used. This measure, used in some but not all previous letters-to-numbers studies (Laughlin, 2011; Laughlin et al., 2002) avoids the distortion potentially introduced by the simpler trials-to-solution approach, and accounts for the full range of possible performance scores. Authors of previous letters-to-numbers studies did not specifically report the number of participants in their studies who failed to solve the problems, nor did they specify if a time limit was imposed that may have impacted

performance. If only a small number of participants—relatively even across conditions—

fail to solve the letters-to-numbers problems, it may be possible to drop their data from

the analysis and proceed with the trials-to-solution measure. However, due to concerns

that this method also distorts or misrepresents results, the letters-per-trial measure may

need to be used as the performance measure. Due to a number of participants in multiple

conditions failing to solve the code (discussed under results for Hypothesis 1), this

alternative measure is used in analyses.

**Confidence**

Confidence was measured via the self-report questionnaire (see Appendices D and

E), with questionnaires administered to individual participants after finishing each letters-

to-numbers problem. Participants were asked to rate their confidence in having performed

the task better than 25%, 50%, and 75% of other participants (compared to other

individuals after the first problem, teams after the second) in the study. Initial analyses

used the "better than 50%" item for measuring confidence, however alternative measures

(such as combining the three confidence items into a single measure, or substituting the

"better than 25%" or "better than 75%" alternative measures) were also explored.

**Team Dynamics**

The interaction between team members was measured using both self-report and

observational methods. For self-report measures, participants completed a questionnaire

after finishing the second task (see Appendix E). The questionnaire included items

addressing perceptions of social dynamics such as degree of conflict vs. cooperation

within the team, leadership, and perceptions of how much (if at all) discussions benefitted

the team. Sessions were also video-recorded to allow for coding of behavioral evidence

of these constructs, including frequency counts of observed behaviors and scale ratings of

team interactions (see Appendix G), by an observational analyst who was blind to

hypotheses. Elements of this coding scheme were tested for reliability in a pilot study.

Matched subsets of the coding data recorded independently by the principal investigator

and a research assistant correlated strongly ($r = .80$ or higher) and consistently across

behavior categories, suggesting high reliability for the coding scheme.

## Study 1 Results

### Hypothesis 1

I predicted that Diverse-strategy dyads would outperform Same-strategy dyads

(both Add Two and Use Three conditions) on the letters-to-numbers task. Many

participants failed to solve one or both letters-to-numbers problems. 70% of participants

failed to solve the first problem individually, and approximately 14% of groups failed to

solve the second problem. Due to these failures to solve one or both letters-to-numbers

problems, the average number of letters solved per trials used was used to measure

performance (instead of total trials used to solve the full code, as in previous studies

utilizing the letters-to-numbers task). The analysis used to test Hypothesis 1 was a 2

(Letters-to-numbers problem: First, Second) x 3 (Strategy diversity condition: Same-

strategy/Add-Two, Same-strategy/Use-Three, Diverse-strategy) mixed-model analysis of

variance (ANOVA) on participants' letters-to-numbers performance, with problem

number as a within-subjects factor and strategy diversity condition as a between-subjects

factor. The within-subjects factor was comprised of two levels: (1) the average of the two

team members' performance on the first problem and (2) the team's measure from the

second problem. The main effect of letters-to-numbers problem was significant, $F (1, 63)$

= 119.07, $p < .001$, partial $\eta^2 = .65$. This suggests that participants' performance on the second letters-to-numbers problem ($M = 1.39$, $SD = .51$) was significantly better than their performance on the first problem ($M = .69$, $SD = .35$). The main effect of strategy diversity condition was not significant, $F(2, 63) = .10$, $p = .902$, partial $\eta^2 = .00$. This suggests that participants' performance did not differ between participants in the Diverse-strategy ($M = 1.03$, $SD = .40$), Same-strategy/Add-Two ($M = 1.07$, $SD = .41$), and Same-strategy/Use-Three ($M = 1.03$, $SD = .47$) conditions. The interaction between letters-to-numbers problem and strategy diversity condition was also not significant, $F(2, 63) = 1.87$, $p = .162$, partial $\eta^2 = .06$, suggesting that the relationship between strategy diversity condition and performance did not differ between the two letters-to-numbers problems. A planned contrast revealed that there was no difference in performance between Diverse-strategy groups and groups in the two Same-strategy conditions, $t(63) = .00$, $p = .432$ (one-tailed), $r = .02$.

Two additional exploratory analyses were conducted. An additional mixed-model ANOVA was conducted, replacing the average of two members' performance on the first problem with the better individual's performance score on the first problem in the within-subjects factor. The main effect of letters-to-numbers problem was significant, $F(1, 63) = 28.60$, $p < .001$, partial $\eta^2 = .31$. This suggests that participants' performance on the second letters-to-numbers problem ($M = 1.39$, $SD = .51$) was significantly better than their performance on the first problem ($M = 1.00$, $SD = .52$). The main effect of strategy diversity condition was not significant, $F(2, 63) = .06$, $p = .944$, partial $\eta^2 = .00$. This suggests that participants' performance did not differ between participants in the Diverse-strategy ($M = 1.18$, $SD = .49$), Same-strategy/Add-Two ($M = 1.22$, $SD = .48$), and Same-

strategy/Use-Three ($M = 1.20$, $SD = .57$) conditions. The interaction between letters-to-numbers problem and strategy diversity condition was also not significant, $F$ (2, 63) = 1.94, $p = .152$, partial $\eta^2 = .06$, suggesting that the relationship between strategy diversity condition and performance did not differ between the two letters-to-numbers problems. A planned contrast revealed that there was no difference in performance between Diverse-strategy groups and groups in the two Same-strategy conditions, $t$ (63) = .00, $p = .394$ (one-tailed), $r = .03$.

Finally, Hypothesis 1 was also tested with a one-way ANOVA, with performance on the second letters-to-numbers problem as the dependent measure and strategy diversity condition as the between-subjects factor. The effect of strategy diversity condition on performance was not significant, $F$ (2, 63) = .57, $p = .568$, partial $\eta^2 = .02$. This suggests that performance on the second letters-to-numbers problem did not differ between participants in the Diverse-strategy ($M = 1.46$, $SD = .46$), Same-strategy/Add-Two ($M = 1.42$, $SD = .50$), and Same-strategy/Use-Three ($M = 1.30$, $SD = .56$) conditions. A planned contrast revealed that there was no difference in performance between Diverse-strategy groups and groups in the two Same-strategy conditions, $t$ (63) = .75, $p = .229$ (one-tailed), $r = .09$. Thus, Hypothesis 1 was not supported.

## Hypothesis 2

I predicted that Diverse-strategy dyads would demonstrate superior letters-to-numbers information processing than Same-strategy dyads. Improved information processing is conceptualized as proposing more complex equations using more letters (Hypothesis 2A) and more exploratory (vs. confirmatory) guesses (Hypothesis 2B). I tested for differences in equation complexity in the second letters-to-numbers problem

with a one-way ANOVA, with strategy diversity condition as the between-subjects

factor. The effect of strategy diversity condition on equation complexity was not

significant, $F (2, 63) = 1.37$, $p = .261$, partial $\eta^2 = .04$. This suggests that participants'

equation complexity did not differ between participants in the Diverse-strategy ($M =$

3.36, $SD = 1.81$), Same-strategy/Add-Two ($M = 2.68$, $SD = 1.08$), and Same-

strategy/Use-Three ($M = 3.03$, $SD = 1.06$) conditions. A planned contrast provided

marginal evidence that Diverse-strategy groups may have used more complex equations

than groups in both Same-strategy conditions, $t (63) = 1.42$, $p = .08$ (one-tailed), $r = .18$.

Thus, Hypothesis 2A was only weakly supported by the marginal results of the planned

contrast.

Hypothesis 2B was tested with a 2 (Letters-to-numbers problem: First, Second) x

3 (Strategy diversity condition: Same-strategy/Add-Two, Same-strategy/Use-Three,

Diverse-strategy) mixed-model ANOVA, with letters-to-numbers problem as a within-

subjects factor and strategy diversity condition as the between-subjects factor. The

within-subjects factor was comprised of 2 levels: (1) the average of the two team

members' proportion of total guesses that were exploratory on the first problem and (2)

the team's proportion of total guesses that were exploratory on the second problem. The

main effect of letters-to-numbers problem was significant, $F (1, 63) = 32.14$, $p < .001$,

partial $\eta^2 = .34$. Somewhat unexpectedly, this suggests that participants proposed more

exploratory guesses on the first letters-to-numbers problem ($M = .68$, $SD = .18$) than they

did on the second problem ($M = .47$, $SD = .28$). The main effect of strategy diversity

condition was not significant, $F (2, 63) = 2.08$, $p = .133$, partial $\eta^2 = .06$. This would

suggest that participants' guessing strategy did not differ between participants in the

Diverse-strategy ($M$ = .64, $SD$ = .21), Same-strategy/Add-Two ($M$ = .57, $SD$ = .21), and Same-strategy/Use-Three ($M$ = .53, $SD$ = .27) conditions. The interaction between letters-to-numbers problem and strategy diversity condition was also not significant, $F$ (2, 63) = .28, $p$ = .759, partial $\eta^2$ = .01, suggesting that the relationship between strategy diversity condition and guessing strategy did not differ between letters-to-numbers problems. Nevertheless, a planned contrast suggested that Diverse-strategy groups proposed more exploratory guesses than groups in both Same-strategy conditions, $t$ (63) = 11.62, $p$ < .05 (one-tailed), $r$ = .23.

