CARNEGIE MELLON UNIVERSITY

PROBLEM REPRESENTATION AND TEAM MENTAL MODEL DEVELOPMENT IN
INDIVIDUAL AND TEAM PROBLEM SOLVING PERFORMANCE

A DISSERTATION SUBMITTED IN
PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

In

PSYCHOLOGY

by

MATTHEW D WOOD

Department of Psychology
Carnegie Mellon University
Pittsburgh, Pennsylvania 15213
April, 2013

ACKNOWLEDGEMENTS

introduction to interdisciplinary research, and for demonstrating some of the ways in which

problems and tools can be mapped across fields to each's mutual benefit. These lessons have

surely helped me to develop a toolkit for professional and personal success for years to come.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Teams play an important role in solving today's complex problems. Disagreement exists about how teams should begin working on these problems to maximize performance. Ideation research encourages team members to produce many solutions to a problem and delay alignment of individual mental models of the problem. Team problem solving research advocates earlier and consistent alignment of individual mental models (team coherence). Problem solving research with individual solvers advocates clarifying and refining requirements for an ill-defined problem (representation construction), often via heuristics such as restating the problem in different ways. The purpose of this dissertation is to compare these different ways of starting on a problem, including the performance outcomes associated with each and the process by which these strategies produce their effect.

The process by which teams identified goals for solving a problem was more important than a coherent team mental model for solution quality across two experiments with three-member teams solving case study problems. Experiment 1A & 1B shows that teams who collaborate early on a problem produce better solutions than those who collaborate after a delay. Team coherence is affected by working together with teammates, but does not impact solution quality. Groups of individuals who work alone the entire time (nominal teams) provided interim self-reports of their best solution that were more like their final problem solutions than other groups. In addition, members of interacting teams identify more "best" solutions of a problem when describing the solving process than "nominal" team (three-person aggregates working individually) members.

Experiment 2 replicates Experiment 1 findings, and also shows that instructions to restate the problem had a similar positive effect on solution quality as early interaction. Working in

isolation with restatement instructions provided comparable results to collaborating with team members. Team coherence, in contrast, was not related to solution quality.

Both interaction and restatement improved quality by helping team members to consider alternative problem representations. Consideration of alternatives in turn may inoculate solvers from fixating on their first impressions of a best solution. Implications are discussed in terms of optimizing individual and team problem solving performance on ill-defined tasks.

Chapter 1. **Introduction**

1.1 Motivation

Teams play an increasingly important role in how complex problems are solved in diverse disciplines, including science and engineering (Paulus, Dzindolet, & Kohn, 2011), business (Reiter-Palmon, Herman, & Yammarino, 2008), and other fields (Fiore et al., 2010). One construct that has been used to study and improve team performance is the team mental model (Cannon-Bowers, Salas, & Converse, 1993). A team mental model refers to the combined set of problem representations held by individual team members, including both shared and unshared elements. Disagreement exists about the extent to which individual mental models of a problem should overlap (herein "team coherence"; Dong, Hill, & Agogino, 2004), and when during the course of problem solving individuals should be encouraged to develop consensus and reconcile their individual mental models. For instance, team mental models research tends to advocate an approach where teams should converge early during solving (Badke-Schaub, Neumann, Lauche, & Mohammed, 2007; Dong, et al., 2004), while brainstorming research suggests that teams should converge later (Hernandez, Shah, & Smith, 2010; Osborn, 1957). Additionally, team mental models research often follows problem-solving teams over a wide variety of timescales, ranging from hours to months (e.g., Dong, 2005)

Representation construction (known as "problem construction" in some domains; Mumford, Reiter-Palmon, & Redmond, 1994), or the process by which solvers clarify and refine problem requirements, is one mechanism that may explain this apparent discrepancy between the consequences of early or late coherence on the quality of problem solving outcomes. Priming individual problem solvers to think about different problem representations through instructions

to restate the problem in as many different ways as possible has been shown to improve solution outcomes (Baer, 1988; Reiter-Palmon & Robinson, 2009). Interaction between team members as they share their initial impressions of a problem at the start of the problem solving process may encourage teams to engage in representation construction.

1.2 Thesis Statement

As mentioned above, brainstorming and team mental models research disagrees about the degree to which teams should develop coherent mental models of a task in the course of solving a problem, as well as the time course of that coherence. Representation construction may present a mechanism that can negotiate these conflicting viewpoints. Individuals who are encouraged to "brainstorm" different ways of restating a problem produce better solutions than individuals who do not; members of teams may engage in an analogous process when they are primed by conflicting viewpoints of their fellow teammates. Representation construction is a two-stage process that has first a divergent stage where early team interaction provokes team members to think about different potential approaches for solving a problem, followed by a convergent stage where team members select amongst and/or combine identified approaches as they work toward a solution. The result is coherent team task mental model that is developed fairly early in the course of the problem solving process (Figure 1.1).

Latent Semantic Analysis (LSA; Deerwester, Dumais, Furnas, Landauer, & Harshman, 1990) is a tool that can be used to provide a quantitative estimate of team task mental model coherence whether these changes occur as a function of time, team interaction, or factors that may facilitate representation construction and therefore coherence. LSA uses information about the co-occurrence of words across a set of documents (e.g., self-reports, transcripts from

conversation) to develop an understanding of synonymy of words given the context of other words that they appear with across all documents. This general approach has been used successfully to model human judgments of similarity (e.g., Landauer, 1998; Lund & Burgess, 1996), and has been proposed as a psychological theory of how humans infer meaning through learning (Landauer, 2007). In the team problem solving literature, it has been used to track team mental model coherence from self-reports and other documents produced through the problem solving process (Dong, et al., 2004; Fu, Cagan, & Kotovsky, 2010). LSA produces a quantitative measure of semantic similarity that can be read like a correlation coefficient, with higher values indicating greater similarity between self-reports.

Using LSA as a computational tool to better track and understand human cognition, it is possible to achieve a better understanding of the relation between team coherence and the quality of resulting solutions in ill-defined problem contexts, as well as to monitor the effect that different interventions have on team coherence, solution quality, or both. Representation construction and strategies that promote representation construction may provide the link that explains the relationship between team coherence and solution quality identified in the literature.

The thesis of this work is that representation construction and problem solving activities which promote representation construction lead to changes in team and individual task mental models as well as the quality of resulting solutions.

Figure 1.1 Proposed Model of the Representation Construction Process in Team Problem Solving.

Each node represents a step in the process. Shaded areas represent stages.

1.3 Overview of Experiments and Hypotheses

Two experiments were conducted in order to better understand the relationship between representation construction, team coherence, team interaction structure, and problem solving outcomes. Three-member teams worked on a complex problem solving task for 28 minutes and were interrupted periodically to provide self-reports describing their current individual task mental model representations of the problem. At the end of each experiment, team members gave their solution to the problem. In Experiment 1A & 1B, team communication structure (nominal or interacting) was manipulated during the 1st and 2nd half of problem solving efforts to better understand the role that interaction has in team mental model development, and subsequently in solution quality. Self-report protocols from Experiment 1B were also analyzed to identify changes to the number of "best" solutions identified throughout the solving process by individual team members as a function of early team interaction. Experiment 2 builds on Experiment 1A/1B to draw direct parallels between how teams engage in the representation construction process naturally, and proscriptive heuristics used with individuals to facilitate representation construction. This was done by manipulating team interaction structure and whether explicit instructions to restate the problem were provided, and measuring the effects of these factors on representation construction behaviors, team coherence, and both interim and final problem solution quality. Three hypotheses are proposed.

1.3.1 Hypothesis One

H_1: *Interacting teams will have more representation construction, greater team coherence, and higher solution quality than nominal teams.*

Teams that interact, especially those who interact early, are able to share ideas about the problem. In sharing different solutions for a problem, team members may come to realize that their teammates are proposing ways of approaching the problem that are different from their own approaches. This process may be analogous to prescriptive heuristics deployed to encourage representation construction in individuals. Individuals, who do not have an opportunity to share viewpoints with others, are less likely to identify approaches to the problem beyond the first few they recognize early on in working with the problem. This potentially leaves nominal teams more prone to fixating on poor-quality ideas that they identified early in the problem solving process. As a result, nominal teams should also have lower team coherence, and produce poorer final problem solutions, than interacting teams.

Experiment 1A and 1B will test H_1 by manipulating the structure (interacting or nominal) and timing (early or late) of interaction while measuring team coherence and solution quality across conditions in two different groups of participants working on different types of problems. H_1 will be tested by coding Experiment 1B self-report protocols for evidence of representation construction, looking specifically at differences across conditions in the number of distinct solutions reported in individual team members' self-reports. Experiment 2 will test H_1 by manipulating team structure while measuring team coherence, solution quality, and the number of distinct solutions identified in self-reports using the methods deployed in Experiment 1A/1B and 2.

1.3.2 Hypothesis Two

     H_2: Instructions to restate the problem will facilitate representation construction in both interacting and nominal teams.

Explicit instructions to restate the problem are an effective method for inducing representation construction amongst individuals (Section 2.6). These instructions should also be effective at facilitating representation construction in interacting teams, as measured by the number of distinct solutions reported by team members throughout the problem solving process. However, since interacting teams are posited to engage in representation construction naturally as a function of sharing alternative viewpoints, explicit instructions to restate the problem may have less of an impact among interacting teams compared to nominal teams, in terms of representation construction behaviors and solution quality outcomes.

Experiment 2 will test H_2 by manipulating the instructions teams receive when working on the problem. Across both nominal and interacting teams, some will be asked to spend a few minutes restating the problem in as many different ways as they can (i.e., Baer, 1988), while others will be asked to begin work on the problem as usual. Representation construction will be measured by coding self-reports provided by individuals throughout the experiment for the number of distinct solutions identified, analogous to the coding scheme employed in Experiment 1B.

1.3.3 Hypothesis Three

H_3: *Representation construction will have a positive impact on team coherence and solution quality.*

Teams that engage in more representation construction should produce higher quality solutions than those who do not, comparable to individuals who engage in representation construction. In teams, the representation construction process will help individual team members recognize differences between their own representations of the problem and those of

7

other team members, and may therefore also facilitate the identification of common goals for solving. Identifying common goals for solving will in turn lead to greater team coherence among these teams compared to teams who engage in less representation construction.

Experiments 1B and 2 will test H_3 by coding self-reports for the number of distinct solutions among individual team members. Experiment 1B will provide information on the extent to which team structure and timing in changes to team structure influence representation construction (and therefore solution quality) as measured by the number of solutions identified by team members, while Experiment 2 will focus on the influence of explicit representation construction instructions and team structure on representation construction behavior and solution quality.

1.4 Dissertation Outline

This dissertation will investigate the relationship between representation construction, team coherence, and solution quality through a series of empirical studies that track the coherence of problem solving teams' task mental models while manipulating two factors that may facilitate or inhibit representation construction: the structure of team interaction, and the effect of instructions to identify different ways of conceptualizing a problem. A summary table of empirical findings and their relationship to hypotheses can be found at the end of the chapter (Table 1.1).

Chapter 2 introduces background literature related to research on team mental model development, representation construction, and related constructs which have been shown to facilitate team performance on problem solving tasks.

Chapter 3 presents two parallel experiments (Experiment 1A & 1B) that manipulate team communication structure as a means of influencing the ability of teams to develop a coherent mental model of a problem. Both experiments show that teams which interact in the process of problem solving have more coherent team mental models related to the task, and that those teams which interacted early during solving produced better solutions than those who interacted later. In addition, analysis of written protocols from Experiment 1B suggests that team interaction increases the number of "best" solutions team members consider while working on a problem, a phenomenon that may cue team members to think about the different ways in which the problem at hand may be represented.

Chapter 4 describes a final experiment (Experiment 2) exploring the influence of two strategies for inducing representation construction processes: team collaboration and explicit instruction. This experiment finds that instructions to restate a problem in different ways impart benefits similar to team interaction in terms of the quality of final problem solutions, while team coherence is unrelated to final problem solution performance.

Chapter 5 reviews findings of the earlier experiments, presents conclusions in light of the extant literature on team coherence, representation construction, and team performance, and additionally identifies several contributions of the present work.

Table 1.1 Summary of Dissertation Results and the Hypotheses Supported.

Hypothesis	Result	Section(s) / Figure(s)
H_1: Interacting teams will have more representation construction, greater team coherence, and higher solution quality than nominal teams.	Interacting teams consider a greater number of "best" interim solutions compared to nominal teams.	Section 3.3.2.5, Figure 3.11, Figure 3.12 Section 4.2.2.5, Figure 4.7, Figure 4.8
	Interacting teams take longer to commit to a particular representation of a problem compared to nominal teams.	Section 3.2.2.2, Figure 3.4 Section 3.3.2.2, Figure 3.8 Section 4.2.2.3, Figure 4.4 Section 4.2.2.4, Figure 4.6
	Nominal teams who never interact with their teammates consider few "best" interim solutions and produce final problem solutions that are most similar to these interim solutions, which together indicate that these teams become fixated on their initial representation of the problem.	Section 3.4, Figure 3.8, Figure 3.11

Hypothesis	Result	Section(s) / Figure(s)
H₁ (continued)	Interacting teams show greater team coherence compared to nominal teams.	Section 3.2.2.1, Figure 3.2 Section 3.3.2.1, Figure 3.6 Section 4.2.2.2, Figure J.1
	Interacting teams produce better solutions than nominal teams.	Section 3.2.2.3, Figure 3.5 Section 3.3.2.3, Figure 3.9 Section 4.2.2.6, Figure 4.10
H₂: Instructions to restate the problem will facilitate representation construction in both interacting and nominal teams.	Individuals working both alone and in teams provide more restatements to the problem when given instructions.	Section 4.2.2.1, Figure 4.2, Figure 4.3
	Representation construction instructions increase the number of "best" interim solutions considered in interacting and nominal teams.	Section 3.3.2.5, Figure 3.11 Section 4.2.2.5, Figure 4.7
H₃: Representation construction will have a positive impact on team coherence and solution quality.	Instructions to restate the problem increase solution quality, but only for nominal teams.	Section 4.2.2.6, Figure 4.10

Chapter 2. Background

The primary aim of this research is to learn more about the relationship between team coherence, problem representation processes, and the structure of team interaction, as well as how these factors influence the development of solution outcomes. Team creative problem solving research focuses on two distinct types of approaches. Ideation and brainstorming research (Linsey, Green, Murphy, Wood, & Markman, 2005; Osborn, 1957; Paulus, et al., 2011) uses tasks that encourage teams to produce many solutions to a problem. In contrast, research in team problem solving (Fiore, et al., 2010; Hayes, 1989; Herman & Reiter-Palmon, 2011; Salas & Cannon-Bowers, 1997) and engineering design (Badke-Schaub, et al., 2007; Dong, et al., 2004) tends to use tasks that encourage development of one or two specific solutions or designs. As a result of these differences in task demands, these two literatures promote different approaches for optimal team problem solving performance. Brainstorming research focuses on understanding the development of a shared mental model of a problem only after lots of broad, unconstrained search of the problem space (Hernandez, et al., 2010; Osborn, 1957), while team problem solving and design research focuses on the importance of the development of a shared mental model early during solving rather than later in the process (Dong, 2005; Fu, et al., 2010).

Brainstorming and ideation research find that a collection of individuals working alone typically produce more ideas than an interacting team, but it is unclear whether either individuals or teams produce ideas of higher quality (Kerr & Tindale, 2004; Paulus, et al., 2011). Although the research on team structure and ideation is robust, less is available on team interaction structure and problem solving processes more generally (Reiter-Palmon, et al., 2008). In particular, there has been relatively little study on the role that factors such as mental model

development have on the design or problem solving solutions that teams produce (Reiter-Palmon, Wigert, & de Vreede, 2011). The following will review mental models, team mental models, and various factors that may contribute to changes in team mental models.