Two additional exploratory analyses were conducted to test Hypothesis 2B. An additional mixed-model ANOVA was conducted, replacing the average of the two team members' proportion of total guesses that were exploratory on the first problem with the better individual's proportion of guesses that were exploratory in the within-subjects factor. The main effect of letters-to-numbers problem was significant, $F$ (1, 63) = 84.01, $p$ < .001, partial $\eta^2$ = .57. Again, this unexpected result suggests that participants proposed more exploratory guesses on the first letters-to-numbers problem ($M$ = .82, $SD$ = .17) than they did on the second problem ($M$ = .47, $SD$ = .28). The main effect of strategy diversity condition was not significant, $F$ (2, 63) = 2.10, $p$ = .131, partial $\eta^2$ = .06. This suggests that participants' guessing strategy did not differ between participants in the Diverse-strategy ($M$ = .71, $SD$ = .21), Same-strategy/Add-Two ($M$ = .63, $SD$ = .21), and Same-strategy/Use-Three ($M$ = .61, $SD$ = .25) conditions. The interaction between letters-to-numbers problem and strategy diversity condition was also not significant, $F$ (2, 63) = .84, $p$ = .436, partial $\eta^2$ = .03, suggesting that the relationship between strategy diversity condition and guessing strategy did not differ between letters-to-numbers

problems. However, a planned contrast suggested that Diverse-strategy groups proposed more exploratory guesses than groups in both Same-strategy conditions, $t$ (63) = 16.56, $p$ < .05 (one-tailed), $r$ = .23.

Finally, Hypothesis 2B was also tested with a one-way ANOVA, with proportion of guesses that were exploratory on the second letters-to-numbers problem as the dependent measure and strategy diversity condition as the between-subjects factor. The effect of strategy diversity condition on guessing strategy was not significant, $F$ (2, 63) = 1.37, $p$ = .261, partial $\eta^2$ = .04. This suggests that guessing strategy on the second letters-to-numbers problem did not differ between participants in the Diverse-strategy ($M$ = .54, $SD$ = .29), Same-strategy/Add-Two ($M$ = .48, $SD$ = .23), and Same-strategy/Use-Three ($M$ = .40, $SD$ = .30) conditions. However, a planned contrast provided marginal evidence that Diverse-strategy groups may have made more exploratory guesses than groups in the two Same-strategy conditions, $t$ (63) = 1.37, $p$ = .08 (one-tailed), $r$ = .03. Thus, Hypothesis 2B was partially supported by the planned contrasts analyses. Participants proposing more exploratory guesses on the first letters-to-numbers problem overall was unexpected.

## Hypothesis 3

I predicted that Diverse-strategy dyads would be more likely to use outside advice (in the form of the more complex but potentially more efficient Add All Unknown strategy) on the second problem than Same-strategy dyads. The tendency to use this more complex strategy was measured as the proportion of trials on which the strategy was used. This hypothesis was tested with a one-way ANOVA, with strategy diversity condition as the between-subjects factor. The effect of strategy diversity condition on the

use of the Add All Unknown strategy was not significant, $F$ (2, 63) = .06, $p$ = .947,

partial $\eta^2$ = .00. This suggests that participants' tendency to use this strategy did not differ

between participants in the Diverse-strategy ($M$ = .10, $SD$ = .16), Same-strategy/Add-

Two ($M$ = .08, $SD$ = .15), and Same-strategy/Use-Three ($M$ = .08, $SD$ = .15) conditions.

A planned contrast revealed that there was no difference in the tendency to use this

strategy between the Diverse-strategy groups and groups in the two Same-strategy

conditions, $t$ (63) = .32, $p$ = .377 (one-tailed), $r$ = .04. Since the result was not significant,

no further analyses examining moderators or mediators of this effect were conducted.

An additional exploratory analysis was conducted to test Hypothesis 3. Since

some participant groups did not use the Add All Unknown strategy at all, and those that

did used it only a small number of times, a Pearson's chi-square analysis was conducted,

with the tendency to use the Add All Unknown strategy treated as a categorical variable

with two levels (participants did or did not use it) and strategy diversity condition

maintained as a categorical variable with three levels. Cell frequencies are provided in

Table 1. There was not a significant association between strategy diversity condition and

whether or not participants used the Add All Unknown strategy, $\chi^2$ (2) = .13, $p$ = .935,

providing additional evidence that there was no difference between strategy diversity

conditions in their tendency to use the strategy. Thus, Hypothesis 3 was not supported.

Table 1. Chi-square cell frequencies for Hypothesis 3.

| Condition | Did not use Add All Unknown | Used Add All Unknown |
|---|---|---|
| Same-strategy/Add-Two | 15 | 7 |
| Same-strategy/Use-Three | 14 | 8 |
| Diverse-strategy | 14 | 8 |

Hypothesis 4

I predicted that the tendency to use the more complex Add All Unknown strategy would mediate the relationship between strategy diversity condition and letters-to-numbers performance. The mediation analysis would have been tested with a path analysis approach and followed steps discussed by Baron & Kenny (1986). Unfortunately, the preliminary conditions necessary to test mediation (testing correlations between variables in the model) were not met. In order to test the prediction that the tendency to use the Add All Unknown strategy mediates the relationship between strategy diversity condition and performance, three preliminary conditions must hold: 1) strategy diversity condition and performance are correlated, 2) strategy diversity condition and the tendency to use the Add All Unknown strategy are correlated, 3) the tendency to use the Add All Unknown strategy and performance are correlated, while controlling for the relationship between strategy diversity condition and performance. The correlation matrix used to test these steps is presented in Table 2. As can be seen from the data, the necessary correlational relationships between variables are not present, and resulted in non-significant path coefficients in the preliminary mediation analyses. Thus, the mediation model predicted by Hypothesis 4 was not supported.

Table 2. Correlations between variables for all participants used to test mediation model for Hypothesis 4. Right-hand column provides standard deviations for all variables. None of the correlation values are statistically significant.

| | Strategy Diversity Condition | Proportion of Trials used Add All Unknown | *SD* |
|---|---|---|---|
| **Strategy Diversity Condition** | | | .823 |
| **Proportion of Trials used Add All Unknown** | .041 | | .152 |
| **Performance** | .033 | .011 | .506 |

Hypothesis 5

I predicted that Diverse-strategy dyads would self-report more productive

discussions (Hypothesis 5A) and better perceived performance (Hypothesis 5B) than

Same-strategy dyads on post-session questionnaires. Hypothesis 5A was tested with a

one-way ANOVA, with strategy diversity condition as the between-subjects factor. Item

9 from the P2-Post questionnaire ("How beneficial was discussing with your teammate in

helping your team perform better?") served as the dependent measure for productive

discussion. The effect of strategy diversity condition on self-reported discussion

productivity was significant, $F (2, 63) = 4.26$, $p < .05$, partial $\eta^2 = .12$. This would

suggest that productivity of discussion differed between participants in the Diverse-

strategy ($M = 8.62$, $SD = 1.54$), Same-strategy/Add-Two ($M = 7.42$, $SD = 1.74$), and

Same-strategy/Use-Three conditions ($M = 8.57$, $SD = 1.29$). A planned contrast revealed

a marginal effect suggesting that Diverse-strategy groups may have reported more

productive discussions than groups in the two Same-strategy conditions, $t (63) = 1.54$, $p =$

.06 (one-tailed), $r = .19$, in support of Hypothesis 5A. A post hoc comparison using the

Tukey HSD test revealed that Same-strategy/Use-Three groups reported significantly

more productive discussions ($M = 8.57$, $SD = 1.29$) than Same-strategy/Add-Two groups

($M = 7.42$, $SD = 1.74$, $p < .05$). The latter difference was not predicted in advance. Thus,

Hypothesis 5A was only weakly supported by the marginal results of the planned

contrast.

Hypothesis 5B was tested with 3 one-way ANOVAs, with items 2, 3, and 4 from

the questionnaire (How confident are you that your team performed better than

25%/50%/75% of teams of students in introductory psychology classes who might sign up for this study?") serving as dependent measures for perceived performance. The effect of strategy diversity condition on self-perceived performance (better than 25% of other participants) was not significant, $F(2, 63) = .08$, $p = .923$, partial $\eta^2 = .00$, suggesting that self-perceived performance did not differ between participants in the Diverse-strategy ($M = 7.36$, $SD = 1.49$), Same-strategy/Add-Two ($M = 7.18$, $SD = 1.75$), and Same-strategy/Use-Three ($M = 7.34$, $SD = 1.68$) conditions. The effect of strategy diversity condition on self-perceived performance (better than 50% of other participants) was not significant, $F(2, 63) = .27$, $p = .763$, partial $\eta^2 = .01$, suggesting that self-perceived performance did not differ between participants in the Diverse-strategy ($M = 6.50$, $SD = 1.70$), Same-strategy/Add-Two ($M = 6.27$, $SD = 1.77$), and Same-strategy/Use-Three ($M = 6.66$, $SD = 1.78$) conditions. The effect of strategy diversity condition on self-perceived performance (better than 75% of other participants) was not significant, $F(2, 63) = .48$, $p = .482$, partial $\eta^2 = .02$, suggesting that self-perceived performance did not differ between participants in the Diverse-strategy ($M = 5.57$, $SD = 1.92$), Same-strategy/Add-Two ($M = 5.14$, $SD = 1.76$), and Same-strategy/Use-Three ($M = 5.82$, $SD = 1.97$) conditions.

An additional exploratory analysis was conducted to test Hypothesis 5B, combining the three confidence measures into a single repeated-measures factor. A mixed-model ANOVA was conducted, with the combined confidence measure as a repeated-measures factor and strategy diversity condition as the between-subjects factor. The main effect of confidence ratings was significant, $F(2, 126) = 111.87$, $p < .001$, partial $\eta^2 = .64$. Not surprisingly, this suggests that confidence in besting these standards differed between ratings. The main effect of strategy diversity condition was not

significant, $F (2, 63) = .34$, $p = .709$, partial $\eta^2 = .01$. This suggests that confidence

ratings did not differ between participants in the Diverse-strategy ($M = 6.48$, $SD = 1.70$),

Same-strategy/Add-Two ($M = 6.20$, $SD = 1.76$), and Same-strategy/Use-Three ($M = 6.61$,

$SD = 1.81$) conditions. The interaction between confidence ratings and strategy diversity

condition was also not significant, $F (4, 126) = 2.59$, $p = .820$, partial $\eta^2 = .03$, suggesting

that the relationship between strategy diversity condition and level of confidence in

besting each standard did not differ between ratings. Greenhouse-Geisser corrections

were employed in this analysis due to the assumption of sphericity being violated

(Mauchly's $W = .452$, $p < .001$). Thus, Hypothesis 5B was not supported.

## Hypothesis 6

I predicted that Diverse-strategy dyads would be rated by observers as having

more productive discussions than Same-strategy dyads. This hypothesis was tested with

several one-way ANOVAs, with strategy diversity condition as the between-subjects

factor. Dependent measures included coded behaviors (Appendix G) as well as team

climate ratings for relationship (conflict vs. cooperation) and discussion activity (active

vs. passive). One coded behavior, Metastrategy statements, was not observed in any

sessions, and thus was omitted from the analyses. Results of these analyses are listed in

Table 3.

For Solutions to Letters statements, a planned contrast revealed that Diverse-

strategy groups made significantly fewer statements than Same-strategy groups, $t (63) = -$

$2.00$, $p < .05$ (one-tailed), $r = .24$. A post hoc comparison using the Tukey HSD test

revealed, somewhat surprisingly, that Same-strategy/Use-Three groups may have made

more Solutions to Letters statements ($M = 14.32$, $SD = 4.98$) than Diverse-strategy

groups ($M = 11.14$, $SD = 4.48$, $p = .053$). For Strategy Sharing (asking partner to share)

statements, a planned contrast revealed that there was no difference between Diverse-

strategy groups and Same-strategy groups, $t(63) = -.53$, $p = .300$ (one-tailed), $r = .00$. A

post hoc comparison using the Games-Howell test (chosen because of unequal variances)

revealed that participants in Same-strategy/Use-Three groups may have asked their

partners to share strategies more ($M = .27$, $SD = .55$) than participants in Same-

strategy/Add-Two groups ($M = .00$, $SD = .00$, $p = .074$). Thus, observed behaviors did

not support Hypothesis 6.

Table 3. Results of ANOVA tests for observed group behaviors (Hypothesis 6). One asterisk indicates marginal significance ($p < .10$), two asterisks indicate significance at $p < .05$. Mean/SD columns include means and standard deviations (in parentheses).

| Behavior | $F$ (2, 63) | $p$ | Partial $\eta^2$ | Mean/SD (Diverse) | Mean/SD (Same/Add-Two) | Mean/SD (Same/Use-Three) |
|---|---|---|---|---|---|---|
| Explicit Strategy | 1.42 | .250 | .04 | 10.68 (3.36) | 8.96 (4.07) | 10.59 (4.03) |
| Strategy Arguments | .52 | .595 | .02 | 1.05 (1.40) | .68 (1.17) | .96 (1.09) |
| Solution to Letters | 2.83 | .066* | .08 | 11.14 (4.48) | 12.59 (3.78) | 14.32 (4.98) |
| Strategy Sharing (shallow) | .85 | .433 | .03 | .32 (.57) | .18 (.40) | .14 (.47) |
| Strategy Sharing (deep) | 1.71 | .190 | .05 | .27 (.70) | .18 (.50) | .55 (.80) |
| Strategy Sharing (asking partner to share) | 3.27 | .045** | .09 | .09 (.29) | .00 (.00) | .27 (.55) |
| Performance Sharing (shallow) | .32 | .728 | .01 | .23 (.75) | .18 (.59) | .09 (.29) |
| Performance Sharing (deep) | 1.55 | .220 | .05 | .00 (.00) | .09 (.29) | .23 (.69) |
| Performance Sharing (asking partner to share) | 1.03 | .362 | .032 | .09 (.29) | .00 (.00) | .05 (.21) |

For team climate ratings, the effect of strategy diversity condition on relationship (conflict vs. cooperation) was significant, $F(2, 63) = 3.46$, $p < .05$, partial $\eta^2 = .10$, suggesting that relationship differed between participants in the Diverse-strategy ($M = 9.32$, $SD = .57$), Same-strategy/Add-Two ($M = 8.82$, $SD = 1.25$) and Same-strategy/Use-Three ($M = 9.45$, $SD = .51$) conditions. A planned contrast revealed that Diverse-strategy groups did not report more conflict or cooperation than Same-strategy groups, $t(63) = .82$, $p = .207$ (one-tailed), $r = .10$. A post hoc comparison using the Tukey HSD test revealed that Same-strategy/Use-Three groups reported significantly more cooperation (less conflict) ($M = 9.45$, $SD = .51$) than Same-strategy/Add-Two groups ($M = 8.82$, $SD = 1.25$, $p < .05$). The effect of strategy diversity condition on discussion activity (active vs. passive) was also significant, $F(2, 63) = 4.69$, $p < .05$, partial $\eta^2 = .13$, suggesting that discussion activity differed between participants in the Diverse-strategy ($M = 8.36$, $SD = 1.59$), Same-strategy/Add-Two ($M = 7.55$, $SD = 2.41$) and Same-strategy/Use-Three ($M = 9.14$, $SD = .77$) conditions. A planned contrast revealed that Diverse-strategy groups did not report more or less active discussions than Same-strategy groups, $t(63) = .05$, $p = .480$ (one-tailed), $r = .01$. A post hoc comparison using the Games-Howell test (again chosen because of unequal variances) revealed that Same-strategy/Use-Three groups reported significantly more active discussions ($M = 9.14$, $SD = .77$) than Same-strategy/Add-Two groups ($M = 7.55$, $SD = 2.41$, $p < .05$). Thus, team climate ratings did not support Hypothesis 6.

## Study 1 Discussion

The purpose of Study 1 was to investigate the role of diversity in group problem-solving process and performance. Participants attempted to solve two letters-to-numbers

problems, the first as individuals and the second as members of a two-person team. For the first problem, participants were given a strategy to solve the problem and were required to use it. Participants were assigned to one of three strategy diversity conditions; in the two Same-strategy conditions the teammates' strategies matched, while in the third Diverse-strategy condition they differed. For the second problem, participants were free to discuss their strategies with their partners and attempt to solve the second letters-to-numbers problem as a team using whichever strategy or combination of strategies they desired, with no restrictions on the strategies that could be used. An additional strategy was offered for the team to consider. It was expected that Diverse-strategy groups would outperform Same-strategy groups on the task due to improved information processing and use of more efficient strategies, prompted by discussion driven by the diversity of opinions within the groups.

Hypothesis 1 predicted that Diverse-strategy groups would perform better than Same-strategy groups, with performance measured by the average number of letters in the code solved per trials used. Results did not support this prediction, as there was no evidence of a main effect of strategy diversity condition, nor an interaction between strategy diversity condition and letters-to-numbers problem. These results suggest that the strategy diversity manipulations may not have affected performance as expected.

Hypothesis 2 predicted that Diverse-strategy groups would demonstrate superior information processing (proposing more complex equations during the task and chancing more exploratory guesses) than Same-strategy groups. For equation complexity, the between-conditions contrast was only marginally significant, but in the predicted direction. For guessing strategy, the between-conditions contrast supported predictions,

partially supporting Hypothesis 2. Thus, Diverse-strategy groups proposed more exploratory guesses, but there was only marginal evidence that they used more complex equations, compared to Same-strategy groups. These results may suggest that interventions designed to influence group information processing may not affect all aspects of that processing equally.

Hypothesis 3 predicted that Diverse-strategy groups would be more likely than Same-strategy groups to use outside advice in the form of a more complex but potentially more useful strategy for the second problem. Results suggested that there was no difference in the use of this outside advice between strategy diversity conditions, and so Hypothesis 3 was not supported. It is possible that intragroup diversity does not impact a group's consideration of outside advice at all, or that there are other factors that might affect a problem-solver's consideration or use of outside advice. One such factor, confidence in one's own strategy, is explored further in Study 2 with individuals. Hypothesis 4 predicted that the tendency to use the more complex Add All Unknown strategy would mediate the relationship between strategy diversity condition and letters-to-numbers performance, with Diverse-strategy groups being more likely to use the Add All Unknown strategy. Preliminary conditions necessary to test mediation were not met, and so Hypothesis 4 was not supported.

In addition to performance and information processing behaviors, Study 1 also examined group activity through the use of self-report questionnaires completed by participants and observer ratings of team dynamics. Hypothesis 5 predicted that participants in Diverse-strategy groups would report more productive group discussions and better self-perceived performance than participants in Same-strategy groups. Results

provided only marginal evidence that Diverse-strategy groups may have had more productive discussions than Same-strategy groups. There was no evidence that self-perceived performance differed between strategy diversity conditions. Since other analyses of group discussion and performance also did not suggest predicted differences between conditions, these results may suggest that participants' perceptions of their own group's dynamics and performance were simply accurate.

Observers analyzed video-recordings of group discussions to examine productivity of discussions and rated general team climate. Out of 9 categories of behaviors, only two showed evidence of differences (one only marginally) between strategy diversity conditions, and the pattern of results did not support predictions in line with Hypothesis 6. For team climate ratings, there was no evidence that Diverse-strategy groups differed from Same-strategy groups in conflict (vs. cooperation) or general activity levels of discussion. It is possible that any behavioral differences between conditions were not captured by the behavioral coding and team climate rating measures, both of which were relatively new measures designed by the principal investigator.