2.1 Mental Models

A mental model (Craik, 1943) is a simplified, internal, cognitive representation of the world held by individuals with respect to a task, topic, or phenomenon of interest. These representations are used to reason about the world, simulate events in order to make predictions, and consider alternatives in order to solve problems (Johnson-Laird, 2006; Norman, 1983). A team mental model of a problem is the intersection of all individual mental models. Cannon-Bowers, Salas, and Converse (1993) decompose team mental models into four main components:

1) Equipment model – the tools available to a team in a problem solving or decision making context

2) Task model – the undertaking the team is engaged with, including goals to achieve, constraints to consider, and strategies to use for task completion

3) Team interaction model – the responsibilities and roles of each team member, and the available communication pathways between team members

4) Team model – the knowledge, skills, and abilities of team members, as well as their preferences and dispositions

In addition to these components, Badke et al (2007) propose two additional features of team mental models: the organizational context in which the task is situated (context model), and the team's self-perceived confidence in its ability to successfully complete the task (competence model). While each of these mental model concepts is a topic of considerable inquiry, the major

focus of research on concept development to-date has been on the task mental model held by individuals and teams, and on how that task mental model changes as a function of team interaction and other factors.

2.2 Team Mental Model Overlap

The extent to which team members' task mental models overlap has been shown to have a positive influence performance on team tasks. Task mental model overlap is referred to as convergence or sharedness in this literature, and is posited to improve task performance by helping team members to take actions based on their own knowledge that are consistent with the actions of teammates, in turn facilitating progress toward a common objective (Cannon-Bowers, et al., 1993; Mathieu, Heffner, Goodwin, Salas, & Cannon-Bowers, 2000). Team convergence on task mental models has been shown to predict task performance more than convergence on team mental models (Section 2.1; DeChurch & Mesmer-Magnus, 2010). Mohammed and colleagues (Mohammed, Ferzandi, & Hamilton, 2010) report that studies which measure team and task mental models separately typically find that sharedness in the task mental model has more direct impact on team performance than sharedness of the team mental model. Mathieu et al. (2000) found that task mental models exert their influence on team performance by affecting team processes, or the ways in which team members work together to solve the problem.

Although much of this research is done with highly structured team tasks (e.g., flight simulators; LePine, 2005; Mathieu, et al., 2000) in order to manipulate or monitor communication processes at a fine-grained level, and is less often conducted using less-structured design tasks whose goals for performance are sometimes ambiguous at the outset of a task, task mental model similarity has also been shown to affect performance in more real-world

14

contexts. Lim and Klein (2006) find this relationship between task mental model and performance to be true not only in laboratory simulation tasks like those traditionally reported in the literature, but also in real-world field settings among military infantry units engaged in field training exercises. Smith-Jentech et al (2005) also reported a relationship between task mental models and performance in real-world air traffic control settings.

Cooke and colleagues (Cooke, Gorman, Myers, & Duran, 2013) hypothesize that team member interaction in the form of verbal communication is equivalent to team cognition, in the sense that processes which exemplify coordination of efforts or requests for information among team members are often expressed in terms of verbal communication among team members, and that individual conceptions of what the current state of a problem is changes as a function of these communications with teammates. This notion is consistent with the work of others in the field (e.g., Klimoski & Mohammed, 1994) that team knowledge is more than the combination of knowledge held by individual team members, but is rather a constructive process whereby interaction amongst team members generates new knowledge that is then known to all team members.

2.3 Team Coherence

Team cognition research shows that team performance in problem solving tasks is often associated with improvements in the coherence of a team's mental model of the problem over time. Improvements in team coherence are typically associated with individual team members negotiating and aligning their conceptualizations of the problem with one another (Salas & Cannon-Bowers, 1997). Literature from a diverse set of applied domains suggests that team coherence is important for successful team problem solving outcomes (e.g., Reiter-Palmon, et al.,

15

2011). However, disagreement exists about when during the course of solving a team should be encouraged to develop a coherent team mental model. Team mental models research suggests that teams should develop a coherent task mental model early during the problem solving process, as this coherent mental model facilitates later coordinated activity among team members (Mathieu, et al., 2000). Research in team cognition also emphasizes convergence on a single problem representation early, and as a result tends to focus on how one or a small set of ideas develop into a single solution over time (Fiore, et al., 2010; Reiter-Palmon, et al., 2008; Salas & Cannon-Bowers, 1997).

Dong (2005) found that team coherence increases naturally over time as team members work together on a design problem in a semester-long design exercise. Individual ideation was measured by collecting emails and design documents, and analyzing the similarity of these documents to each other using latent semantic analysis (LSA; Deerwester, et al., 1990). LSA is a technique that produces a correlation-like measure of semantic similarity by comparing entropy-weighted frequency counts of words in a set of documents via multidimensional scaling (Landauer, Foltz, & Laham, 1998). Dong et al. (2004) also found a positive correlation between the semantic coherence of team members' individual mental models and the quality of resulting designs. In other words, teams with more coherent task mental models produced final project designs that were of superior quality across a wide range of design domains.

A more recent study by Fu and colleagues (2010) provides an example of the potential benefit of early team coherence in team problem solving. Fu et al. asked teams of engineering students to design a peanut sheller for deployment in resource poor, developing countries (see Linsey, et al., 2005). Along with the problem statement teams were provided with either a good example of a design solution, a poor example, or no example of a design solution. Examples

16

influenced team coherence as measured by LSA at the start of solving, and led to differences in both team coherence and final design quality. Specifically, good examples were associated with greater team coherence over the course of solving, and also higher quality final designs, than poor solutions. In conclusion, teams with coherent task mental models appear to develop better problem solutions, although it is unclear whether coherent task mental models cause high quality problem solutions or if a good solution promotes a more coherent task mental model.

2.4 Ideation

In contrast to research on team cognition and mental model coherence, ideation research suggests that consensus among team members should be delayed until further on in the solving process (Osborn, 1957; Paulus, et al., 2011), and this advice has become a ubiquitous part of the problem solving process for teams across a wide range of organizational and professional contexts (Brown & Paulus, 2002). While knowledge about the relationship between individual and team cognition in design and problem solving research in general is underdeveloped, a much more robust understanding of the relationship between individual and team cognition, as well as its impact on team outcomes, is available in the brainstorming literature. This area of inquiry has focused on differences in the quantity, quality, and variety of ideas developed by problem solving teams as a function of working independently or interactively on ideation tasks. First developed as a component of Osborn's proscriptive Creative Problem Solving model (Osborn, 1957), brainstorming is associated with a series of rules that explicitly exclude criticism and critical thinking and therefore discourage team coherence during that phase of the problem solving process. These rules include (1) focusing on quantity (2) ruling out criticism, (3) encouraging unusual ideas, and (4) combining and improving suggested ideas. Once a set of

ideas has been generated, a subset is selected for further refinement, combination, and eventual implementation.

Practitioners in team cognition have since developed variations on this set of core tenets. Some, especially those in engineering design, include a graphical component in addition to textual information. For instance, C-sketch (Shah, Vargas-Hernandez, Summers, & Kulkarni, 2001) uses five-person design teams where each member sketches a solution candidate, then passes that design to another team member to extend or modify, until all team members have an opportunity to refine or add to the work of other members. C-sketch was inspired by the textual method 6-3-5 (Otto & Wood, 2001) that uses six-member teams, where each member writes and passes three ideas a total of five times to other team members. These rotational ideation methods are different from gallery approaches that ask team members to share all concepts simultaneously with other members between ideation episodes or traditional brainstorming approaches where ideas are shared on-the-fly (Linsey, et al., 2005). Linsey et al. (2011) found that rotational methods that include both words and sketches tend to produce the most ideas, but that sketch-only methods produce higher quality ideas, and that high-quality ideas tend to be the result of development of earlier design concepts. Although these methods encourage coherence over time with their variations on earlier-suggested concepts, they do so with some delay.

The team ideation literature has also provided insights on the effectiveness of different team structures in the generation and selection of ideas for further refinement. Two types of teams that are often compared in this context are nominal teams (whose members work independently) and interacting teams (whose members work together). Nominal teams, unlike interacting ones, are unable to establish a coherent mental model of a problem because the members cannot communicate about the problem with each other. Both empirical work (Kerr &

18

Tindale, 2004; Putman & Paulus, 2009) and simulation (Brown & Paulus, 1996; Brown, Tumeo, Larey, & Paulus, 1998) research indicates that nominal teams generate more ideas than interacting teams in brainstorming contexts; the performance difference for interacting teams is a phenomenon known as *productivity loss* (Diehl & Stroebe, 1987). Structured ideation models attempt to overcome this issue. However, the ideation literature focuses primarily on properties of the resulting pool of ideas (e.g., quantity, quality, variety, novelty; Hernandez, et al., 2010; Kerr & Tindale, 2004; Paulus, et al., 2011) and tends not to consider process variables that describe the activity that is occurring to produce these ideas. One exception are computational models of brainstorming teams which rely on memory mechanisms used to explain cued recall effects in order to replicate phenomena like productivity loss (Brown & Paulus, 1996; Brown, et al., 1998).

Interacting teams tend to perform better than individuals if evaluating and selecting ideas when fewer alternatives are available, but this is only true when the products of an earlier ideation session are already available for evaluation (Mumford, Feldman, Hein, & Nagao, 2001) and does not consider ideas that are created, developed, and refined on-the-fly. Most research focused on team interaction structure is only interested in comparing performance of interacting teams to nominal teams (Reiter-Palmon, et al., 2008). Such research has less to say on how interaction and/or independent work may be combined to produce outcomes that are different (and possibly better) than that of either exclusively collective or exclusively independent work. If problem solving teams are able to recognize the best ideas from the search process prescribed by brainstorming, encouraging team members to work independently to explore the problem space, followed by team interaction to select parts of the space to pursue and refine for a solution should lead to the best solution outcomes.

However, recent research suggests that both nominal and interacting teams choose randomly when asked to select the highest quality ideas from a pool generated during brainstorming. This is despite the fact that nominal teams tended to produce more high-quality ideas (Faure, 2004; Putman & Paulus, 2009; Rietzschel, Nijstad, & Stroebe, 2006). It should be noted that the above studies implement a process that is only partially representative of the problem solving process in general. The focus was only on the generation and identification of alternatives, and subsequent selection of a subset of them based on their quality. Problem solving teams in the real world frequently combine and modify ideas while searching the problem space for alternatives. In addition, the process of generation then selection of ideas artificially decomposes the process into two phases, whereas problem solving teams outside the laboratory are more likely to iterate through idea generation and selection multiple times through the course of problem solving.

2.5 Fixation

One reason that interacting problem solving teams may have difficulty in selecting from a candidate pool of brainstormed ideas based on quality may be because the team focuses on one category of problem solutions, and is then unable to solve the problem effectively because members cannot help themselves but to think about the fixated category. Amongst individual problem solvers this phenomenon is known as fixation, and has been shown to powerfully influence the quality and likelihood of successful problem solving outcomes. Problem solving fixation was first described in the psychological literature by Woodworth and Schlosberg as a reluctance of the solver to relax initial assumptions used to approach the problem (1962). A seminal paradigm illustrating the deleterious effects of fixation was developed by Steve Smith

20

and colleagues (Jansson & Smith, 1991; S. M. Smith & Blankenship, 1989, 1991). In experiments using verbal puzzle problems, Smith and Blankenship (1989, 1991) showed that fixation could be induced by providing misleading hints along with the initial presentation of the problem. Both Jansson and Smith (1991) and Smith, Ward, and Schumacher (1993) extended these results to conceptual design, demonstrating the impact of poor examples in the conceptual design of mechanical devices, children's toys, and fictional lifeforms.

Some team cognition research shows that teams are also prone to the deleterious effects of fixation early in the problem solving process. Over three experiments using a brainstorming task, Kohn and Smith (2011) found that three- to four-person groups tend to explore ideas from fewer categories when working together than they would when working independently. In addition, suggested ideas provided by a confederate team member caused participant team members to produce ideas from the same category as the suggested idea. Other research by C. M. Smith and colleagues (C. M. Smith, Bushouse, & Lord, 2010) introduced helpful or misleading hints alongside rebus puzzles to three-person groups as well as individuals, similar to Smith and Blankenship (1989), with comparable results. Misleading hints tended to inhibit progress on the problem. Fu et al. (2010) show a similar negative impact of misleading design examples on the quality of proposed solutions among design teams.

Teams and individuals are both prone to the effects of fixation when the affecting examples are provided early on in the solving process. Much of the research above used an unrelated interruption task to ameliorate the impact of misleading examples by flushing these problematic ideas out of the working memory of individuals or team members (the incubation effect; Silveira, 1971; Wallas, 1926). Experts in creative domains develop strategies to avoid fixation effects like those exemplified by design fixation. Isaak and Just (1995) hypothesize that

21

inventors self-impose constraints to identify a design goal or evaluate a solution, and relax these

constraints to generate possible solution candidates. While a design example like those seen in

the experiments above may make it more difficult for individuals to relax example constraints, a

strategy that identifies and relaxes some of these self-imposed constraints should reduce fixation.

Kim, Kim, Lee, and Park (2007) found that expert designers work on one component of a design

then move onto another component, without necessarily finishing the first, in order to avoid

design fixation. Novices, in contrast, tend to complete one design component before moving

onto the next component. They refer to the expert's behavior as the *Limited Commitment Mode*

control strategy, to emphasize the fact that experts wait before committing to a particular design

component's configuration until several components have been considered.

Both of the above strategies are deliberate ways which expert solvers have been shown to

employ for reducing fixation on complex problem solving tasks. The conclusion is that in order

to reduce problem solving fixation in teams, similar strategies that reduce the likelihood that

inappropriate problem information like misleading hints become the focus of solving efforts in

the first place might be employed. Strategies that encourage representation construction may

provide one such way to reduce the likelihood of fixation.

2.6 Representation Construction

Attention to the representation construction process may be one way to negotiate

discrepancies between the brainstorming and problem solving literatures, as well as facilitate

performance of problem solving teams by reducing fixation. Representation construction may

also help explain the relationship in the team problem solving and design literatures between

team mental model coherence and team solution quality. Problem solving research with

individuals demonstrates that representation construction, or the process of clarifying and refining requirements for an ill-defined problem (Mumford, et al., 1994), has important consequences for the quality of later problem solving activities. Seminal research in this area shows that the difficulty of well-defined problems with the same underlying structure can be modified by changing surface features of the problem, which encourages solvers to frame the problem in different ways (Kotovsky, Hayes, & Simon, 1985). For ill-defined problems, a common way to develop a clear understanding of problem requirements is to ask individuals to consider as many different ways of restating the problem as possible (Baer, 1988). Not only do representation construction heuristics like restatement lead to improved solving outcomes, but experts in a wide variety of domains have been shown to spend more time than novices on representation construction (Reiter-Palmon & Robinson, 2009), and creativity training programs that include a representation construction component have been shown to be effective (Scott, Leritz, & Mumford, 2004).

Team problem solving research shows that team performance is associated with increased alignment of individual team members' mental models of a problem (Salas & Cannon-Bowers, 1997). Work on team coherence described earlier (Dong, 2005; Fu, et al., 2010) suggests that team coherence can affect the quality of final design solutions, although the data they present linking coherence and quality makes it difficult to draw a direct causal link. Research on the problem solving process with individuals demonstrates that representation construction, or the process of "… defining the goals, objectives, and parameters of the problem-solving effort (Mumford, et al., 1994, p. 6)," has important consequences for the quality of later problem solving activities. Team members who interact early in the problem solving process have the opportunity to share the differences in how they interpreted the problem early on in problem

solving, and are thus exposed to different ways of representing the problem. In contrast, individuals who work alone are not given assistance with representing the problem in different ways, and are therefore more likely to fixate on one particular problem frame and neglect alternative approaches as they develop their initial conceptions toward a solution (Jansson & Smith, 1991; S. M. Smith & Blankenship, 1989, 1991). This may lead individuals who work together early in the course of solving to generate better problem representations, and therefore better solutions to the problem than a comparably sized group of individuals working in isolation. In contrast, a delay before initial interaction may encourage individuals to become entrenched in their initial representation of the problem, and reduce the likelihood of late-interacting teams adopting a common problem representation compared to early-interacting teams. The result may be use of suboptimal problem representations and less team coherence for teams whose members begin work on a problem independently, as well as a negative effect on the quality of the problem solutions those teams ultimately produce, compared to teams whose members begin work on a problem by interacting with each other.