CHAPTER THREE

STUDY 2

Study 2 explored the influence of confidence when considering outside advice at the individual level. Confidence was considered as a potential moderator or mediator of effects observed in previous advice-taking studies; instead of factors increasing/decreasing a decision-maker's tendency to consider or use outside advice directly, such factors may increase/decrease the decision-maker's confidence in his/her judgment or ability, and this in turn may drive them to seek advice. As in Study 1, participants attempted to solve two letters-to-numbers problems. However, they attempted to solve both problems as individuals (never in teams with other participants, unlike Study 1). Participants were randomly assigned to either high-confidence or low-confidence conditions. Confidence in one's assigned strategy was manipulated by altering the letters-to-numbers program to provide more (high-confidence) or less (low-confidence) useful answers to equations that participants submitted to the program as they attempted to solve the code. The following predictions were tested:

Hypothesis 7: Low-confidence individuals will outperform high-confidence individuals on the second letters-to-numbers problem, with performance measured in the same way as in Study 1: Letters solved per trial used, and using fewer trials to solve the code overall.

Hypothesis 8: Low-confidence individuals will be more likely than high-

confidence individuals to use outside advice (in the form of the more complex but potentially more efficient Add All Unknown strategy).

Hypothesis 9A: The tendency to use the more complex Add All Unknown strategy will mediate the relationship between self-reported confidence and letters-to-numbers performance.

Hypothesis 9B: Self-reported confidence will mediate the relationship between strategy confidence condition and letters-to-numbers performance.

Hypothesis 9C: Both self-reported confidence and the tendency to use the more complex Add All Unknown strategy will mediate the relationship between strategy confidence condition and letters-to-numbers performance, with lower confidence predicting the tendency to use the Add All Unknown strategy. See Figure 1 for the predicted path analysis model.

Figure 1. Mediation path analysis model for Hypothesis 9C. Solid lines indicate predicted significant relationships.
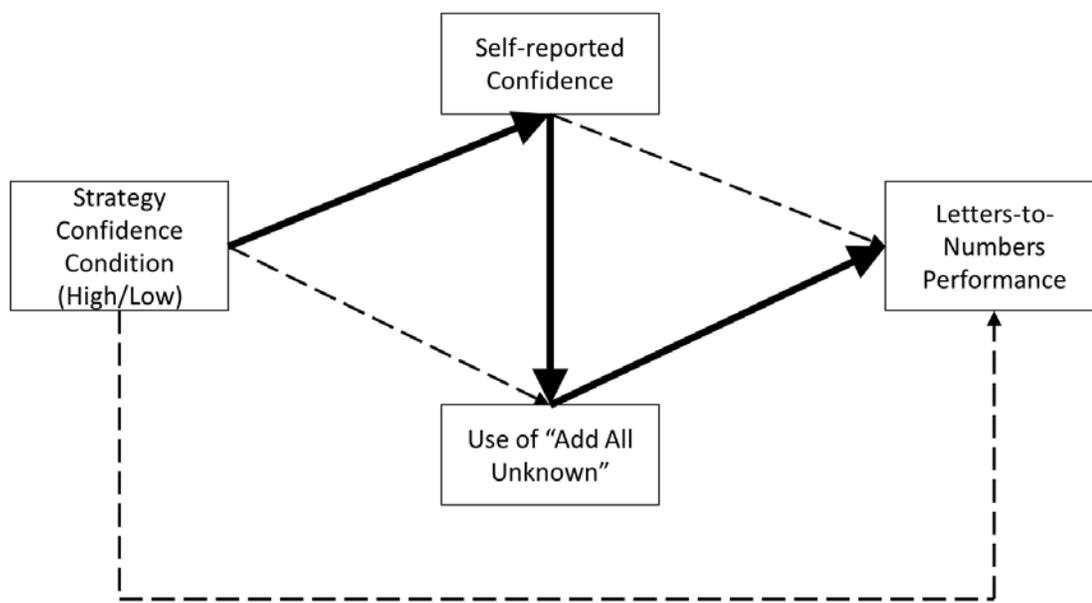
**Study 2 Method**

Participants

Sixty participants were recruited for Study 2, based on the same criteria for sample size calculation as Study 1 (80% power with a small to moderate effect size). Undergraduate students from introductory psychology classes were recruited for the study. Participants were randomly assigned to one of two strategy confidence conditions and attempted to solve two letters-to-numbers problems. For their participation, students earned credits toward fulfilling a research participation course requirement. Participants also had the opportunity to enter the same raffle as Study 1 participants.

Experimental Task

Participants in Study 2 attempted to solve two letters-to-numbers problems with the same general structure as in Study 1.

Experimental Conditions

Instead of manipulating equation strategies, all participants were instructed to use the Add Two strategy as described in Study 1. Participants were assigned to one of two strategy confidence conditions (high or low). For participants assigned to the high-confidence condition, the letters-to-numbers program was manipulated to always produce a two-digit answer to the first equation that participants submitted to the program, as in Study 1. For low-confidence participants, the program delayed (as long as was possible with the unused numbers available) the occurrence of a two-digit answer. The latter was intended to make the problem more difficult and cause participants to question the effectiveness of their problem-solving strategy.

Procedure

With the exception of the experimental manipulation, procedures for Study 2 were similar to those for Study 1. Participants were administered the same informed consent process, compensated with both research participation credits and a chance to win a raffle prize (if desired), completed the tutorial and two letters-to-numbers problems (being forced to use the Add Two strategy on the first problem, but having the choice to use the Add All Unknown strategy on the second problem), and answered a questionnaire after each problem (see Appendices D and F). Participants in Study 2 were the only participant in the laboratory room for the experiment session.

Measures

**Letters-to-Numbers Process and Performance**

The measures for participants' progress through the letters-to-numbers problems and their performance were the same as in Study 1.

**Confidence**

Confidence was also measured the same way as in Study 1, although both questionnaires asked participants to assess themselves as individuals (instead of as a team for the second problem as in Study 1).

**Study 2 Results**

Hypothesis 7

I predicted that individuals in the low-confidence condition would outperform individuals in the high-confidence condition on the second letters-to-numbers problem. Again, performance was poorer than expected. 50% of participants failed to solve the first problem, and 25% failed to solve the second problem. The average number of letters

solved per trials used again served as a performance measure (for the same reasons given

in Study 1). This hypothesis was tested with a one-way ANOVA, with strategy

confidence condition as a between-subjects factor. The effect of strategy confidence

condition on performance was marginally significant, $F (1, 58) = 2.96$, $p = .091$, partial $\eta^2$

$= .05$. This suggests that participants in the low-confidence condition ($M = 1.31$, $SD =$

.54) may have performed better than participants in the high-confidence condition ($M =$

1.09, $SD = .45$). Thus, Hypothesis 7 was only weakly supported by this marginal result.

## Hypothesis 8

I predicted that individuals in the low-confidence condition would be more likely

than individuals in the high-confidence condition to use outside advice (in the form of the

more complex but potentially more efficient Add All Unknown strategy) on the second

problem. The tendency to use this strategy was measured as the proportion of trials on

which the strategy was used. This hypothesis was tested with a one-way ANOVA, with

strategy confidence condition as the between-subjects factor. The effect of strategy

confidence condition on the tendency to use this strategy was significant, $F (1, 58) =$

7.57, $p < .01$, partial $\eta^2 = .12$. This suggests that participants in the low-confidence

condition ($M = .15$, $SD = .25$) used this strategy more than participants in the high-

confidence condition ($M = .02$, $SD = .05$). Thus, Hypothesis 8 was supported.

## Hypothesis 9

I predicted that the tendency to use the more complex Add All Unknown strategy

would mediate the relationship between self-reported confidence and letters-to-numbers

performance (Hypothesis 9A), with lower self-reported confidence correlating with

increased likelihood of using the Add All Unknown strategy, and increased use of this

strategy correlating with better performance. Confidence was measured with items 2, 3, and 4 from the P1-Post questionnaire ("How confident are you that you performed better than 25%/50%/75% of all of the students in introductory psychology classes who might sign up for this study?").

I conducted three one-way ANOVAs as a manipulation check for the three confidence measures, with strategy confidence condition as the between-subjects factor. For the "better than 25%" measure, the difference between strategy confidence conditions was not significant, $F (1, 58) = 1.13$, $p = .292$, partial $\eta^2 = .02$. This suggests that the self-reported confidence for participants in the low-confidence condition ($M = 5.37$, $SD = 2.30$) and those in the high-confidence condition ($M = 4.70$, $SD = 2.55$) did not differ. For the "better than 50%" measure, the difference between strategy confidence conditions was not significant, $F (1, 58) = 2.20$, $p = .143$, partial $\eta^2 = .04$. This suggests that the self-reported confidence for participants in the low-confidence condition ($M = 4.60$, $SD = 2.18$) and those in the high-confidence condition ($M = 3.77$, $SD = 2.18$) did not differ. For the "better than 75%" measure, the difference between strategy confidence conditions was not significant, $F (1, 58) = .69$, $p = .411$, partial $\eta^2 = .01$. This suggests that the self-reported confidence for participants in the low-confidence condition ($M = 3.73$, $SD = 2.21$) and those in the high-confidence condition ($M = 3.27$, $SD = 2.15$) did not differ. The pattern of means (with low-confidence participants reporting higher—though not statistically significantly—self-reported confidence) was unexpected.

A path analysis approach would have been used to test relationships between strategy confidence condition, self-reported confidence, the tendency to use the Add All Unknown strategy, and performance based on the mediation approach discussed by

Baron & Kenny (1986). Multiple variations of this model were considered, using the

three different confidence measures separately, as well as analyzing data from

participants in the two strategy confidence conditions combined in one sample and as two

separate samples. Despite this varied approach, the data did not support the proposed

mediation model(s). Preliminary conditions for mediation (testing correlations between

variables in the model) did not provide the necessary circumstances for conducting a

mediation analysis. In order to test the prediction that the tendency to use the Add All

Unknown strategy mediates the relationship between self-reported confidence and

performance (Hypothesis 9A), three preliminary conditions must hold: 1) self-reported

confidence and performance are correlated, 2) self-reported confidence and the tendency

to use the Add All Unknown strategy are correlated, 3) the tendency to use the Add All

Unknown strategy and performance are correlated, while controlling for the relationship

between self-reported confidence and performance. In order to test the prediction that

self-reported confidence mediates the relationship between strategy confidence condition

and performance (Hypothesis 9B), similar conditions would need to hold: 1) strategy

confidence condition and performance are correlated, 2) strategy confidence condition

and self-reported confidence are correlated, 3) self-reported confidence and performance

are correlated, while controlling for the relationship between strategy confidence

condition and performance. Correlation matrices used to test these steps are presented in

Tables 4 and 5. Table 4 presents correlations between variables for all participants in both

strategy confidence conditions. Table 5 presents correlations between variables for high-

confidence (above diagonal) and low-confidence (below diagonal) participants

separately. As can be seen from the data, the necessary correlational relationships

between variables are not present, and resulted in non-significant path coefficients in the preliminary mediation analyses. Thus, Hypotheses 9A and 9B were not supported. Hypothesis 9C assumed that Hypotheses 9A and 9B would be supported, and so it was also not supported. Overall, Hypothesis 9 was not supported.

An additional path analysis model was considered to explore an alternative mediation model. This model would test the hypothesis that the tendency to use the more complex Add All Unknown strategy mediates the relationship between strategy confidence condition (replacing self-reported confidence in the previous analyses) and letters-to-numbers performance, with participants in the low-confidence condition being more likely to use the Add All Unknown strategy than participants in the high-confidence condition, and increased use of this strategy correlating with better performance. Correlations between variables used to test this model are included in Table 4.

Although the direct relationship between strategy confidence condition and performance was not statistically significant, a bootstrapping approach was used to test the possible indirect effect of strategy diversity condition on performance through the use of the Add All Unknown strategy, following procedures suggested by Preacher & Hayes (2004). The analysis provides results for the bootstrapping test as well as the Sobel test, an alternative approach to testing indirect mediation effects (Sobel, 1982). The bootstrapping test, based on 5000 resamples, suggested that the indirect effect was not significant; the 95% Confidence Interval included zero: [-.22, .04]. Increasing the number of resamples did not significantly alter results. The results of the Sobel test also suggested that the indirect effect was not significant, $z = -1.56$, $p = .119$. Thus, the alternative mediation model was not supported.

Table 4. Correlations between variables for all participants used to test mediation models. Right-hand column provides standard deviations for all variables. Two asterisks indicate significant correlations at $p < .01$. One asterisk indicates significant correlations at $p < .05$.

| | Strategy Confidence Condition (Low/High) | Confidence (50%) | Confidence (25%) | Confidence (75%) | Proportion of Trials used Add All Unknown | SD |
|---|---|---|---|---|---|---|
| **Strategy Confidence Condition (Low/High)** | | | | | | .504 |
| **Confidence (50%)** | -.191 | | | | | 2.198 |
| **Confidence (25%)** | -.138 | .898** | | | | 2.428 |
| **Confidence (75%)** | -.108 | .888** | .668** | | | 2.175 |
| **Proportion of Trials used Add All Unknown** | -.340** | -.062 | -.033 | -.059 | | .193 |
| **Performance** | -.220 | .273* | .299* | .181 | .312* | .502 |

Table 5. Correlations between variables used to test mediation models, with correlations for high-confidence (above diagonal) and low-confidence (below diagonal) participants presented separately. The two right-most columns provide standard deviations for all variables in each group. Two asterisks indicate significant correlations at $p < .01$. One asterisk indicates significant correlations at $p < .05$.

| | Confidence (50%) | Confidence (25%) | Confidence (75%) | Proportion of Trials used All Unknown | Performance | SD (Low Conf.) | SD (High Conf.) |
|---|---|---|---|---|---|---|---|
| **Confidence (50%)** | | .895** | .899** | -.226 | .404* | 2.175 | 2.176 |
| **Confidence (25%)** | .900** | | .670** | -.360 | .419* | 2.297 | 2.548 |
| **Confidence (75%)** | .880** | .658** | | -.033 | .295 | 2.212 | 2.149 |
| **Proportion of Trials used All Unknown** | -.151 | -.046 | -.139 | | -.181 | .253 | .054 |
| **Performance** | .108 | .154 | .055 | .375* | | .537 | .446 |

**Study 2 Discussion**

Study 2 further explored the role of confidence in an individual problem-solver's tendency to consider outside advice. Individual participants attempted to solve two letters-to-numbers problems. Similar to Study 1, participants were required to use a specific equation strategy for solving the first problem, and were offered the alternative Add All Unknown strategy to consider using when attempting to solve the second problem, with no restrictions on equations that could be used for the second problem. Unlike Study 1, all participants were given the same strategy to use on the first problem (the Add Two strategy). Participants were randomly assigned to one of two strategy confidence conditions, intended to promote high or low confidence in the strategy used for the first problem as an effective strategy to use again for the second problem. It was predicted that low-confidence participants would perform better on the second problem (solving for more letters per trials used), and that this may result from an increased tendency to consider and use the outside advice, with this use of outside advice mediating the relationship between self-reported strategy confidence and performance. Results provided marginal evidence that low-confidence participants may have performed better than high-confidence participants on the second problem, possibly supporting Hypothesis 7. Results also suggested that low-confidence participants were more likely than high-confidence participants to consider and use the outside advice, supporting Hypothesis 8. However, the predicted mediation model was not supported by the data, contradicting Hypothesis 9. Preliminary conditions necessary to test mediation effects were not met, despite various exploratory approaches to conducting this analysis. An alternative mediation model was considered, proposing that the use of the outside advice would

mediate the relationship between strategy confidence condition (instead of self-reported

confidence) and performance. However, this model was also not supported.

CHAPTER FOUR

GENERAL DISCUSSION

The purpose of these studies was to investigate factors that may influence the extent to which both groups and individuals will consider or use outside advice when attempting to solve a problem. A long history of research within psychology and other disciplines supports evolutionary perspectives that, throughout our evolutionary history, we tend to affiliate with others and seek their aid when confronting problems. This is characteristic of modern humans as well, with recent studies in psychology and organizational behavior demonstrating our preference to work with others over facing our problems alone. However, there appear to be circumstances in which we stubbornly shun the help or advice of others. The present studies investigated some factors that may contribute to this behavior.

Study 1 focused on the behavior of collaborative problem-solving groups (two-person dyads), and the role of intragroup diversity in groups' tendency to consider or use advice from a source outside the group. Diversity was presented in the form of group members' experience with one of two particular strategies to use to solve the problem that the group was tasked with solving. An experimental manipulation was introduced to give group members experience with either the same or different strategies. When paired with partners to perform the task as a team, these Same-strategy or Diverse-strategy groups were also given outside advice (a third strategy) to consider. Groups were

free to discuss their own strategies and the given advice while attempting to solve the problem. Previous research suggests that diversity within groups may prompt conflict and subsequent discussion to resolve that conflict, potentially leading to innovation and improved performance (Carton & Cummings, 2012; Franz, 2012). In contrast, relatively homogeneous groups may never experience such conflict, leading to more cooperative but ultimately suboptimal performance as these groups prioritize consensus over potentially disruptive innovation or deliberation. Following this logic, it was expected in Study 1 that Diverse-strategy groups, compared to Same-strategy groups, would be more likely to consult and utilize outside advice to resolve their inter-member differences, demonstrate superior information processing, have more productive group discussions, and perform better on the task. Results generally did not support these predictions; a few possible explanations are considered.

First, it is possible that the predicted relationship between intragroup diversity and group performance does not manifest when performance is constrained to intellective tasks like the letters-to-numbers problem. To the author's knowledge, this is the first study examining the relationships between intragroup diversity, advice-taking, and performance on an intellective task. Intellective tasks are notable for having demonstrably correct solutions, and a series of logical intermediate steps to arrive at those solutions. When groups perform such tasks, a correct solution or efficient method for performing the task successfully can be demonstrated by an individual to their group members (Laughlin, 2011). I expected that such tasks might be particularly useful for studying various aspects of group process and performance. While performing such tasks, participants could presumably demonstrate ideas to fellow group members more easily

than on tasks with no definitive solution or ambiguous processes to arrive at that solution.

If group members differ in their solution strategies and approaches to the task, and an

outside source of information was available, group members may be able to take

advantage of the demonstrable nature of the task to argue in favor of a specific strategy

(that is also demonstrable), and such discussions might drive more innovative solutions

than if there was no discussion. This perspective may be misguided; the demonstrable

nature of the task may not impact group discussions or individual perspectives on

consulting outside advice. Alternatively, the letters-to-numbers task, at least as it was

presented in this study, may be less demonstrable than previously thought. The

demonstrability of a group task depends in part on group members' abilities to

understand, perform, and demonstrate the task to their fellow group members (Larson,

2010, Laughlin, 2011; Laughlin & Ellis, 1986). Poor performance on the task overall in

the studies presented here may suggest that these criteria were not met. The predicted

effects may manifest with other intellective tasks with a higher degree of demonstrability.

Future research should continue probing how characteristics of a task might influence

participants' behavior while performing it, beyond hypotheses related to more fine-

grained effects.