Chapter 3. Team Interaction Structure

3.1 Introduction

The goal of the present work was to clarify the role that team coherence plays in affecting performance on complex problem solving tasks. Students in engineering (Experiment 1A) and psychology (Experiment 1B) worked together or independently during the 1st or 2nd half of their work on the problem. In Experiment 1A, teams developed a conceptual design for a peanut sheller to be deployed in developing nations (Linsey, et al., 2005), and were interrupted occasionally for self-reports of their current representation of the problem (Fu, et al., 2010). The experiment used a 2x2 between-groups factorial design. Team members either worked together or independently on the problem during the 1st half and/or 2nd half of their solving efforts. Experiment 1B used an analogous design to Experiment 1A, but used a more general sample of undergraduate students from introductory psychology classes working on a campus improvement problem. The originality of proposed final problem solutions were also evaluated in Experiment 1B to more closely link these results with those of the brainstorming literature, and a count of "best" interim solutions that individuals identified during the course of their solving in self-reports was made to identify behavior consistent with representation construction.

3.2 Experiment 1A

3.2.1 Method

    3.2.1.1 Participants. Participants were n = 80 upperclass undergraduate and masters' level graduate students recruited from engineering design methods classes at Carnegie Mellon University. Students participated for 1¼ hours in exchange for a chance in a raffle for one of four $25.00 Amazon gift cards. Participants were asked to bring their own laptop computers to give individual self-reports of their problem solving efforts throughout the experiment.

    3.2.1.2 Design. This experiment utilized a 2x2 between-groups factorial design, where team members were either able to work together (interacting structure) or independently (nominal structure) on the problem during the 1st half and/or 2nd half of the experiment. This created four team structure conditions (expressed herein as 1st half structure – 2nd half structure): nominal-nominal, interacting-nominal, nominal-interacting, interacting-interacting (Table 3.1).

    3.2.1.3 Materials. Participants were asked to design a peanut sheller for use in developing countries (Peanut Sheller problem). This problem has been used successfully in the past with comparable student groups to study various aspects of design cognition, including the efficacy of different group ideation techniques (Linsey, et al., 2005), and the role of preliminary examples in design fixation (Fu, et al., 2010). The problem asks participants to develop an inexpensive solution for shelling peanuts with an eye on increased efficiency relative to hand-shelling, ease of manufacturing, and several other relevant factors (Appendix A).

Table 3.1 Design for Experiment 1A & 1B.

2nd HALF

	nominal	interacting
nominal	nominal-nominal	nominal-interacting
interacting	interacting-nominal	interacting-interacting

1st HALF

Instructions were provided to volunteers through a combination of spoken and written means. The problem statement and instructions to work alone (nominal condition) or with team members (interacting condition) throughout the experiment were provided via a series of numbered, folded sheets of paper in a 9x12 inch envelope given to each participant. Each sheet was stapled to itself to discourage participants from looking ahead. In addition, participants in all conditions were asked to periodically provide a self-report of their best approach or solution to the Peanut Sheller problem. These self-reports were typed into a series of Microsoft Word document files pre-loaded onto each volunteer's laptop computer, with a prompt at the top of each file (Appendix B).

3.2.1.4 Procedure. Several teams participated in a session. Individuals signed up for a session, and were randomly assigned to three-member teams upon arrival. In cases where there were more individuals than could be used to form three-member teams, the remaining individuals were assigned to an extra condition that was like the nominal-nominal condition in every way, with the exception that these individuals were not affiliated with any team (extra-nominal condition). The result was 24 three-member teams (n = 72 individuals), with six teams
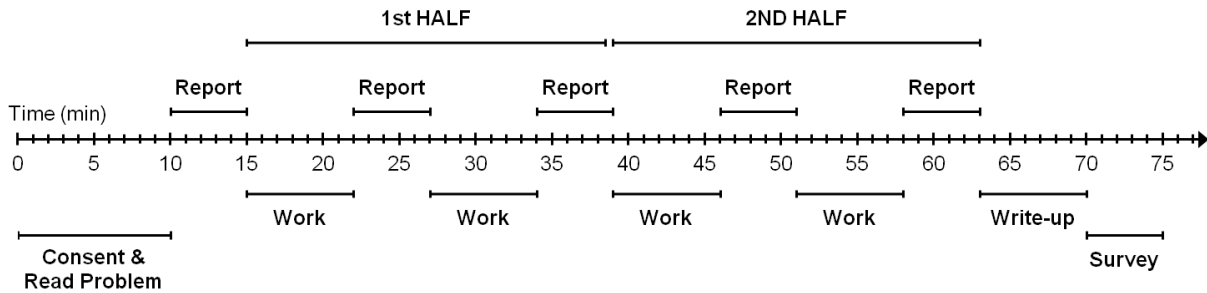
Figure 3.1 Procedure for Experiment 1A & 1B.

in each of the four experimental conditions, and eight individuals in the extra-nominal condition.

Teams were randomly assigned to condition, and all teams in a given session participated in the

same condition. This was done to keep team members naïve to the factors being manipulated

and researcher hypotheses.

Participants were given 28 minutes to work on the design problem, split up into four

smaller seven-minute epochs, with one additional seven-minute epoch to write up their final

design solution (Figure 3.1). For each work epoch, they were informed that they would proceed

on the problem either by themselves (nominal condition), or with the members of their team

(interacting condition), and that they would be notified at the start of each epoch whether they

would be working alone or as a team. In sum, participants worked alone or together during both

of the first two epochs (1st half) and/or both of the last two epochs (2nd half) of their solving

efforts. Between epochs of work on the problem, participants provided independent self-reports

about the problem solving process.

After working on the problem, teams wrote up their solution(s) to the design problem.

Interacting teams during the 2nd half of the experiment produced one team design solution during

this final write-up epoch. Nominal teams during the 2nd half of the experiment produced three

distinct solutions, one for each individual team member. All teams and individuals were

instructed to include as part of their solution description: a sketch of the design, labels of major

elements, and a few sentences describing how the solution works. Teams then completed a brief

questionnaire to supply demographic information (e.g., age, major) and a rating of their

familiarity with the other members of their team.

3.2.2 Results

The experiment produced two different sets of data: a set of self-reports over time for

each team that was used to evaluate team coherence, and one or a set of design solutions for each

team that was used to evaluate team performance in terms of solution quality. Self-report data

was analyzed using LSA (below). The quality of final design solutions were evaluated by

trained raters (Section 3.2.2.3). LSA metrics and design quality were subsequently compared

across conditions using an analysis of variance approach. Data from participants in the extra-

nominal condition was used to train the LSA model, as well as for design solution rater

calibration, but was not otherwise part of statistical analyses of the results.

Team members provided self-reports of their team's current best approach or solution to

the Peanut Sheller problem after reading the problem, between problem solving sessions, and

before providing their final design solution (Figure 3.1 & Appendix B). LSA (Deerwester, et al.,

1990; Landauer, et al., 1998) was used to compare these self-reports and transcripts of the final

design descriptions to each other, consistent with the method of Fu and colleagues (2010). LSA

was initiated by forming a word-by-document matrix, where the cells held a frequency count of

the number of instances of each word (represented by matrix rows) in each self-report or final

design transcript document (matrix columns). An "entropy weighting" step was performed to differentially weight word-type occurrences based on their inferred importance, giving greater weight to words that are infrequent in the corpus in general but very frequent in the document in question. Singular value decomposition (SVD) was performed on the transformed matrix, setting the number of dimensions to the number of documents (n = 458), and resulting in three matrices: the left (U) and right (V) orthonormal singular matrices, and the diagonal matrix of singular values (S). A measure of cosine similarity (Dong, 2005) was applied to the S and transpose of V matrices, producing a matrix of semantic similarity values between documents with cells populated with values ranging from -1 to +1. This measure is interpreted in a manner similar to a correlation coefficient, with positive values indicating synonymy, negative values indicating antonymy, and the magnitude reflecting the strength of the relationship between documents (Deerwester, et al., 1990). Though the corpus of documents used here is relatively small compared to typical applications of LSA, similarly sized corpora have been used in the past with satisfactory results (Fu, et al., 2010; Strait, Haynes, & Foltz, 2000). In addition, LSA and similar approaches (e.g., Lund & Burgess, 1996) have been used successfully in the past to model similarity judgments made by human raters in complex domains (Quesada, 2007), as well as priming and other memory-based phenomena that are influenced by semantic similarity (McNamara, 2005).

Two metrics were developed using LSA to better understand the team design process. A metric of team semantic coherence was used to understand changes in the similarity of individual team members' mental models of the design task at each point in time as a function of time and experimental condition. Team coherence was computed by taking the mean of all pairwise cosine similarities of self-reports in the final LSA matrix for the members of each team and for

each point in time. The result is five semantic coherence values for each team, corresponding to each time during the procedure that a self-report was elicited. Teams whose members have very similar mental models will have high team semantic coherence values; those whose members have dissimilar mental models will have low semantic coherence values.

A metric was also developed to better understand the change in individual team member's problem representations over time by comparing final design solutions provided by each individual or team to self-reports made by each individual throughout the design problem solving process. Referred to herein as "similarity-to-solution," this metric reflects the extent to which team member's mental models over time were consistent with the final design solution that was provided at the end of the experiment. The result of these comparisons is five similarity-to-solution values for each team member, corresponding to each time a self-report was elicited. Individuals who began the experiment with a specific solution to the design problem and stayed with that solution throughout should show high and stable similarity-to-solution values across time. In contrast, individuals whose ideal design solutions changed from their initial conception, either as a function of time or interaction with other team members, should show similar modulations in similarity-to-solution values.

3.2.2.1 Team coherence. In order to test the hypothesis that team interaction structure would influence team coherence (H_1), a repeated-measures factorial ANOVA was conducted on coherence scores for each team with the two team interaction structure factors (1st half structure, 2nd half structure) entered as a between-groups effects, and time entered as a within-groups effect. This analysis identified a significant main effect of 2nd half team interaction structure, with 2nd half interacting teams having higher team coherence over the course of the entire experiment ($M = 0.473$, $SD = 0.045$) compared to 2nd half nominal teams ($M = 0.414$, $SD =$

0.045), $F(1,20) = 10.113$, $p = 0.005$, $\eta^2 = 0.336$. This main effect was qualified by two statistical

interactions. The first of these was a Time by 2nd half team interaction structure statistical

interaction, $F(4,80) = 4.155$, $p = 0.004$, $\eta^2 = 0.172$. Follow-up within-groups contrasts suggests

coherence for 2nd half nominal teams dropped from t_2 (after 1st half and before 2nd half) to t_3
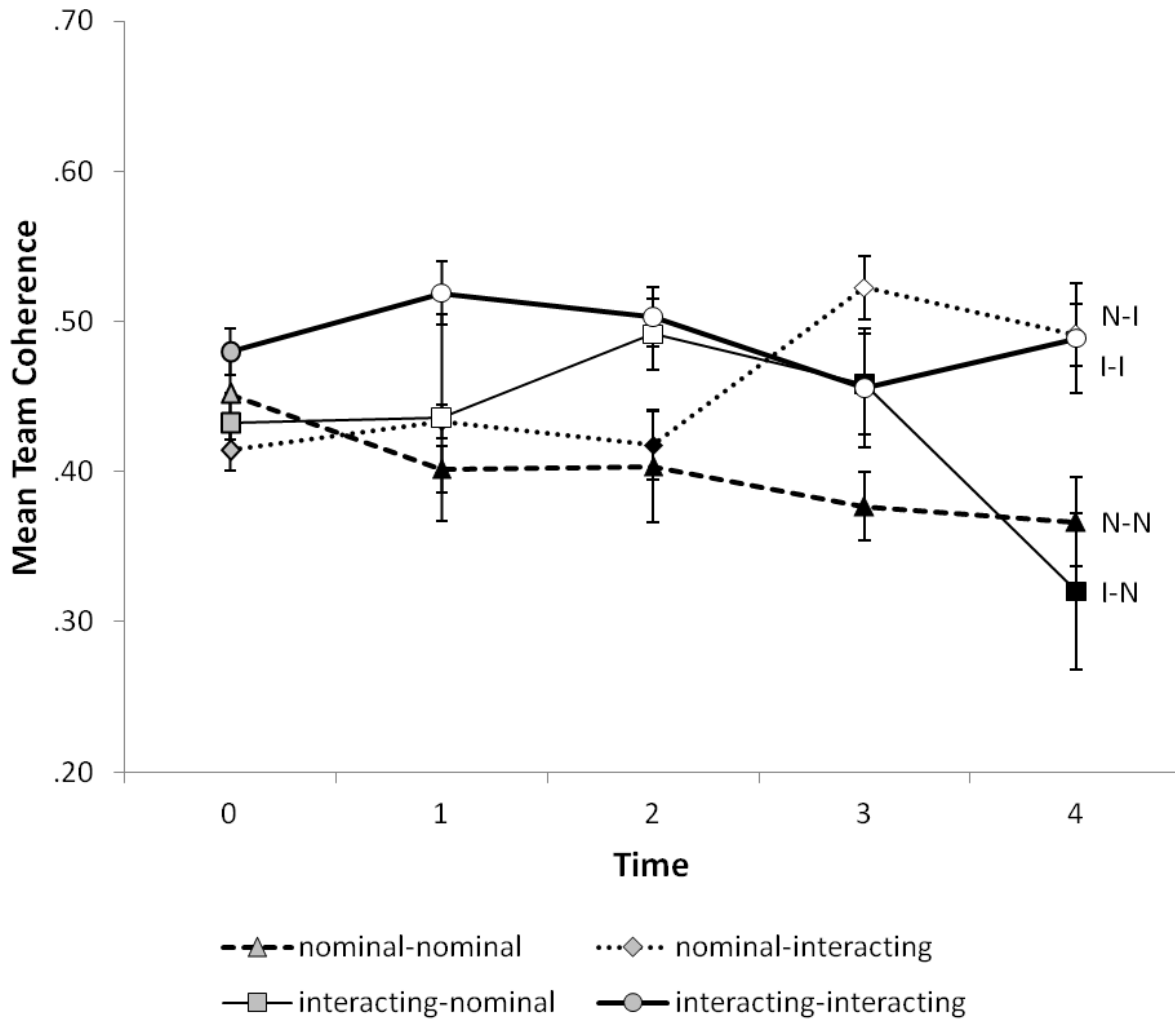
Figure 3.2 Experiment 1A Team Coherence Over Time.