Second, there may be mechanical limitations of the letters-to-numbers task itself.

Earlier I discussed the issue of measuring performance for participants that fail to solve

the letters-to-numbers code(s) and the use of a substitute performance measure (average

numbers of letters solved per trials used, instead of the more traditional measure of letters

used to solve the code in its entirety). It is unclear if this is an appropriate alternative

measure. If data from previous letters-to-numbers studies were to be analyzed with this

alternative performance measure as a dependent variable, the results may not support the same conclusions. Further, previous studies did not report failure rates of participants (the number of participants, if any, who failed to solve similar letters-to-numbers codes) in those studies. It is not clear if the proportion of participants who failed to solve one or both letters-to-numbers problems in the studies presented here was unique to these studies; the results of previous letters-to-numbers studies—specifically, mean performance scores measured as trials used to solve the code—suggest that failure rates for those studies were not as high as in the two studies presented here. Considering that "successful" performance of participants on a task can never be guaranteed, future problem-solving researchers are strongly encouraged to consider the tasks used to measure performance and the performance measures specific to a given task, to avoid similar difficulties in interpreting performance data. In some cases a specific task may not be appropriate for a variety of reasons, or some adjustments to its administration may be required. If such adjustments are made, researchers should be cautious when interpreting findings in comparison to results from previous studies using a variation of the same task. Other variations of the letters-to-numbers task to resolve these difficulties, or the use of an alternative task to measure individual and group performance, may be beneficial for future research endeavors.

Third, experimental manipulations of intragroup diversity may not have been sufficiently robust to produce the intended effects on advice-taking, group behaviors, or performance. Previous theories that consider the influence of diversity on group process and performance suggest a curvilinear relationship, with moderate diversity being optimal for groups to reap the benefits (Carton & Cummings, 2012). If there is too little

diversity, the group may not innovate nor engage in effortful discussion. If there is too much diversity, irreconcilable conflict may stifle group progress. Of these two extremes, the former seems more likely in Study 1. The observational analyst (viewing recordings of group discussions) noted relatively low levels of discussion overall within groups, and highly cooperative (vs. conflictual) behavior. Regardless of strategy diversity condition, groups also rarely consulted, discussed, or used the outside advice. Although individual participants were temporarily required to adhere to a specific strategy, they might not have internalized this strategy as their own, nor felt motivated to defend it as a good strategy in group discussion. Thus, any predicted effects of diversity may not have been observed simply because the diversity itself was not truly present. Future studies on intragroup diversity should consider alternative methods of manipulating or creating diversity within groups, or perhaps strategic recruitment of participants to take advantage of existing diversity within participant samples, to create stronger perceptions of differences between group members.

Potential difficulties in operationalizing experimental conditions and limitations of the letters-to-numbers task may have extended to Study 2, conducted with individual participants. The goal of Study 2 was to examine how an individual problem-solver's confidence in his/her own problem-solving strategy might affect his/her tendency to consider an alternative strategy presented in the form of outside advice. It was predicted that individuals induced to have low confidence in their given strategy on the first problem would be more likely to consider outside advice and perform better on the second problem, compared to those induced to have higher confidence in their given strategy on the first problem. Low-confidence individuals were expected to be frustrated

or disappointed with their apparent ineffective strategy, and to defect to the strategy offered by the outside advice, leading to better performance on the second problem compared to high-confidence individuals who might have remained loyal to their given strategy.

Results of Study 2 did support some predictions. Low-confidence individuals did use the strategy offered by the outside advice more than high-confidence individuals, although participants in both conditions used the alternative strategy offered by the advice less than was expected. Analysis of performance yielded only marginally significant statistical differences between conditions, providing inconclusive evidence of a relationship between strategy confidence condition and performance. I also predicted a mediation effect, with the use of outside advice mediating the relationship between self-reported strategy confidence and performance. The proposed mediation model could not be fully tested since preliminary conditions for mediation were not met. An alternative mediation model proposing that the use of outside advice would mediate the relationship between strategy confidence condition (instead of self-reported strategy confidence) and performance was also not supported. The proposed mediation effects may not operate as expected, or the measures used may not have accurately represented the constructs underlying the proposed models.

Similar to Study 1, experimental manipulations in Study 2 may not have been robust enough to produce desired effects. As participants were working to solve the first letters-to-numbers problem, the computer program used to administer the task dynamically adjusted difficulty, making the problem more difficult for low-confidence participants and easier for high-confidence participants as they attempted to solve it. The

difference between the "difficult" and "easy" versions of the task may not have been dramatic enough to affect perceptions of participants as intended, resulting in confidence levels between low-confidence and high-confidence participants not being different enough in magnitude to influence consideration of outside advice in the predicted manner. Presenting participants with more than one letters-to-numbers problem to solve using a specific strategy might strengthen such a manipulation. Alternative methods of manipulating confidence might also be developed, although ethical considerations may prevent researchers from inducing radical changes in participants' confidence in their abilities or other thoughts or behaviors. Overly strong manipulations may also be too obvious to participants, and lead to biased behavior due to experimental demand. Future research should continue exploring how to best achieve balance between these concepts.

Also similar to Study 1, some participants in Study 2 failed to solve one or both of the letters-to-numbers problems. If a task is intended for groups of participants to complete collaboratively, individuals may be even less likely than groups (with any number of members) to be able to solve it. Researchers that seek to compare performance of individuals and groups, or at least test performance of individuals and groups on the same tasks, should consider whether the tasks can be used for both. The letters-to-numbers problems may simply have been too difficult and artificially reduced variance in performance as a result. Attempts were made to make the problems easier to solve with certain experimental manipulations, but these may not have been sufficient. Further efforts to make the problem easier (e.g., if participants start the problem knowing which letter represents the number 1) may resolve such issues, however it is also possible that the problem may be made too easy as a result.

Another issue potentially affecting both studies was how confidence was measured. After completing each letters-to-numbers problem, participants responded to questionnaire items asking them to rate their perceived performance against 25%, 50%, and 75% of other participants that might participate in the study. Hypotheses were tested with each of these measures separately, as well as with a measure that combined the three confidence measures into a single repeated-measures factor. These may not have been satisfactory measures for a confidence variable. Since confidence is such a widely studied construct with many ways of defining and measuring it (Forsyth, 1999; Kruglanski & Webster, 1996; Park & Hinsz, 2006; Tindale, 1989), there are no clear better alternatives than what was presented here. Future efforts into evaluating the validity and reliability of confidence measures, as well as more nuanced features (e.g., balancing the directness of a measure with the need to avoid revealing purposes of the study) may benefit selection of such measures for future studies. The measures used in these studies were exploratory and were not rigorously tested for such criteria; the use of instruments or measures with a more proven history might have strengthened the studies.

The concerns described thus far are primarily methodological in nature. However, the mixed results of both studies might have also resulted from flawed theoretical predictions. It seems premature to dismiss the possibility of relationships between intragroup diversity and advice-taking, confidence and advice-taking, or how either of these may impact individual or group performance given the significant findings in this and other studies on advice-taking by individuals and groups. Such effects may simply not manifest under the specific circumstances of the present studies. It is possible that relationships between the main variables of interest are more complex than previously

thought, and that additional intermediary variables may drive effects observed in other studies. For example, previous studies have considered how an advisor's apparent expertise might affect a decision-maker's tendency to consider their advice (Sniezek & Van Swol, 2001). Inclusion of such additional variables may allow for a more fine-tuned understanding of why we only sometimes consult an advisor when attempting to make a decision or solve a problem.

It is also hoped that these studies contribute to the already expansive research literature on both group diversity and confidence. Diversity and its impact on interpersonal relationships continues to be a popular avenue of research in social psychology and organizational behavior, as well as in other disciplines. Continued interest in this topic (and the persistence of contradictory findings) speaks to the complexity of the topic. Researchers should continue to explore new ways to define and measure diversity to allow investigation into its effects (both positive and negative) on various aspects of group life.

Confidence is also a persistently popular topic for research inquiry. The studies reported here examined experimentally manipulated confidence specific to a certain problem-solving task. Given the vast amount of motivational techniques intended to improve our confidence in ourselves and our abilities, future studies should continue exploring how confidence can be changed or manipulated, as well as how it can be best measured. It may be more beneficial for research studies to consider specific types or subtypes of confidence, such as self-confidence versus confidence in one's group during group tasks, instead of treating it as a broader construct.

The studies presented here produced generally mixed findings, supporting some theoretical predictions while failing to provide support for others. It is unclear whether this is due to flaws in theoretical perspectives, flaws in methodological approach, some combination of both, or simply an accurate picture of reality. Since methodological shortcomings seem more likely to be the cause of failed hypotheses, some adjustments and improvements to methods were offered. Alternative theoretical perspectives were also presented for those wishing to explore these avenues.

APPENDIX A

"ADD TWO LETTERS" STRATEGY SHEET

# The "Add Two Letters" Strategy

Explanation:
Create equations by adding two letters together.

Examples:
A + B =
A + A =
C + J =

*Remember to include the equal sign (=) at the end of the equations.

APPENDIX B

"USE THREE LETTERS" STRATEGY SHEET

# The "Use Three Letters" Strategy

Explanation:
Create equations that include 3 letters, using addition or subtraction.

Examples:
A + B + C =
A + A + A =
B + CD =
BC + D =
C + D – E =
E – F + G =
A – DJ =

*Remember that two letters together (like CD) is a 2-digit number, NOT multiplying C x D.

*Remember to include the equal sign (=) at the end of the equations.

APPENDIX C

"ADD ALL UNKNOWN LETTERS" STRATEGY SHEET

# The "Add All Unknown Letters" Strategy

Explanation:
Create equations by adding together all of the letters that
you don't know.