Each line represents a different team interaction structure condition. White markers indicate times at which teams were interacting for the previous work epoch; black markers indicate a nominal team interaction structure for the prior work epoch. Line style indicates 1st half interacting (solid) or nominal (dashed) team structure. Line weight indicates whether 1st & 2nd half structure were the same (heavy) or different (light). Error bars indicate 1SE.
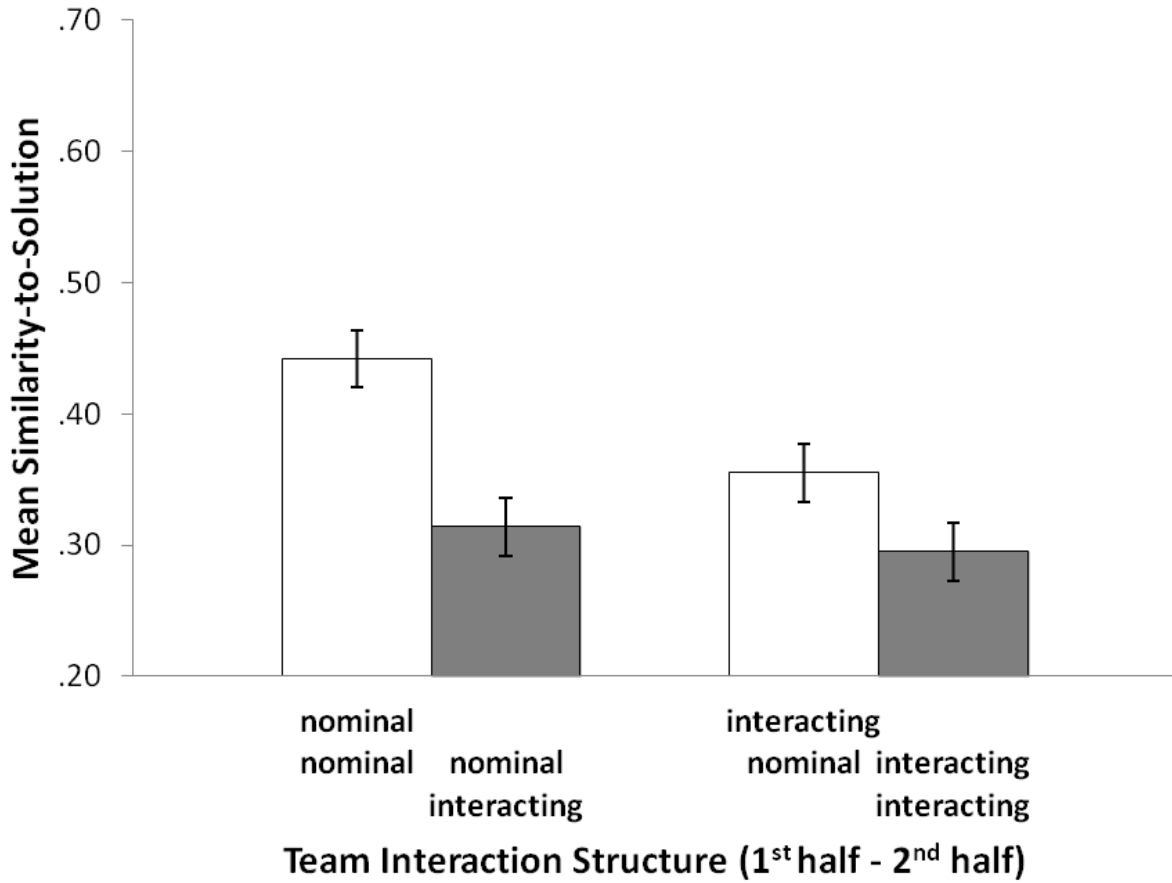
Figure 3.3 Experiment 1A Mean Similarity-to-Solution.

Bar grouping indicates 1st half team interaction structure (nominal left, interacting right), bar color indicates 2nd half team interaction structure (nominal white, interacting grey). Error bars indicate 1SE.

(midway through 2nd half), and t_3 to t_4 (end of 2nd half), while coherence for 2nd half interacting teams tended to increase (Figure 3.2). This was further qualified by a Time by 1st half team interaction structure by 2nd half team interaction structure statistical interaction, $F(4,80) = 2.711$, $p = 0.036$, $\eta^2 = 0.119$. This statistical interaction is likely explained by the behavior of the nominal-interacting and interacting-nominal groups, who saw sharp changes in team coherence from t_2 to t_3 consistent with the change in team structure, either showing a sharp increase

33

(nominal-interacting) or sharp decrease (interacting-nominal) as they began to work together or separately respectively.

Team coherence results suggest that interaction between team members increases team coherence. These results are consistent with H_1, as teams who work together have the opportunity to share ideas for solving the problem with each other. In contrast, team members assigned to the nominal condition showed low or decreasing team coherence scores across time as they were not able to share their ideas.

3.2.2.2 Similarity-to-solution. In order to better understand the relationship between team interaction structure and individual mental models over time, a repeated-measures factorial ANOVA was conducted on similarity-to-solution scores for each team with the two team interaction structure factors (1st half structure, 2nd half structure) entered as between-subjects effects, and time entered as a within-subjects effect. This analysis identified a significant main effect of 1st half interaction structure, $F(1,68) = 5.558$, $p = 0.021$, $\eta^2 = 0.076$, where members of 1st half nominal teams had higher similarity-to-solution scores over the course of the experiment ($M = 0.378$, $SD = 0.016$) than 1st half interacting teams ($M = 0.325$, $SD = 0.016$). A main effect of 2nd half interaction structure was also identified, $F(1,68) = 17.591$, $p < 0.001$, $\eta^2 = 0.206$, whereby 2nd half nominal team members also had higher similarity-to-solution scores over the course of the experiment ($M = 0.398$, $SD = 0.016$) than 2nd half interacting individuals ($M = 0.304$, $SD = 0.016$). The effect of 2nd half interaction structure is conflated, at least in part, with the similarity-to-solution LSA measure. 2nd half interacting team members had their individual self-reports compared to a common team design solution to produce the similarity-to-solution metric, whereas 2nd half nominal team members had their individual self-reports compared to
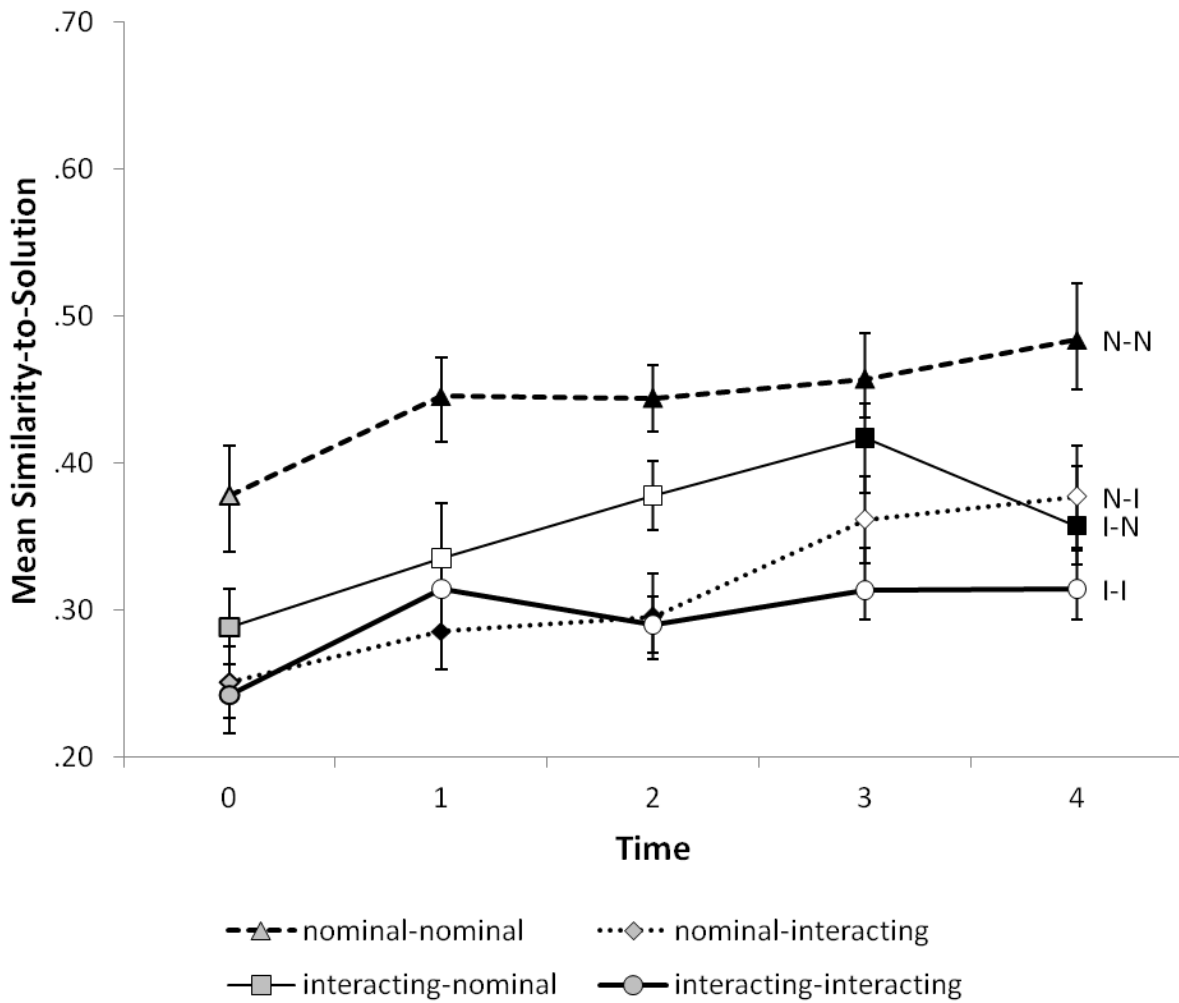
Figure 3.4 Experiment 1A Similarity-to-Solution Over Time.

Each line represents a different team interaction structure condition. White markers indicate times at which teams were interacting for the previous work epoch; black markers indicate a nominal team interaction structure for the prior work epoch. Line style indicates 1st half interacting (solid) or nominal (dashed) team structure. Line weight indicates whether 1st & 2nd half structure were the same (heavy) or different (light). Error bars indicate 1SE.

their individual design solutions. There was no statistical interaction present between 1st half and 2nd half interaction structure (Figure 3.3). In addition to the main effects of the experimental manipulation, the analysis also identified a main effect of time, $F(4,272) = 14.999$, $p < 0.001$, η^2

= 0.181. Similarity-to-solution scores tended to increase over time, regardless of the particular experimental condition (Figure 3.4).

Similarity-to-solution results suggest that members of nominal-nominal teams are likely to focus on a representation of the problem that is translated into their final design solution. In contrast, teams who experience any amount of interaction with other team members have much lower similarity-to-solution scores. This likely means that members of interacting teams entertained several different ways to represent the problem during the solving process, of which only one or two representations were selected for incorporation into the final design solution. It is possible for interacting teams to entertain multiple representations while demonstrating high team coherence values, so long as each team member describes a few of the same representations in their self-reports of the process.

3.2.2.3 Final quality assessment. Final designs were judged for quality using a Pugh chart by a masters' level student in mechanical engineering and a Ph.D. student in psychology familiar with research in mechanical design and past research using this particular design problem. The designs were assessed on several criteria relevant to deployment of the shelling device in a resource-scarce environment typical of the developing nations described in the problem statement (e.g., cost, feasibility, power source). These criteria correspond to the design requirements explicated in the design problem statement (Appendix C). Each design was compared on each criterion on a -2 to +2 scale. For each criteria, "-2" indicates that the rated design is "much worse" than the datum solution, "0" indicates that the rated design and datum solution are comparable, and "+2" indicates that the rated design is "much better" than the datum solution.

The Pugh chart was adapted from that used by Fu et al. (2010), with the exception that a datum solution was identified from design solutions in the dataset that was roughly prototypic, or average relative to other design solutions, on most Pugh chart performance dimensions. The two raters reviewed all designs in the dataset and identified the prototypic datum based on its design features and a rough assessment of its performance. This was done because most designs in this dataset performed worse than the best known existing solution used by Fu et al. (2010; Full Belly Project, 2010). In addition, a relative rating approach takes into account the relative expertise level of the participant set. Such approaches have been successfully used in the past to evaluate products in a wide variety of domains with success (e.g., Amabile, 1996; Redmond, Mumford, & Teach, 1993; Wood, Hocker, Hunter, & Ligon, 2011). After identifying the datum solution, the

Table 3.2 Intraclass Correlation Coefficients on Training Set for Experiment 1A.

Criterion	α
Cost	0.851
Removes shell w/o damaging peanut	0.696
Separates shell & nut	0.563
Amt. of peanuts shelled per hr	0.694
Energy source required to operate	0.625
Amt. of energy required	0.879
Feasibility	0.838
Size/Portability	0.631
# of people needed to operate	0.790
Time to set up & build	0.615
TOTAL	0.684

Pugh chart scales were adjusted to accommodate the range of performance on each criterion in the dataset. Abstracted examples of characteristics that fit each level of each criterion were identified by both raters and incorporated into the Pugh chart where appropriate.

Both raters independently coded a random selection of 25% of all designs using this revised Pugh chart to assess rater reliability. Reliability was assessed using intraclass correlation (ICC) with relative agreement (McGraw & Wong, 1996; Shrout & Fleiss, 1979) on 25% (n = 15) of all final designs. A relaxed ICC criterion of $\alpha = 0.6$ overall ratings was established because of the imprecise nature of evaluations on some criteria and the incomplete nature of some final design descriptions. Interrater agreement was acceptable in total (ICC, $\alpha = 0.684$), and was higher than the criterion on all but one criteria (Table 3.2). Remaining designs were randomly assigned to one of the two raters for evaluation.

In an initial analysis, quality ratings were averaged across criteria to produce one value (aggregate quality). Aggregate quality values were entered into a 2x2 between-subjects ANOVA, with the team interaction structure entered as the two factors (1st half, 2nd half). For 2nd half nominal teams, each individual received the aggregate quality score from their own final design. For 2nd half interacting teams, which produced one final design per team, the aggregate quality score for the team's design was assigned to each individual. This procedure was used, as opposed to averaging nominal team member scores, in order to better estimate the variance across all designs. The ANOVA identified a marginally significant effect of 1st half structure, $F(1,68) = 3.073$, $p = 0.084$, $\eta^2 = 0.043$, but identified neither an effect of 2nd half team interaction structure nor a 1st half by 2nd half team interaction structure statistical interaction. Aggregate quality scores for 1st half interacting teams ($M = 0.011$, $SD = 0.496$) was greater than that for 1st half nominal teams ($M = -0.207$, $SD = 0.550$; Figure 3.5).

In order to identify which criteria were most influential in producing quality differences across conditions, a multivariate analysis of variance (MANOVA) was conducted, entering each criterion as a separate dependent variable into the analysis. This analysis identified a main effect of 1st half structure, $F(10,59) = 4.057$, $p < 0.001$, $\eta^2 = 0.407$. In addition, a significant statistical interaction between 1st half and 2nd half team interaction structure was detected, $F(10,59) =$
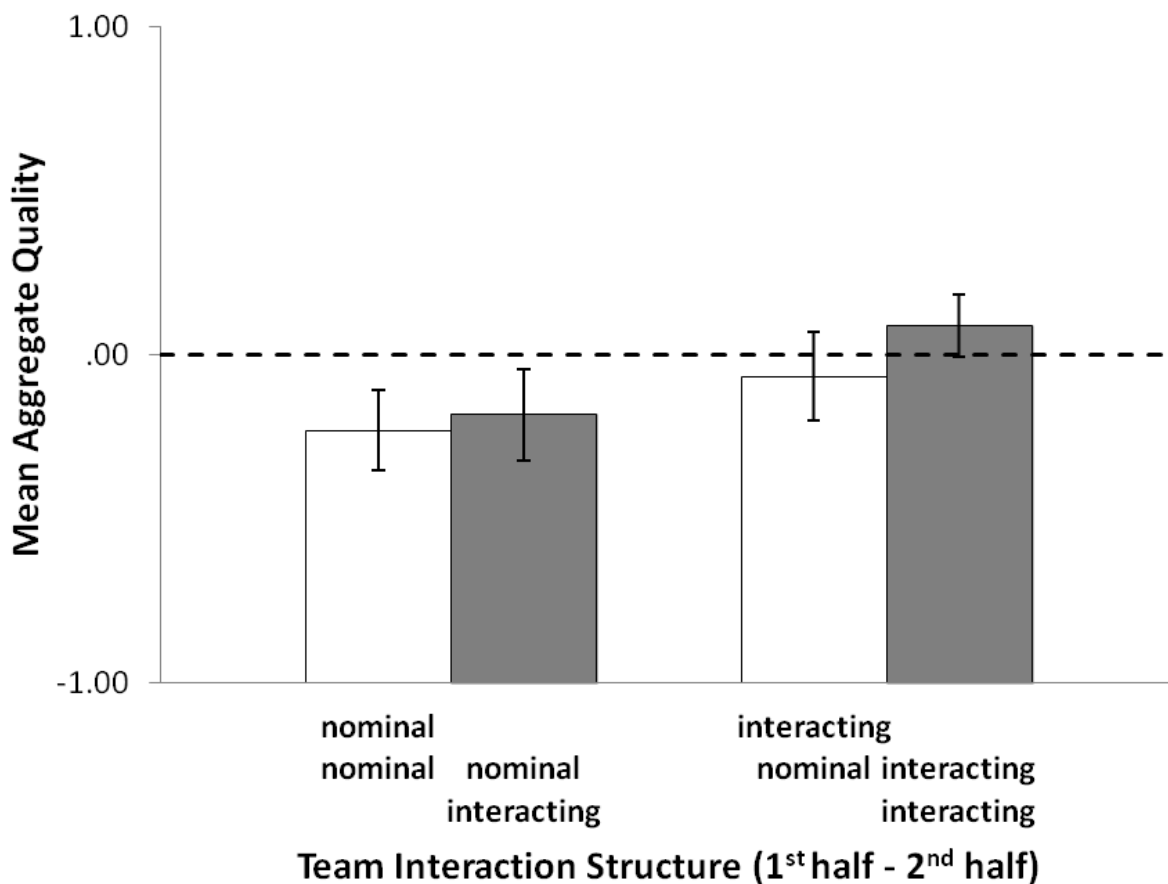
Figure 3.5 Experiment 1A Aggregate Quality Ratings.

Bar grouping indicates 1st half team interaction structure (nominal left, interacting right), bar color indicates 2nd half team interaction structure (nominal white, interacting grey). Error bars indicate 1SE.

39

3.655, $p = 0.001$, $\eta^2 = 0.383$. Discriminant analysis was used to detect multivariate simple

effects as a function of team interaction structure. First, a function was constructed to

discriminate teams on the basis of 1st half structure to provide a basis of comparison for further

discriminant analyses that check for differences in 1st half structure, holding each level of 2nd half

structure constant. Consistent with the MANOVA, a function was able to successfully

Table 3.3 Correlations Between Discriminant Function and Criteria for Experiment 1A.

Negative function values indicate nominal team structure, positive values indicate interacting team structure. Correlations shown for functions that discriminate on the basis of 1st half team structure overall, as well as 1st half structure when holding 2nd half structure constant at one of its levels (nominal or interacting).

	r		
Criterion	*1st half*	*1st half at 2nd half nominal*	*1st half at 2nd half interacting*
Cost	0.316	0.363	0.133
Removes shell w/o damaging peanut	-0.018	-0.250	0.054
Separates shell & nut	0.558	0.631	0.268
Amt. of peanuts shelled per hr	0.023	-0.066	0.035
Energy source required to operate	-0.096	0.263	-0.152
Amt. of energy required	0.086	0.219	0.012
Feasibility	0.173	0.333	0.046
Size/Portability	-0.088	0.285	-0.159
# of people needed to operate	0.170	-0.315	0.172
Time to set up & build	0.117	-0.023	0.104

differentiate between 1st half interacting and 1st half nominal teams, canonical $R^2 = 0.356$, $\Lambda = 0.644$, $X^2(10) = 28.645$, $p < 0.001$. Correlations between dependent variables and the discriminant function show that teams which interacted in the first part of the experiment produced designs that not only were of higher aggregate quality, but more specifically cost less than those generated by 1st half nominal teams. In addition, 1st half interacting teams, compared to 1st half nominal teams, also produced designs that that were better able to separate shells from nut fruit once the shell was cracked (Table 3.3 & Table 3.4).

Table 3.4 Discriminant Functions with Standardized Coefficients for Experiment 1A.

Negative function values indicate nominal team structure, positive values indicate interacting team structure. Correlations shown for functions that discriminate on the basis of 1st half team structure overall, as well as 1st half structure when holding 2nd half structure constant at one of its levels (nominal or interacting).

Criterion	β		
	1st half	1st half at 2nd half nominal	1st half at 2nd half interacting
Cost	2.339	0.511	4.820
Removes shell w/o damaging peanut	-0.352	-0.475	-1.000
Separates shell & nut	0.672	0.669	-1.609
Amt. of peanuts shelled per hr	0.222	0.328	3.449
Energy source required to operate	-0.889	-0.596	1.983
Amt. of energy required	1.078	0.925	-0.048
Feasibility	-1.653	-0.053	-2.346
Size/Portability	-0.718	0.063	-4.488
# of people needed to operate	-0.675	-0.778	---
Time to set up & build	0.294	-0.205	4.044

Next, discriminant functions were constructed to detect differences in 1st half structure, holding each level of 2nd half structure constant. A function could be constructed to discriminate nominal-interacting and interacting-interacting teams from each other (2nd half interacting), canonical $R^2 = 0.818$, $\Lambda = 0.188$, $X^2(10) = 49.303$, $p < 0.001$. However, a function could not be constructed to discriminate nominal-nominal and interacting-nominal teams (2nd half nominal), canonical $R^2 = 0.213$, $\Lambda = 0.787$, $X^2(10) = 6.942$, $p = 0.731$. This suggests that 2nd half interacting teams only saw a performance benefit if they also interacted in the 1st half, explaining the MANOVA interaction reported earlier.

3.2.3 Discussion

Support was found for H_1. Interacting teams tended to produce higher team coherence values than nominal teams, and this effect was larger for 2nd half structure than 1st half structure. Members of interacting teams also tended to have lower similarity-to-solution scores than nominal team members; this was especially true for members of nominal-nominal teams. It should be noted that 2nd half interaction structure was conflated with the similarity-to-solution LSA measure, explaining in part the effect of 2nd half interaction structure on similarity-to-solution. 2nd half interacting team members had their individual self-reports compared to a common team design solution to produce the similarity-to-solution metric, whereas 2nd half nominal team members had their individual self-reports compared to their individual design solution. The timing of interaction also had an impact on the quality of design solutions. Whereas effects of team coherence were more pronounced as a function of 2nd half team interaction condition, it was 1st half interaction which produced differences in quality across groups. 1st half interacting teams had higher quality solutions than 1st half nominal teams, and

this suggests that early interaction is more important than later interaction in determining solution quality. The statistical interaction between 1st half and 2nd half team interaction structure suggests that more interaction is better for design solution quality.

The similarity-to-solution data may provide insight into why interaction was associated with higher quality solutions. Nominal-nominal individuals had solutions that were very similar to their final design solutions as measured by LSA, suggesting that what these individuals believed to be their best solution or approach to the design problem did not change much over the course of the experiment. In addition, they produced some of the lowest quality solutions. It may be that interaction prevents fixation on low quality solutions, perhaps by exposing the individual to alternative approaches or solutions that they may not have considered otherwise.

3.3 Experiment 1B

In order to replicate and extend the results of Experiment 1A, the study was replicated using a more general sample of undergraduate students from introductory psychology classes. Two additional sets of analyses were also conducted. First, the originality of proposed final problem solutions were evaluated to more closely link these results with those of the brainstorming literature. In addition, written self-reports were analyzed to develop a better understanding of the number of alternatives that team members were considering throughout the problem solving process to identify behavior consistent with that proposed to induce representation construction (Figure 1.1).

3.3.1 Method

     3.3.1.1 Participants. Participants were n = 83 undergraduate students from introductory psychology courses. Volunteers received credit towards a class requirement for participation in the study. Volunteers signed up using an online research participation client.

     3.3.1.2 Design. This experiment used a design analogous to that of Experiment 1A. Three-member teams were randomly assigned by block to condition, with one team in each experimental session. This experiment utilized the same 2x2 between-groups factorial design as Experiment 1A, where team members worked together (interacting) or independently (nominal) during the 1st half and/or 2nd half of the experiment.

     3.3.1.3 Materials. Teams in this study were asked to solve a campus improvement problem. Specifically, they were asked to improve the availability of parking on campus (Appendix D). Campus improvement problems have been used successfully in past problem solving research with university populations (e.g., Mumford, et al., 2001), because students have familiarity with these problems by virtue of being a member of the university community. The Parking Problem was used instead of the Peanut Sheller problem in order to match the problem to the expertise of the participant sample. All students have some experience with the university's parking system (e.g., driving oneself, parents/friends who are visiting), and/or with alternatives to driving to campus (e.g., public bus, biking, walking). All other instructions and materials were similar to those used in Experiment 1A.

     3.3.1.4 Procedure. The procedure was identical to Experiment 1A, with two exceptions. First, each session was run with a maximum of one team at a time, and therefore each team was randomly assigned to condition by block. Second, teams were assigned to an extra-nominal condition in the event that one or more of the three volunteers did not appear for their assigned

session similar to Experiment 1A. The result was n = 24 teams split across each of the four cells of the main experimental design, with n = 11 individuals assigned to the extra-nominal condition. Similar to Experiment 1A, data from extra-nominal participants was used to train the LSA model and calibrate raters, but was not used in any inferential analyses.

3.3.2 Results

Experiment 2 produced a set of self-reports over time for each team, and one or a set of design solutions for each team, similar to Experiment 1A. Self-report data was analyzed using the same LSA procedure used in Experiment 1A (see Section 3.2.2). The quality of final problem solutions were evaluated by three trained undergraduate raters, with rating scales that were developed in a way that emphasized relative comparisons across the set of all possible solutions in the dataset for each criterion (see Section 3.2.2.3). LSA metrics and design quality were subsequently compared across conditions using analysis of variance. Self-report data was analyzed using the same LSA procedures as used in Experiment 1A. An LSA space was trained on all self-reports as well as transcriptions of final problem solutions. The number of SVD dimensions was set to the number of documents (n = 656).

3.3.2.1 Team coherence. As in Experiment 1A, team coherence scores were input into a 2x2 repeated-measures factorial ANOVA, with team interaction structure in the 1st and 2nd half of the experiment as between-groups factors. The analysis identified a main effect of 1st half structure, with 1st half interacting teams showing higher levels of team coherence on average over the course of the experiment ($M = 0.450$, $SD = 0.054$) than 1st half nominal teams ($M = 0.394$, $SD = 0.054$), $F(1,20) = 13.124$, $p = 0.002$, $\eta^2 = 0.396$. A main effect of 2nd half structure was identified as well, $F(1,20) = 6.126$, $p = 0.022$, $\eta^2 = 0.234$, with 2nd half interacting teams

again showing greater coherence over the experiment ($M = 0.441$, $SD = 0.054$) than 2nd half nominal teams ($M = 0.403$, $SD = 0.054$). The ANOVA also identified a Time by 1st half team interaction structure statistical interaction, $F(4,80) = 3.673$, $p = 0.008$, $\eta^2 = 0.155$, and a Time by

Figure 3.6 Experiment 1B Team Coherence Over Time.

Each line represents a different team interaction structure condition. White markers indicate times at which teams were interacting for the previous work epoch; black markers indicate a nominal team interaction structure for the prior work epoch. Line style indicates 1st half interacting (solid) or nominal (dashed) team structure. Line weight indicates whether 1st & 2nd half structure were the same (heavy) or different (light). Error bars indicate 1SE.
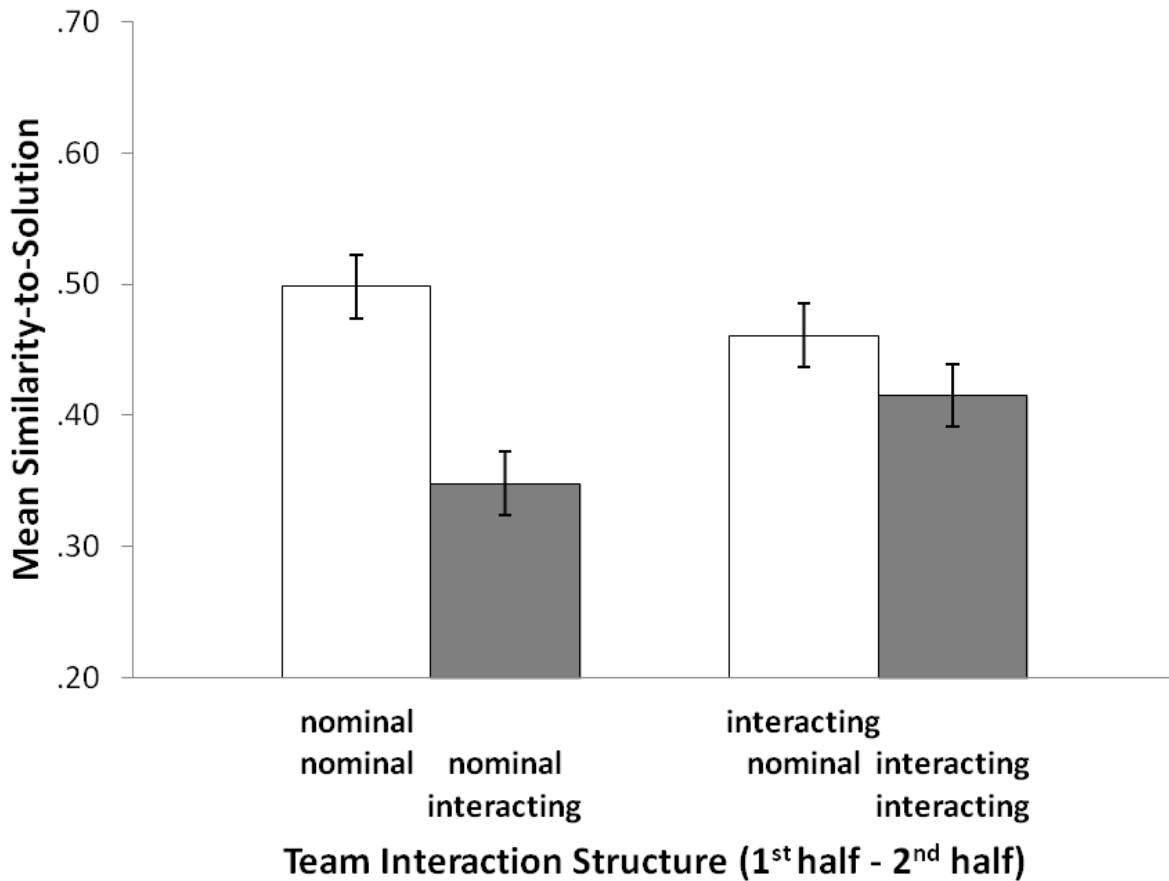
Figure 3.7 Experiment 1B Mean Similarity-to-Solution.

Bar grouping indicates 1st half team interaction structure (nominal left, interacting right), bar color indicates 2nd half team interaction structure (nominal white, interacting grey). Error bars indicate 1SE.

2nd half interaction structure statistical interaction, $F(4,80) = 5.353$, $p = 0.001$, $\eta^2 = 0.211$.

Similar to Experiment 1A, team coherence rose when teams worked together on a problem, and stayed low or fell when team members worked independently (Figure 3.6). Unlike Experiment 1A, a three way Time by 1st half by 2nd half team interaction structure statistical interaction was not detected. In conclusion, support was found for H$_1$ in this sample.