Examples:

(starting with no letters known):
A + B + C + D + E + F + G + H + I + J =

(if you already know A, B, E, and F):
C + D + G + H + I + J =

*Remember to include the equal sign (=) at the end of the
equations.

APPENDIX D

P1-POST QUESTIONNAIRE

P1-Post Questionnaire
(After First Problem, Both studies)

**The following questions will ask you about the Letters-to-Numbers problem you just completed. Please read the questions carefully and answer them honestly. For each question, circle a number on the answer scale to choose your answer.**

1. How **Easy or Difficult** did you find the problem to be?
*Very Easy – Very Difficult (1-10).*

Your performance on the problems is measured by how many letters in the code you figured out, and how many trials it took you to do it.

2. How **confident** are you that you performed better than **25%** of all of the students in introductory psychology classes who might sign up for this study?
*Not at all confident – Extremely confident (1-10).*

3. How **confident** are you that you performed better than **50%** of all of the students in introductory psychology classes who might sign up for this study?
*Not at all confident – Extremely confident (1-10).*

4. How **confident** are you that you performed better than **75%** of all of the students in introductory psychology classes who might sign up for this study?
*Not at all confident – Extremely confident (1-10).*

APPENDIX E

STUDY 1 P2-POST QUESTIONNAIRE

Study 1 P2-Post Questionnaire
(After second problem, Study 1 only)

**The following questions will ask you about the SECOND Letters-to-Numbers problem you just completed. Please read the questions carefully and answer them honestly. For each question, circle a number on the answer scale to choose your answer. For open-ended questions, please write answers on the lines.**

1. How Easy or Difficult did you find the problem to be?
    *Very Easy – Very Difficult (1-10).*

Your performance on the problems is measured by how many letters in the code you figured out, and how many trials it took you to do it.

2. How **confident** are you that your team performed better than **25%** of teams of students in introductory psychology classes who might sign up for this study?
    *Not at all confident – Extremely confident (1-10).*

3. How **confident** are you that your team performed better than **50%** of teams of students in introductory psychology classes who might sign up for this study?
    *Not at all confident – Extremely confident (1-10).*

4. How **confident** are you that your team performed better than **75%** of teams of students in introductory psychology classes who might sign up for this study?
    *Not at all confident – Extremely confident (1-10).*

5. How much did your team consider or use the strategy given to you for the second problem?
    *Not at all – Very Much (1-10).*

6. Would you consider your team discussion as being more conflict or cooperation?
    *Mostly conflict – Mostly cooperation (1-10).*

7. Would you consider your team discussion as being more awkward or smooth?
    *Mostly awkward – Mostly smooth (1-10).*

8. Did your or your teammate take on more of a leadership role?
    *I lead – Neither – My teammate lead (1-10).*

9. How beneficial was discussing with your teammate in helping your team perform better?

*Not at all beneficial – Very beneficial (1-10).*

10. How well did you know your teammate before today?

*We're total strangers – We know each other very well (1-10).*

Gender: _____
Age: _____
Year in School:     Freshman    Sophomore   Junior      Senior
Major: _____

Do you know (or think you can guess) what the experiment is about?

_____
_____
_____

APPENDIX F

STUDY 2 P2-POST QUESTIONNAIRE

## Study 2 P2-Post Questionnaire
## (After second problem, Study 2 only)

**The following questions will ask you about the Letters-to-Numbers problem you just completed. Please read the questions carefully and answer them honestly. For each question, circle a number on the answer scale to choose your answer. For open-ended questions, please write answers on the lines.**

1. How **Easy or Difficult** did you find the problem to be?
    *Very Easy – Very Difficult (1-10).*

Your performance on the problems is measured by how many letters in the code you figured out, and how many trials it took you to do it.

2. How **confident** are you that you performed better than **25%** of all of the students in introductory psychology classes who might sign up for this study?
    *Not at all confident – Extremely confident (1-10).*

3. How **confident** are you that you performed better than **50%** of all of the students in introductory psychology classes who might sign up for this study?
    *Not at all confident – Extremely confident (1-10).*

4. How **confident** are you that you performed better than **75%** of all of the students in introductory psychology classes who might sign up for this study?
    *Not at all confident – Extremely confident (1-10).*

5. How much did you consider or use the strategy given to you for the second problem?
    *Not at all – Very Much (1-10).*

Gender: _____
Age: _____
Year in School:        Freshman        Sophomore    Junior            Senior
Major: _____

Do you know (or think you can guess) what the experiment is about?

APPENDIX G

VIDEO CODING SCHEME

Video Coding Scheme

- **Explicit strategy – specific thoughts/suggestions about strategies. Items in this category can be directly implemented and are concrete. Includes equation ideas.**

  Examples:
  "We should use more complex equations."
  "Maybe using bigger numbers will be more effective."
  "Should we use longer equations?"
  "We need to solve for more letters at a time."
  "Maybe it is better to guess about letters we don't know."
  "Let's try the 'Add All Unknown' strategy to see if it works."
  "We should do A+C for our next equation."
  "Let's try to figure out what F is."

- **Metastrategy – General thoughts regarding strategy or characteristics of strategy. Items are more abstract than those in the Explicit strategy category, and cannot be directly implemented as concise actions.**

  Examples:
  "Let's evaluate the quality of our equations."
  "There have to be better strategies out there than what we're doing."
  "Is there a way to do this in fewer trials?"

- **Strategy Arguments (supporting/opposing strategy ideas) – Opinions regarding why any** presented ideas are good or bad strategies for improving performance.

  Examples (strategy arguments in bold):
  "We can't use multiplication in our equations **because the task doesn't let us use multiplication**."
  "We should use longer equations **because it helps us solve the code faster**."
  "The 'Add All' strategy **seems too complicated**, let's not use it."

- **Solutions to Letters**

  Examples:
  "I think that A=2 because…"
  "B can't be 4 because…"
  "What if A was 3 and C was 6?"

- **Strategy Sharing – teammates share/discuss the strategy they used to solve the first problem.**
  - Shallow sharing – identifying strategy used

- o Deep sharing – comparing strategies, evaluating effectiveness/utility of each
- o Asking partner to share

- **Performance Sharing – teammates share/discuss their performance on the first problem.**
  - o Shallow sharing – mentioning performance (# of trials, good/bad, etc.)
  - o Deep sharing – reflections on why performance was good/bad, or could be improved
  - o Asking partner to share

- **Time spent discussing before action on second problem.**
  - o "Action" defined as program input. Delay between initiating program and a participant doing something with the computer.

- **Requesting assistance from experimenter.**
  - o Reason and content of request.

# General Team Climate Coding

*For each item in bold, circle the number or description that best describes what was observed.* ***ONLY APPLIES TO TEAMS (not individuals).***

**Energy**:
1    2    3    4    5    6    7    8    9    10
[Low energy]                      [High energy]

**Interaction**:
-5    -4    -3    -2    -1    1    2    3    4    5
[Awkward]                      [Smooth]

**Discussion/deliberation**:
1    2    3    4    5    6    7    8    9    10
[Passive]                      [Active]

**Member contribution**:
-5    -4    -3    -2    -1    1    2    3    4    5
[A contributes more]           [B contributes more]

**Leadership**:
-5    -4    -3    -2    -1    1    2    3    4    5
[A leads more]                [B leads more]

**Relationship**:
-5    -4    -3    -2    -1    1    2    3    4    5
[More conflict]             [More cooperation]

**Performance**:
Succeeded          Failed          Unclear

How to code behaviors:

Energy
        LOW – Team is sluggish or uninterested in task.
        HIGH – Team is energetic and engaged in task.

Interaction

      SMOOTH – Members seem comfortable working together.

      AWKWARD – Members are uncomfortable or nervous working together.

Discussion/deliberation

      ACTIVE – Discussion involves a lot of talking, debating, and "back-and-forth" between members.

      PASSIVE – Discussion involves little talking, members do not challenge each other's ideas, member just goes with partner's ideas.

Member contribution

      EQUAL – Members contribute roughly equally to the conversation and the task.

      UNEQUAL – One member dominates or carries discussion and activity.

Leadership

      A LEADS – Member farthest from camera leads team more.

      B LEADS – Member closest to camera leads team more.

      NONE/NEITHER – Neither member exhibits leadership, or they lead team equally.

Relationship

      CONFLICT – Members seem unable or unwilling to agree, "not on the same page."

      COOPERATION – Members able to reach agreement through discussion (even if they demonstrate some difference of opinion).

Performance

      SUCCEEDED – Team solved the problem.

      FAILED – Team did not solve the problem.

      UNCLEAR – Not clear from video if team solved the problem or not ("finished" or "done" does not always mean "solved.").

Notes – note any other interesting features of video, or any issues with coding system (e.g., if none of the coding options for an item fit what was observed).

REFERENCE LIST

Ashby, R. (1958). Requisite variety and its implications for the control of complex systems. *Cybernetica*, *1(12)*, 1-17.

Ashforth, B. E., & Mael, F. (1989). Social identity theory and the organization. *The Academy of Management Review*, *14(1)*, 20-39.

Baer, D., & McEachron, D. L. (1982). A review of selected sociobiological principles: Application to hominid evolution I. The development of group social structures. *Journal of Social and Biological Structures*, *5*, 69-90.

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

Bihary, J. G., Larson, J. R., Jr., & Tindale, R. S. (2015). *Individual vs. group advice-taking*. Poster presented at the Midwestern Psychological Association Eighty-Seventh Annual Meeting, Chicago, IL.

Brauner, E. (1996). *Attitudes and mental models in group interaction*. Paper presented at the Nags Head Conference on Groups, Networks, and Organizations, Highland Beach, IL.