3.3.2.2 Similarity-to-solution. Similarity-to-solution scores were also analyzed with a 2x2 repeated-measures factorial ANOVA, consistent with Experiment 1A. This analysis identified a main effect of 2nd half interaction structure, $F(1,68) = 16.986$, $p < 0.001$, $\eta^2 = 0.200$, qualified by a statistical interaction between 1st half and 2nd half interaction structure, $F(1,68) = 4.836$, $p < 0.031$, $\eta^2 = 0.066$. Members of 2nd half nominal teams ($M = 0.480$, $SD = 0.017$) had higher similarity-to-solution scores on average over the course of the experiment than members of 2nd half interacting teams ($M = 0.382$, $SD = 0.017$), though as noted with Experiment 1A (Section 3.2.2.2) this 2nd half main effect may be more of a function of how similarity-to-solution is computed rather than an effect of working together or alone in the 2nd half of the experiment per se. Teams whose members interacted in the 1st half had similarity-to-solution scores that were more similar to each other on average than members of 1st half nominal teams. In addition, nominal-nominal teams had the highest scores on average of the four conditions ($M = 0.498$, $SD = 0.024$), while nominal-interacting teams had the lowest ($M = 0.348$, $SD = 0.024$; Figure 3.7).

In addition to the above between-subjects effects, the analysis identified a main effect of Time, $F(4,272) = 10.165$, $p < 0.001$, $\eta^2 = 0.130$, and a statistical interaction of Time by 1st half team interaction structure, $F(4,272) = 2.827$, $p = 0.025$, $\eta^2 = 0.040$. Just as in Experiment 1A, similarity-to-solution scores tended to increase over time. The Time by 1st half team interaction structure statistical interaction is consistent with the statistical interaction on average similarity-to-solution scores reported earlier. 1st half nominal team members had a greater discrepancy in their similarity-to-solution scores than 1st half interacting team members. Nominal-nominal individuals had the highest similarity-to-solution scores over the 1st half of the experiment, nominal-interacting individuals had the lowest scores, while 1st half interacting individuals had similarity-to-solution scores over the 1st half that fell in between the scores of nominal-
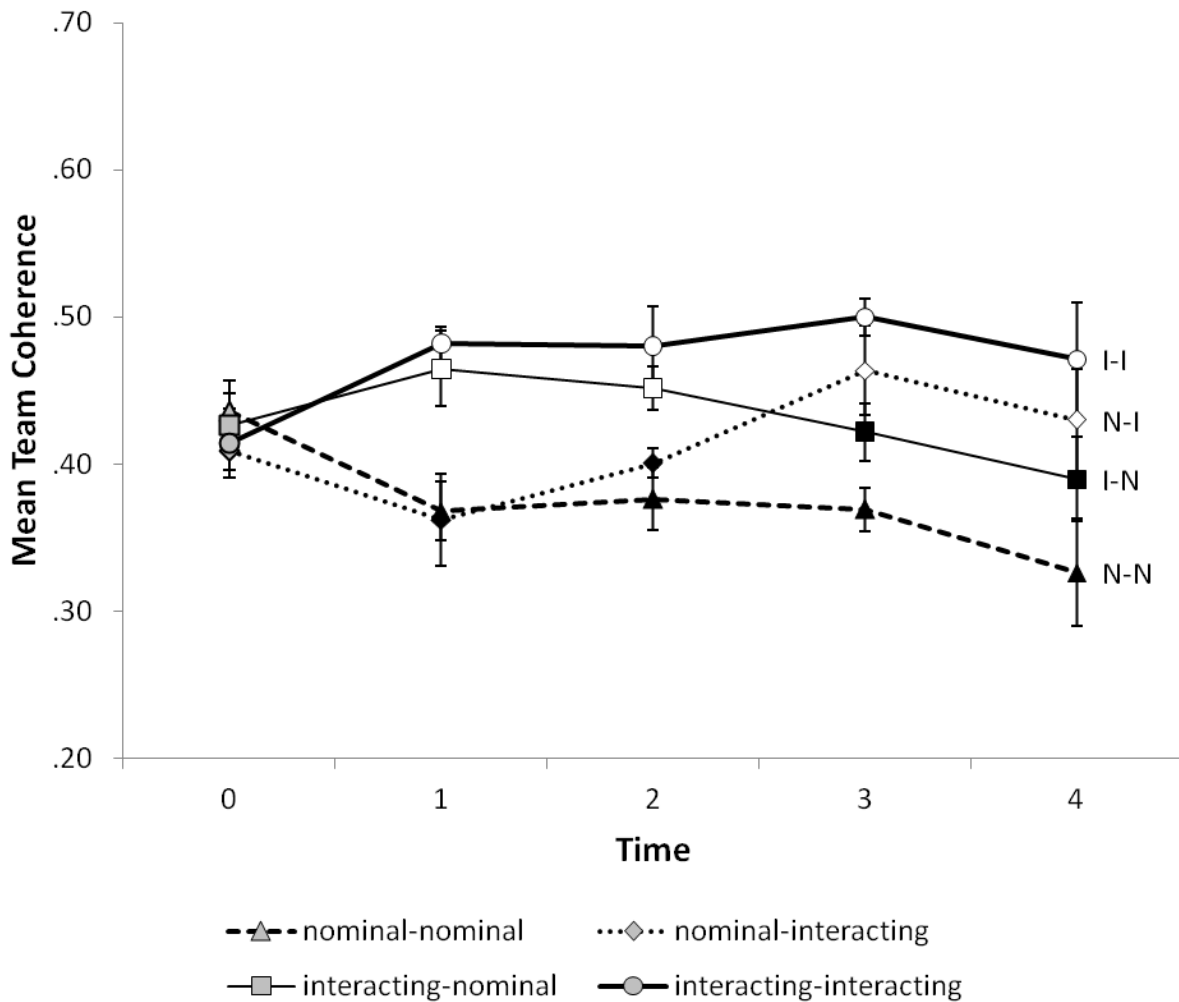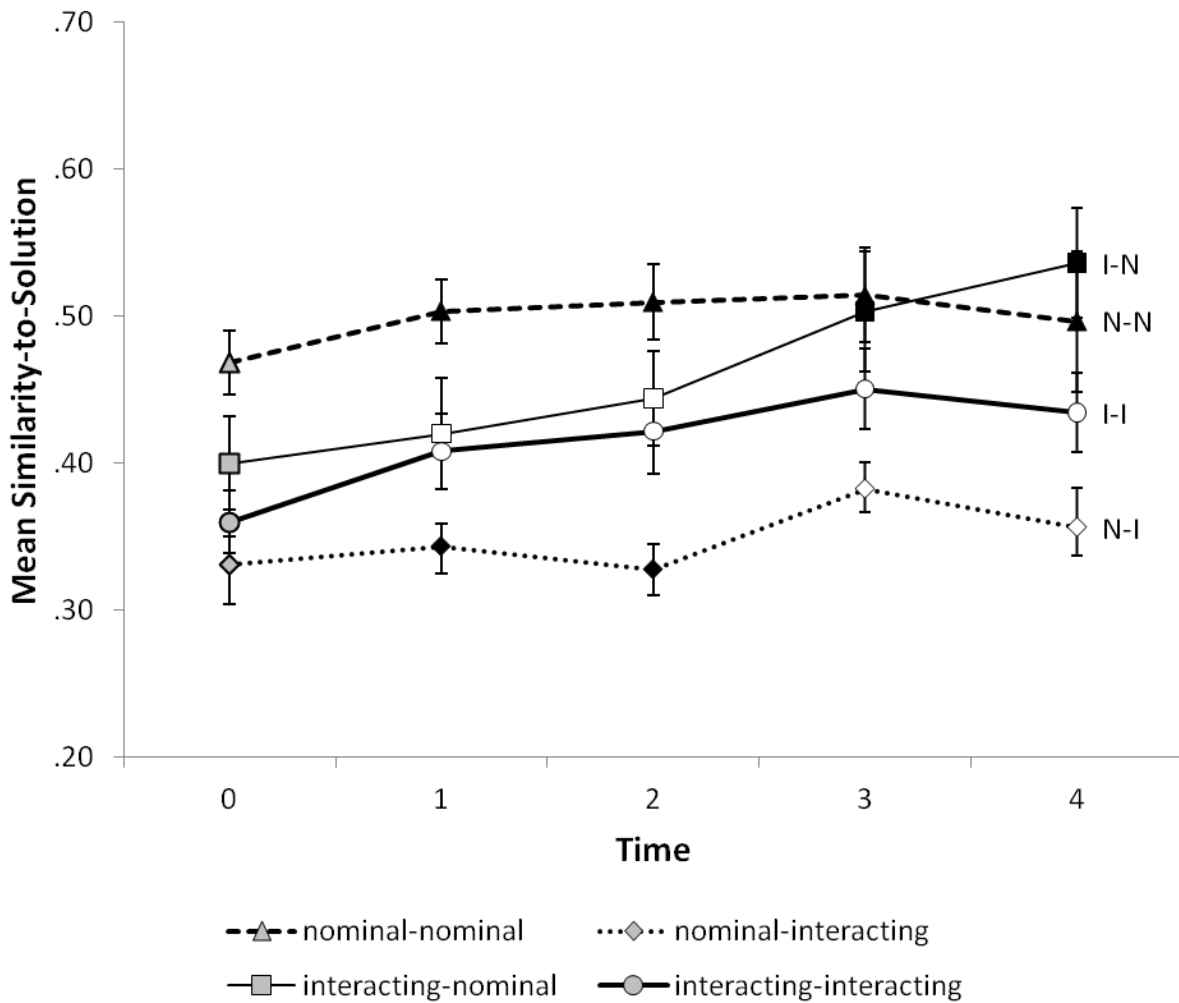
Figure 3.8 Experiment 1B Similarity-to-Solution Over Time.

Each line represents a different team interaction structure condition. White markers indicate times at which teams were interacting for the previous work epoch; black markers indicate a nominal team interaction structure for the prior work epoch. Line style indicates 1st half interacting (solid) or nominal (dashed) team structure. Line weight indicates whether 1st & 2nd half structure were the same (heavy) or different (light). Error bars indicate 1SE.

interacting and nominal-nominal individuals. In the 2nd half of the experiment, similarity-to-solution scores for interacting-nominal individuals trended toward those of nominal-nominal individuals while 2nd half interacting teams scores stayed relatively flat (Figure 3.8). No other

statistical interactions with respect to Time or team interaction structure condition were

identified.

   3.3.2.3 Final problem solution quality assessment. To assess the quality of each team's

final problem solution, trained raters scored each team's solutions relative to the rest of the

sample using a variation of Amabile's Consensual Assessment Technique (1996; Hennessey &

Amabile, 1988) employed by Mumford and colleagues (e.g., Lonergan, Scott, & Mumford,

2004; Redmond, et al., 1993) in order to develop a decision matrix based off a Pugh chart. First,

all the generated solutions were reviewed to gain a sense of the ways in which solutions varied

from each other. Then, using the University master plan and state transportation agency

guidance, performance criteria on which solutions in the dataset varied were identified. Once

Table 3.5 Intraclass Correlation Coefficients on Training Set for Experiment 1B.

Criterion	α
Overall Quality	0.866
Parking availability	0.830
Parking need	0.787
Cost to university	0.875
Space requirements	0.830
Environmental impact	0.684
Traffic impact	0.751
Aesthetics	0.819
Fairness to campus groups	0.771
Community satisfaction	0.747
TOTAL	0.811

Table 3.6 Intraclass Correlation Coefficient on Full Dataset for Experiment 1B.

Criterion	α
Quality	0.672
Parking availability	0.809
Parking need	0.703
Cost to university	0.863
Space requirements	0.739
Environmental impact	0.182
Traffic impact	0.557
Aesthetics	0.664
Fairness to campus groups	0.410
Community satisfaction	0.628
TOTAL	0.696
TOTAL after removing: - environmental impact - traffic impact - fairness to campus groups	0.763

these criteria were identified, the datum for each criterion was set to a prototypic average

solution for that criterion. This was done because the variety of solutions was considerably less

than those found among Peanut Sheller solutions, and as a result no solution existed which

represented the average response on all dimensions. After identifying the datum solution, a

benchmark exemplar was enumerated for each level of the decision matrix for each criterion to

accommodate the range of performance on each criterion in the dataset, just as in Experiment 1A

(Appendix E).

Three undergraduate raters were trained by reviewing the solution set and the decision matrix scales as part of a one-hour classroom session. Interrater agreement was then evaluated by asking raters to assess a random selection of 25% of final problem solutions, using ICC with relative agreement like Experiment 1A, but with a more typical training threshold of $\alpha = 0.7$.
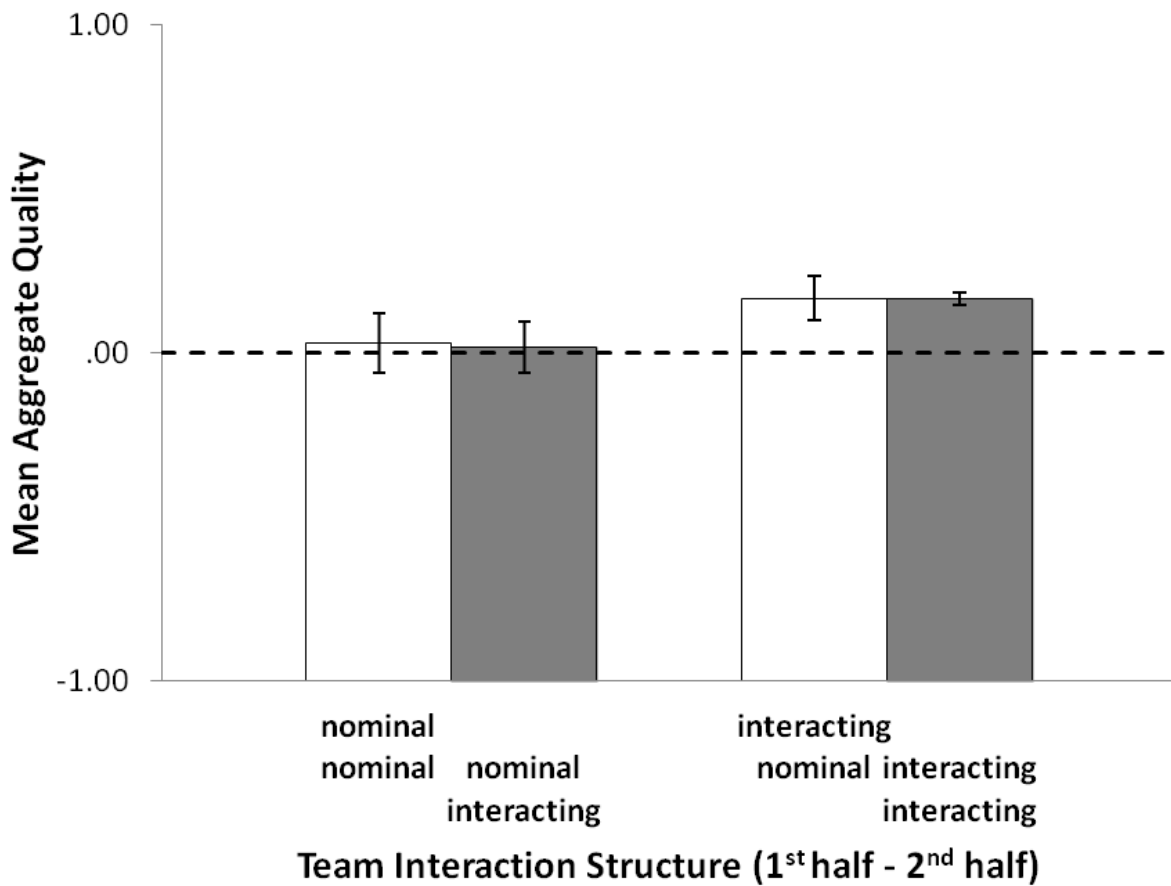
Figure 3.9 Experiment 1B Aggregate Quality Ratings.

Bar grouping indicates 1st half team interaction structure (nominal left, interacting right), bar color indicates 2nd half team interaction structure (nominal white, interacting grey). Error bars indicate 1SE.

Interrater agreement was quite high across all dimensions ($\alpha = 0.811$), and was only below criterion on one dimension (environmental impact, $\alpha = 0.684$; Table 3.5).

Raters then proceeded to evaluate the entire dataset. In order to develop a final quality metric for each final problem solution, ratings were averaged across raters and across criteria for those criteria where reliability was greater than $\alpha = 0.6$ across raters in the full dataset. Three criteria were dropped as a result (environmental impact, traffic impact, fairness to campus groups; Table 3.6). Mean quality was entered into a 2x2 ANOVA, with 1st half and 2nd half team interaction structure as between-subjects factors. As was found in Experiment 1A, a main effect of 1st half team interaction structure was identified, $F(1,20) = 4.119$, $p = 0.046$, $\eta^2 = 0.057$. Interacting teams in the 1st half of the experiment had higher quality solutions ($M = 0.167$, $SD = 0.207$) than nominal teams ($M = 0.0.025$, $SD = 0.357$). The analysis did not identify either a main effect of 2nd half team interaction structure, or a statistical interaction between 1st half and 2nd half team interaction structure (Figure 3.9). This analysis provides support for H_1.

As with Experiment 1A, a 2x2 between-subjects MANOVA was conducted in order to identify specific criteria driving results. This analysis did not identify a main effect of 1st half team interaction structure or a 1st half by 2nd half team interaction structure statistical interaction, but did identify a main effect of 2nd half team interaction structure, $F(7,62) = 2.290$, $p = 0.038$, $\eta^2 = 0.205$. A discriminant function was constructed to discriminate on the basis of 2nd half team interaction structure in order to identify the dependent variables which most influenced the earlier MANOVA analysis. The function successfully differentiated between 2nd half nominal and 2nd half interacting teams, canonical $R^2 = 0.451$, $\Lambda = 0.797$, $X^2(7) = 15.087$, $p = 0.035$. This result appears to contradict the earlier ANOVA result, and also H_2. A closer look at the discriminant function shows that the MANOVA identified common properties of designs for 2nd

Table 3.7 Standardized Discriminant Function Coefficients and Correlations Between Discriminant Function Coefficients and Outcome for Experiment 1B.

Negative discriminant function values indicate "nominal" interaction structure in 2ⁿᵈ half, positive values indicate "interacting."

Criterion	β	r
Quality	-0.898	-0.125
Parking availability	1.501	0.571
Parking need	0.043	-0.210
Cost to university	0.906	-0.227
Space requirements	0.140	-0.358
Aesthetics	0.127	0.082
Community satisfaction	0.626	0.455

half interacting teams that were also frequently present in the dataset in general. Correlations between dependent variables and the discriminant function show that teams which interacted in the 2ⁿᵈ half of the experiment produced solutions that provided more parking spaces for vehicles and made the community happy, but also required more physical space on campus (Table 3.7). Qualitatively, solutions which included building or expanding parking garages represented a large minority of the dataset. Many of these solutions appear to have come from 2ⁿᵈ half interacting teams, and the particular conceptual solutions these teams proposed appear to be better than similar solutions proposed by members of 2ⁿᵈ half nominal teams.

3.3.2.4 Final problem solution originality assessment. In order to explore the impact of team interaction structure on solution originality, raters were also asked to score designs on the basis of their originality. The decision matrix scale for evaluating originality (Appendix F) was developed and evaluated for interrater reliability in the same manner as quality metrics for this

dataset. Interrater reliability on originality for both the training set (α = 0.802) and the full

dataset (α = 0.696) was acceptable, so originality scores were averaged across raters to produce a

single originality score for each final problem solution. These scores were then entered into a

2x2 between-subjects ANOVA with 1st and 2nd half team interaction structure entered as factors.
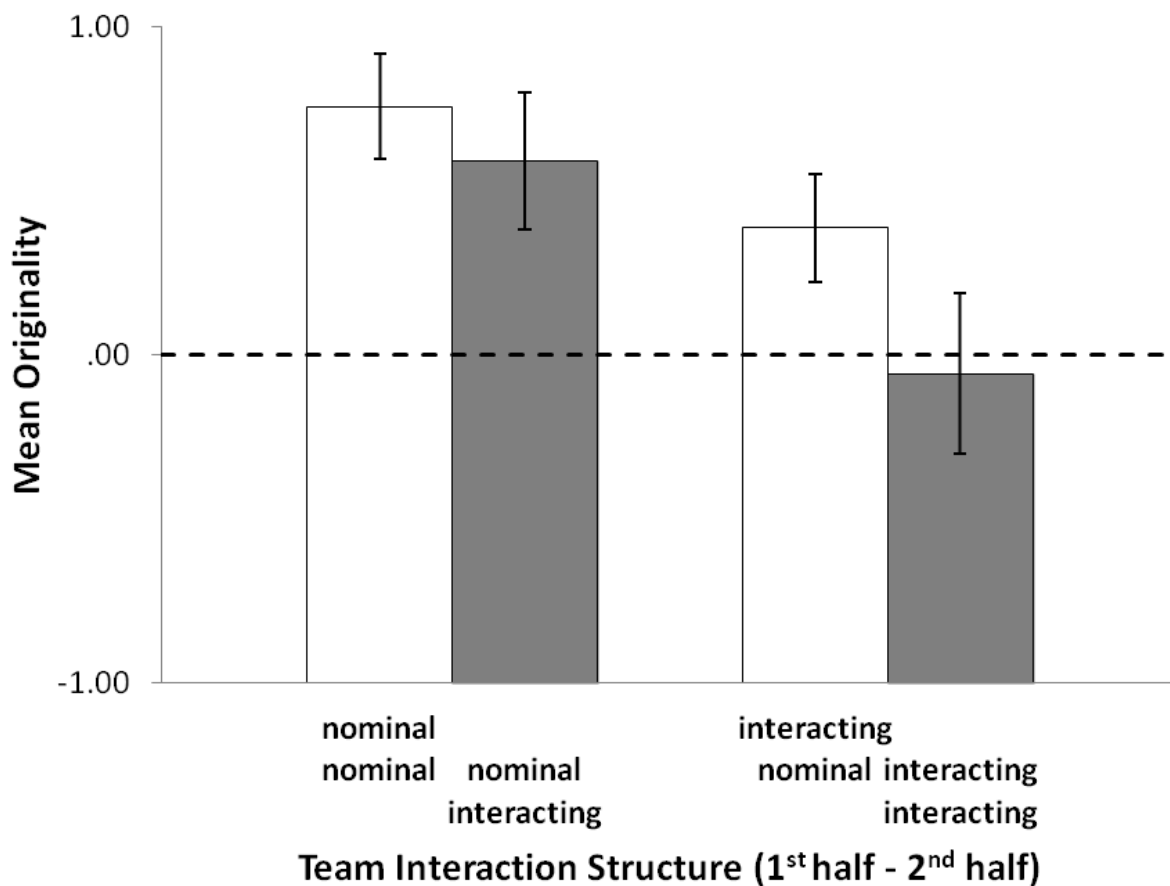
Figure 3.10 Experiment 1B Mean Originality Ratings.

Bar grouping indicates 1st half team interaction structure (nominal left, interacting right), bar color indicates 2nd half team interaction structure (nominal white, interacting grey). Error bars indicate 1SE.

This analysis identified a main effect of 1st half team interaction structure, $F(1,68) = 6.617$, $p = 0.012$, $\eta^2 = 0.089$. Members of 1st half nominal teams had final problem solutions of higher originality ($M = 0.676$, $SD = 0.787$) than 1st half interacting teams ($M = 0.167$, $SD = 0.900$). A main effect of 2nd half team interaction structure was not detected, and there was no statistical interaction between 1st half and 2nd half interaction structure with respect to originality (Figure 3.10). Qualitatively, it appears that 1st half nominal teams tended to have greater diversity in the types of solutions they proposed compared to 1st half interacting teams. This greater diversity of solution types in turn increased the likelihood that the type of solution proposed was unique (i.e., more original) compared to others in the dataset. This originality effect may also be illustrative of an evaluation apprehension effect whereby interacting team members propose more conventional solutions concerned that more novel suggestions may be ridiculed, and then once the team negotiates and agrees on the basic structure of a solution, team members spend the rest of their efforts detailing that solution, continuing with this representation in the 2nd half of the experiment even if they transition to working independently. In contrast, members of nominal teams do not need to negotiate their solutions with other team members, and may be more likely not only to propose more novel solution configurations, but also to stick with these less conventional solutions.

3.3.2.5 Analysis of self-reports. Anecdotal evidence from Experiment 1B suggested that early team interaction may promote representation construction by encouraging individual team members to recognize and discuss differences in the solutions they think of after reading the problem, leading to a discussion of what each team member was trying to accomplish with their solution. This process is similar to that advocated by researchers of the representation construction process in individuals, except that individuals are encouraged to develop problem

56

restatement alternatives rather than solution alternatives through instructions to restate a problem in many different ways. These restatements help the solver to identify instances where different rewordings of the question suggest different goals for solving (Section 2.6).

In order to identify more concrete evidence behavior that may facilitate representation construction, self-reports were coded by trained raters for the number of distinct goals and solutions contained in each. These self-reports were elicited by asking team members to describe their "best solution or approach" to the problem (Appendix B), and the resulting reports are posited to describe only solutions that are under serious consideration. It is hypothesized, consistent with H_1, that members of interacting teams will describe more solutions in their individual self-reports of the problem solving process, compared with members of nominal teams.

Three trained undergraduate raters evaluated self-reports for the number of distinct solutions contained within each. These raters were familiarized with the definitions for a goal or

Table 3.8 Solution Definition, Example, and Intraclass Correlations for Training Set and Full Dataset.

			α	
Criterion	*definition*	*example*	*training set*	*full set*
Solution	A distinct full or partial answer to the problem is elaborated.	*"Make the UC garage valet to fit more cars."*	0.864	0.849
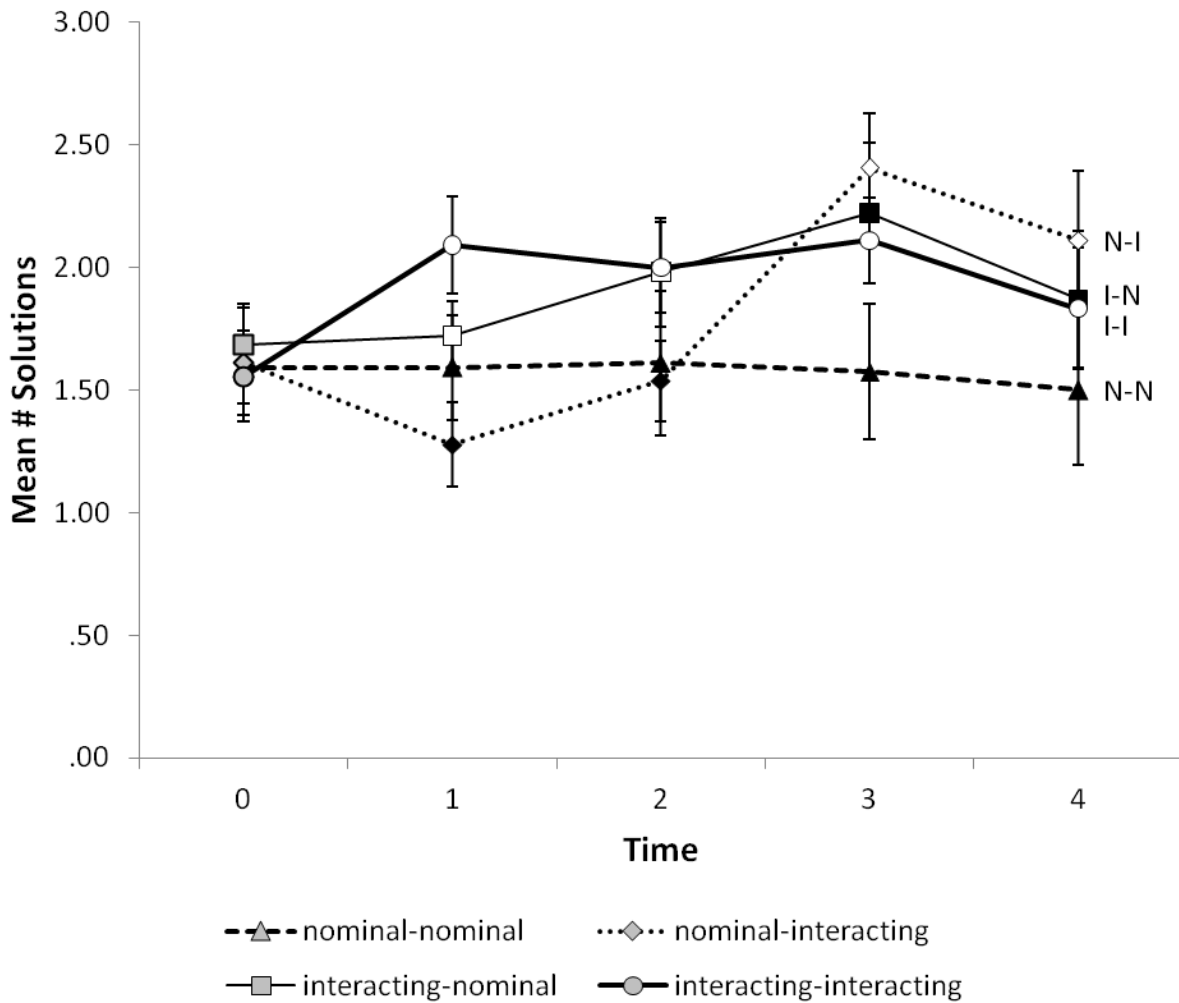
Figure 3.11 Experiment 1B Number of Solutions Over Time.

Each line represents a different team interaction structure condition. White markers indicate times at which teams were interacting for the previous work epoch; black markers indicate a nominal team interaction structure for the prior work epoch. Line style indicates 1st half interacting (solid) or nominal (dashed) team structure. Line weight indicates whether 1st & 2nd half structure were the same (heavy) or different (light). Error bars indicate 1SE.

solution (Table 3.8). Raters were then asked to evaluate a training set of self-reports from a randomly selected team in each condition of the design from Experiment 1B, including an extra-nominal team. This represented about 20% of the self-report data. Ratings for number of solutions in the training set were above the training criteria of $\alpha = 0.7$, and raters proceeded to

assess the full set of self-reports. The relaxed criterion of $\alpha = 0.6$ was used for agreement on the full dataset, consistent with rater assessments of solution quality for this data.

In order to evaluate the effect of team interaction structure on representation construction behavior, in this case the number of explicitly stated solutions that were identified which may have helped team members to evaluate how they were representing the problem, mean rater counts of solutions from self-reports were entered into a repeated-measures factorial ANOVA, where the two team interaction structure factors (1st half, 2nd half) were entered as between-subjects effects, and Time entered as a within-subjects effect. The ANOVA identified a significant main effect of Time on the number of solutions mentioned in self-reports, $F(4,272) = 3.830$, $p = 0.005$, $\eta^2 = 0.053$. In general, the number of solutions mentioned by participants increased over the course of the experiment (Figure 3.11).