Bunderson, J. S., & Boumgarden, P. (2010). Structure and learning in self-managed teams: Why "bureaucratic" teams can be better learners. *Organization Science*, *21*, 609-624.

Cannon-Bowers, J. A., Salas, E., & Converse, S. A. (1993). Shared mental models in expert team decision making. In N. J. Castellan, Jr. (Ed.), *Individual and group decision making: Current issues* (pp. 221-246). Hillsdale, NJ: Erlbaum.

Caporael, L. R. (1997). The evolution of truly social cognition: The core configurations model. *Personality and Social Psychology Review*, *1*, 276-298.

Carton A. M., & Cummings, J. N. (2012). A theory of subgroups in work teams. *Academy of Management Review*, *37*, 441-470.

Christensen, C., Larson, J. R., Jr., Abbott, A., Ardolino, A., Franz, T., & Pfeiffer, C. (2000). Decision-making of clinical teams: Communication patterns and diagnostic error. *Medical Decision Making*, *20*, 45-50.

Forsyth, D. R. (1999). *Group dynamics* (3rd ed.). New York: Brooks/Cole

Franz (2012). *Group dynamics and team interventions: Understanding and improving team performance*. Wiley-Blackwell.

Gino, F., & Moore, D. (2007). Effects of task difficulty on use of advice. *Journal of Behavioral Decision Making*, *20*, 21-35.

Greitemeyer, T., & Schulz-Hardt, S. (2003). Preference consistent evaluation of information in the hidden profile paradigm: Beyond group-level explanations for the dominance of shared information in group decisions. *Journal of Personality and Social Psychology*, *84*, 322-339.

Harvey, N. & Fischer, I. (1997). Taking advice, accepting help, improving judgment and sharing responsibility. *Organizational Behavior and Human Decision Processes, 70*, 117-133.

Hinsz, V. B. (1996). *Metacognitions in groups and the potential of shared mental models*. Paper presented at the Society of Experimental Social Psychologists Pre-Conference on Small Groups, Sturbridge, MA.

Hollingshead, A. B. (1996). The rank-order effect in group decision making. *Organizational Behavior and Human Decision Processes*, *68*, 181-193.

Janis, I. (1982). *Groupthink* (2nd ed.). Boston: Houghton-Mifflin.

Kelly, J. R., & Karau, S. J. (1999). Group decision making: The effects of initial preferences and time pressure. *Personality and Social Psychology Bulletin*, *25*, 1343-1354.

Kenrick, D. T., Maner, J., & Li, N. P. (2005). Evolutionary social psychology. In D. M. Buss (Ed.), *The handbook of evolutionary psychology* (pp. 803-827). Hoboken, NJ: Wiley.

Kerr, N. L., & Tindale, R. S. (2004). Small group decision making and performance. *Annual Review of Psychology*, *55*, 623-656.

Kruglanski, A. W. (1990). Motivations for judging and knowing: Implications for causal attribution. In E. T. Higgins & R. M. Sorrentino (Eds.), *Handbook of motivation and cognition: Foundations of social behavior* (vol. 2, pp. 333-368). New York: Guilford Press.

Kruglanski, A. W. (2004). *The psychology of closed mindedness*. New York: Psychology Press.

Kruglanski, A. W., Dechesne, M., Orehek, E., & Pierro, A. (2009). Three decades of lay epistemics: The why, how and who of knowledge formation. *European Review of Social Psychology*, *20*, 146-191.

Kruglanski, A. W., & Fishman, S. (2009). The need for cognitive closure. In M. R. Leary, & R. H. Hoyle (Eds.), *Handbook of individual differences in social behavior*, (pp. 343-353). New York: Guilford Press.

Kruglanski, A. W., & Webster, D. M. (1996). Motivated closing of the mind: "Seizing" and "freezing." *Psychological Review*, *103*, 263-283.

Larson, J. R., Jr. (2010). *In search of synergy in small group performance*. New York: Psychology Press.

Laughlin, P. R. (1980). Social combination processes of cooperative problem-solving groups on verbal intellective tasks. In M. Fishbein (Ed.), *Progress in social psychology* (pp. 127-155). Hillsdale, NJ: Erlbaum.

Laughlin, P. R. (2011). *Group problem solving*. Princeton, NJ: Princeton University Press.

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

Laughlin, P. R., Bonner, B. L., & Miner, A. G. (2002). Groups perform better than the best individuals on letters-to-numbers problems. *Organizational Behavior and Human Decision Processes, 88*, 605-620.

Laughlin, P. R., Hatch, E. C., Silver, J. S., & Boh, L. (2006). Groups perform better than the best individuals on letters-to-numbers problems: Effects of group size. *Journal of Personality and Social Psychology, 90*, 644-651.

Laughlin, P. R., Zander, M., L., Knievel, E. M., & Tan, T. K. (2003). Groups perform better than the best individuals on letters-to-numbers problems: Informative equations and effective reasoning. *Journal of Personality and Social Psychology, 85*, 684-694.

Levin, S., Federico, C. M., Sidanius, J., & Rabinowitz, J. L. (2002). Social dominance orientation and intergroup bias: The legitimation of favoritism for high-status groups. *Personality and Social Psychology Bulletin*, *28*, 144-157.

Mannix, E. A. (1993). Organizations as resource dilemmas: The effects of power balance on coalition formation in small groups. *Organizational Behavior and Human Decision Processes*, *55*, 1-22.

Mannix E. & Neale, M. A. (2005). What difference makes a difference? The promise and reality of diverse teams in organizations. *Psychological Science in the Public Interest*, *6(2)*, 31-55.

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

McLeod, P. L., Baron, R. S., Marti, M. W., & Yoon, K. (1997). The eyes have it: Minority influence in face-to-face and computer mediated group discussion. *Journal of Applied Psychology*, *82*, 706-718.

Minson, J. A., & Mueller, J. S. (2012). The cost of collaboration: Why joint decision making exacerbates rejection of outside information. *Psychological Science*, *23*, 219-224.

Packer, C. (1977). Reciprocal altruism in Papio anubis. *Nature*, *265*, 441-443.

Park, E. S., & Hinsz, V. B. (2006). "Strength and safety in numbers": A theoretical perspective on group influences on approach and avoidance motivation. *Motivation and Emotion*, *30*, 135-142.

Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, & Computers*, *36(4)*, 717-731.

Scholten, L., van Knippenberg, D., Nijstad, B. A., & De Dreu, C. K. W. (2007). Motivated information processing and group decision-making: Effects of process accountability on information processing and decision quality. *Journal of Experimental Social Psychology*, *45*, 539-552.

Schulz-Hardt, S., Brodbeck, F. C., Mojzisch, A., Kerscheriter, R., & Frey, D. (2006). Group decision making in hidden profile situations: Dissent as a facilitator for decision quality. *Journal of Personality and Social Psychology*, *91*, 1080-1093.

Sedikides, C., & Skowronski, J. J. (1997). The symbolic self in evolutionary context. *Personality and Social Psychology Review*, *1*, 80-102.

Sidanius, J., & Pratto, F. (1999). *Social dominance: An intergroup theory of social hierarchy and oppression*. New York: Cambridge University Press.

Sniezek, J. A., & Van Swol, L. M. (2001). Trust, confidence, and expertise in a judge-advisor system. Organizational behavior and human decision processes, 84(2), 288-307.

Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. In S. Leinhart (Ed.), *Sociological methodology 1982* (pp. 290-312). San Francisco: Jossey-Bass.

Soll, J. B., & Larrick, R. P. (2009). Strategies for revising judgment: How (and how well) people use others' opinions. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*, 780-805.

Stasser, G., & Stewart, D. D. (1992). Discovery of hidden profiles by decision-making groups: Solving a problem versus making a judgment. *Journal of Personality and Social Psychology*, *63*, 426-434.

Stasser, G., Stewart, D. D., & Wittenbaum, G. M. (1995). Expert roles and information exchange during discussion: The importance of knowing who knows what. *Journal of Experimental Social Psychology*, *31*, 244-265.

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

Stewart D. D., & Stasser, G. (1998). The sampling of critical, unshared information in decision-making groups: The role of an informed minority. *European Journal of Social Psychology*, *28*, 95-113.

Tajfel H., & Turner, J. C. (1986). The social identity theory of inter-group behavior. In S. Worchel & L. W. Austin (Eds.), *Psychology of intergroup relations*. Chicago: Nelson-Hall.

Tindale, R. S. (1989). Group vs. individual information processing: The effects of outcome feedback on decision-making. *Organizational Behavior and Human Decision Processes*, *44*, 454-473.

Tindale, R. S., Kameda, T., & Hinsz, V. B. (2003). Group decision making. In J. Cooper & M. Hogg (Eds.) *Sage Handbook of Social Psychology* (pp. 381-403). London: Sage Publications.

Tindale, R. S., Meisenhelder, H. M., Dykema-Engblade, A. A., & Hogg, M. A. (2001). Shared cognitions in small groups. In M. A. Hogg & R. S. Tindale (Eds.), *Blackwell handbook in social psychology: Group processes* (pp. 1-30). Oxford, UK: Blackwell Publishers.

## VITA

Joseph G. Bihary was born and raised in the suburbs of Long Island, NY. He attended the State University of New York at Stony Brook, where he earned a Bachelor of Arts in Psychology in 2010. As an undergraduate, Dr. Bihary worked as a research assistant studying auditory processing in Dr. Arthur Samuel's cognitive psychology laboratory.

Dr. Bihary continued his studies at Loyola University Chicago, working with Dr. James R. Larson, Jr. He earned a Master of Arts in Social Psychology in 2012, and a Doctor of Philosophy in Social Psychology in 2016. His primary research interests are in small group dynamics and performance, and the application of psychological research to organizational settings.