In addition, a Time by 1st half by 2nd half statistical interaction was identified, $F(4,272) = 2.975$, $p = 0.020$, $\eta^2 = 0.042$. Members of 1st half interacting teams tended to increase the number of solutions they mentioned between t_0 and t_2, whereas 1st half nominal team members had the same number of distinct solutions in their self-reports from t_0 to t_2. Over the 2nd half of the experiment, members of teams who interacted in the 1st half tended to maintain a high level of distinct solutions in their self-reports, while the number of solutions mentioned in the 2nd half of the experiment by members of 1st half nominal teams was dependent on 2nd half condition. Nominal-nominal team members maintained a low number of solutions in their self-reports over the 2nd half of the experiment, while nominal-interacting team members had an increase in the number of solutions in each of their self-reports from t_2 to t_3. Post-hoc contrast with Sidak correction comparing nominal-nominal and nominal-interacting teams show a statistically significant difference at t_3, $F(1,68) = 5.807$, $p = 0.019$, $\eta^2 = 0.079$, suggesting that nominal teams

began to consider alternative solutions as soon as they started working with their teammates. A similar analysis between nominal-nominal and interacting-nominal teams shows a marginally significant difference at t_3, indicating that teams that interacted in the 1st half continued to consider more alternatives even after beginning to work by themselves, $F(1,68) = 3.513$, $p = 0.065$, $\eta^2 = 0.049$. No other main effects or interactions were identified (Figure 3.12).
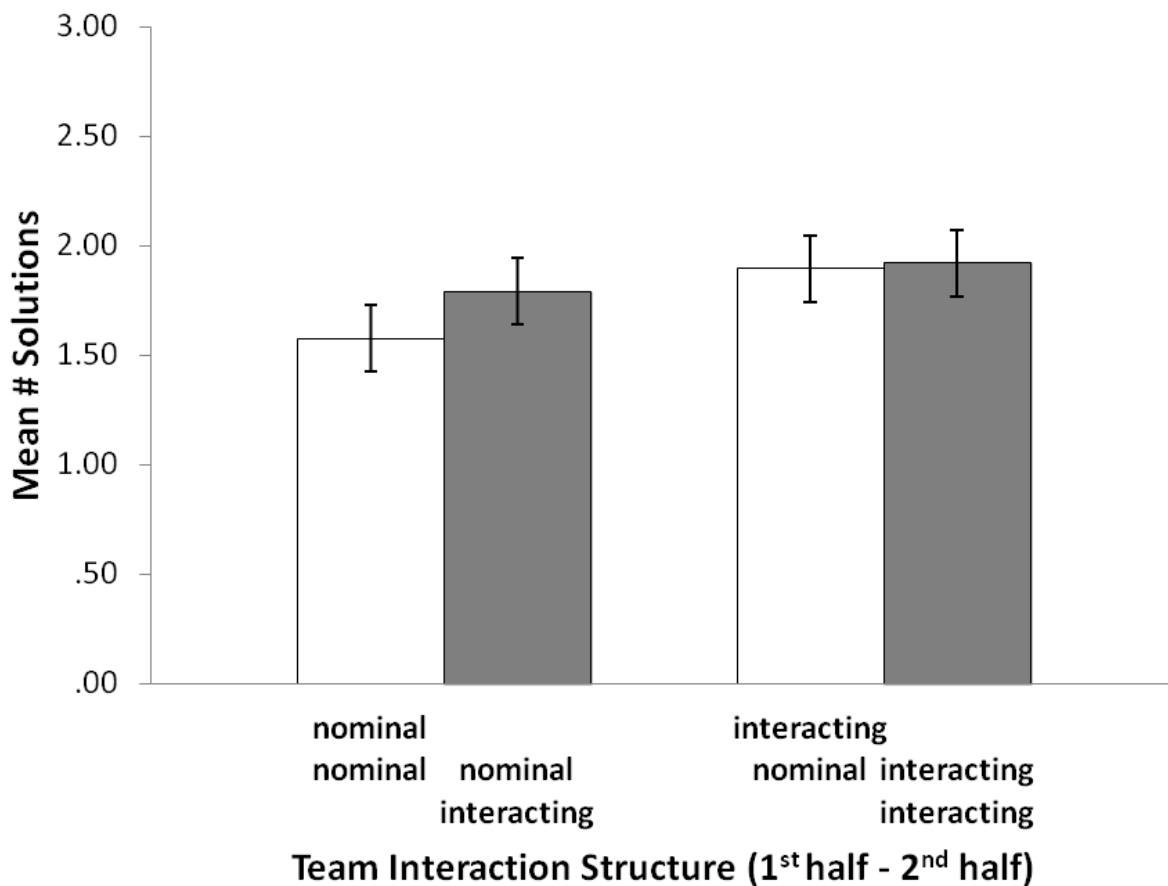
Figure 3.12 Experiment 1B Mean Number of Solutions.

Bar grouping indicates 1st half team interaction structure (nominal left, interacting right), bar color indicates 2nd half team interaction structure (nominal white, interacting grey). Error bars indicate 1SE.

3.4 Discussion

Experiment 1B provided support for H$_1$. Early interaction produced greater team coherence. In addition, early interaction was associated with higher quality, but less original solutions. Similarity-to-solution scores suggest that, similar to Experiment 1A, members of nominal-nominal teams were more likely than those of other teams to fixate on a solution or approach to the problem that they have identified early on in the course of solving, although interacting-nominal individuals showed similar levels of commitment to their solutions as nominal-nominal individuals by the end of the experiment. This is consistent with the idea that interaction prevents fixation.

The analysis of self-report protocols for the number of distinct solutions provides additional evidence consistent with H$_1$. It shows that the number distinct solutions tended to increase over time for all but nominal-nominal team members, and further varied depending on experimental condition. Changes in the number of solutions contained in written protocols over time are consistent with H$_1$ on the role of team interaction on the development of several solutions to a problem. The increased number of solutions in self-reports for interacting teams is consistent with the proposition that the solutions generated by team members act as a cue to help teams consider the different ways in which a problem could be represented (Figure 1.1).

A closer look at the LSA data from Experiment 1B along with solution counts suggest that nominal-nominal team members may have become fixated on problem solutions which they identified very early on in the solving process. The individuals in nominal-nominal teams showed very little variance over time in the number of solutions they mention in self-reports. Members of nominal-nominal teams also had among the highest LSA similarity-to-solution values, and additionally these values fluctuate little over time when compared to other conditions

61

(Figure 3.11). These pieces of data suggest that not only were members of nominal-nominal teams focused on a very narrow representation of the problem, but also that these individuals usually entertained only one (or perhaps two) ways to solve the problem at any point in time.

It is interesting to note that for interacting-nominal teams, the increase in the number of solutions during the 1st half of the experiment stayed consistent into the 2nd half of the experiment, even though team members were no longer sharing ideas with each other at this point. Similarity-to-solution data in the previous chapter (Figure 3.8) shows that interacting-nominal teams also tended to write self-reports that were more like their solution upon discontinuing interaction, so it may be that the solutions mentioned across the 2nd half of the experiment are the same as or similar to those they identified with their team in the 1st half of the experiment. Future work should develop methods that leverage either self-report to track these individual solutions as they evolve over time to inform whether these interacting-nominal teams are continuing to produce new solutions, or rather are reiterating the same solutions expressed earlier. Such a metric is introduced as part of the analysis of Experiment 2 (final path, Section 4.2.2.4), where it is used to estimate the amount of time spent by individuals on refining their final problem solutions.

3.5 Summary and Conclusions

Together, these experiments provide a robust replication of the impact of early team interaction on team performance across two distinct samples. Team interaction was shown to increase team coherence during both the 1st and 2nd half of problem solving efforts in both engineering and psychology student samples. However, high coherence was not always directly associated with team performance. It instead appears that when (early versus late) a team

develops coherence is just as important as the degree or amount of coherence a team develops on a common team task mental model. Early team interaction and corresponding early coherence was associated with better final design or problem solutions in both samples. Interaction and corresponding high coherence throughout solving produced an additional benefit for solution quality only among the engineering student sample in Experiment 1A.

Anecdotal evidence from Experiment 1B verbal protocols provides insight on the origins of the benefit from interacting early during the course of solving. Members of 1st half interacting teams, after reading the problem statement and providing a self-report with their first impressions of the problem, often came together in the first work session on the problem and took turns sharing their solutions. Commonly, one participant would open with "What answer did you give?" or "What did you write?" Once a participant provided a solution that was very different from the other members of the team, the team often began to realize that the different solutions were aimed at achieving different goals. The resulting conversation tended to be a negotiation of exactly what goals a solution ought to achieve, followed by work to define a solution that achieved those goals.

The similarity-to-solution data empirically provides corroboration to the notion that early interaction helps individuals to identify solutions or approaches that they may not have considered in isolation. Solutions from individuals in nominal-nominal teams were most similar, compared to individuals from other conditions, to their own self-reports on the best approach or solution to the problem throughout the solving process. This suggests that nominal-nominal individuals tended to start with one or a small collection of solution concepts, and stay with those concepts more than individuals in conditions where interaction with teammates was permitted. In addition, a count of distinct solutions from self-reports shows that nominal-nominal

individuals tended to entertain the fewest number of potential "best" solutions for solving the problem. As such, interaction provides not only a prompt to identify different possible goals for solving, but may also serve to prevent fixation on a single specific goal and/or solution. Youmans (2011) found a similar benefit for interacting compared to nominal teams working on a design task. He found that this effect was due in part to the fact that interacting teams, compared to nominal teams, integrated fewer elements of an example into their solutions that was introduced with the design task. Research by Smith and colleagues (Jansson & Smith, 1991; S. M. Smith & Blankenship, 1991) finds that design or solution examples are a powerful manipulation that often leads individuals to produce solutions with salient features of the presented examples. In the context of the studies in this paper, it may be solution conceptions perseverate when they are identified early in the course of solving, and example paradigms provide some certainty as to the type of solutions that individuals choose as the locus of this perseveration.

The process of enumerating and negotiating different goals for solving has been demonstrated to improve the quality of solutions proposed by individuals in the industrial/organizational psychology literature (Reiter-Palmon & Robinson, 2009). The process is referred to in this line of inquiry as problem construction (Mumford, et al., 1994), and is typically induced by asking individuals to restate the problem in as many different ways as possible (e.g., Baer, 1988). This exercise provides conditions for individuals to discover for themselves that different solving goals can be enumerated from the same problem statement, leading to a process of evaluating goals similar to that seen in 1st half interacting teams in Experiment 1B. Anecdotal evidence from verbal protocols and lower similarity-to-solution scores over time suggests that teams who began working together later in Experiment 1B

(nominal-interacting) also shared their ideas, although they did not tend to discuss the goal of their efforts and instead combined their individual solutions together to produce a final problem solution that usually was a simple summation of their individual solutions. While the specific mechanism differs between individuals and teams, the process of identifying and evaluating goals for solving appears consistent between early team interactions and individuals engaged in activities that promote representation construction. The data from these experiments suggests that the team representation construction process may mediate the relationship between team coherence and solution quality identified in past research (e.g., Dong, et al., 2004), rather than a relationship whereby team coherence leads to solution quality directly (Figure 3.13).

The finding that teams which work together early produce higher quality, but less original solutions has some relation to research in brainstorming on the effects of different team interaction structures on the quality and variety of ideas. While this research finds that interacting teams produce higher quality solutions, the literature suggests that this is likely not a result of teams' ability to identify high-quality solutions for further refinement. Work by Putman
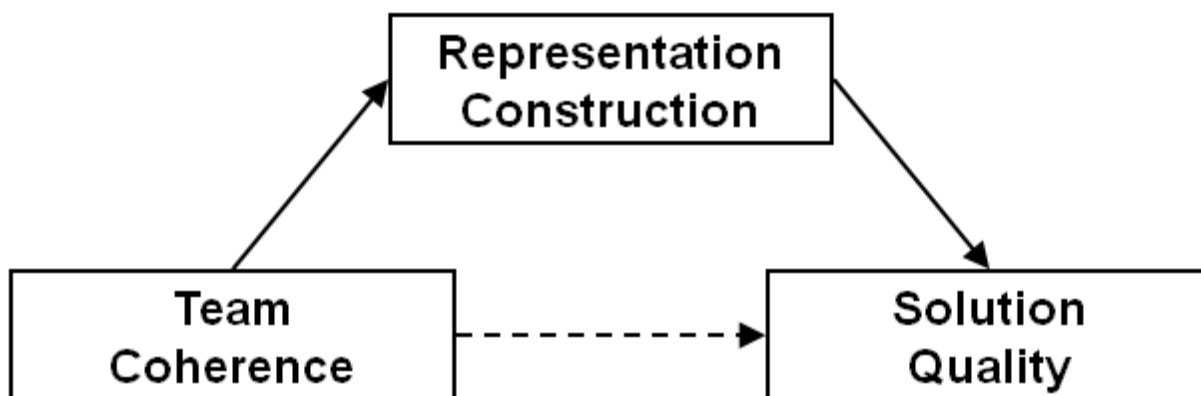
Figure 3.13 Proposed Representation Construction Mediation Model.

and Paulus (2009) and others (Faure, 2004) suggests that both nominal and interacting teams select randomly among ideas when asked to select the highest quality ideas from a pool generated during brainstorming. This result has been shown across a wide range of quality metrics, including feasibility, impact, and others. Rietzschel et al. (2006) finds the inability for individuals to select the best ideas occurs despite the fact that nominal teams tend to produce more high-quality ideas, although the proportion of high-quality ideas to total ideas generated was similar across nominal and interacting teams (see also Faure, 2004; Putman & Paulus, 2009). Since it is unlikely that team members were able to select a high-quality solution early in solving the problem to be refined later, the benefits of early interaction seen in the two experiments outlined above must be a result of developing consensus on other beneficial aspects of the problem very early during the course of solving.

The above experiments find that more original ideas are associated with nominal team interaction structure early in solving, which is consistent with some past brainstorming research. Putman and Paulus (2009) find that interacting teams tend to produce fewer ideas than nominal teams in the context of a brainstorming session, and that nominal teams were better at identifying more original ideas from a pool of brainstormed ideas. Some research on "entrainment" extends these findings and shows that nominal teams maintain a high rate of idea production when brought together for an interacting session (Baruah & Paulus, 2008, 2009). However, work by Paulus and Yang (2000), as well as Leggett et al. (1996) shows that interacting-nominal teams produce more ideas in brainstorming than nominal-interacting teams. The more original solutions produced by 1st half nominal team members in Experiment 1B is consistent with the Putnam and Paulus data for purely nominal or purely interacting teams, but appears to run directly contrary to those of Leggett and others (1996; Paulus & Yang, 2000). One important

distinction between Paulus and colleagues' work from the literature and the above experiments is that brainstorming studies do not often call for participants to refine their brainstormed ideas into a final solution, while the task for participants in the current experiments was to develop a single final problem solution, for which they were free to develop as many or as few alternatives as they wished along the way.

Chapter 4. Representation Construction

4.1 Introduction

Experiments 1A and 1B demonstrated that teams who interact early during problem solving tend to develop higher quality solutions to problems. Though changes in team interaction structure were associated with changes in team coherence, team coherence was not otherwise associated with the quality of final problem solutions. Members of interacting teams did tend to describe a greater number of "best" solutions in their self-reports of the problem solving process than nominal teams, and anecdotal evidence from Experiment 1B suggests that considering a variety of solutions is also associated with considering the particular goals for solving the problem, consistent with the two-stage process of team representation construction described in Chapter 1 (Figure 1.1).

The goal of Experiment 2 was to draw direct parallels between representation construction observed in teams, and proscriptive representation construction heuristics used with individuals to facilitate the process. This was done to see if the benefits of early team interaction and representation construction heuristics were additive. If not, this would be evidence to suggest that interacting teams naturally engage in behavior that is advocated by heuristics used to promote representation construction. If the benefits of early interaction and heuristics are additive, it suggests that the benefit imparted by early interaction is a result of a process distinct from that promoted by representation construction heuristics.

Team interaction structure and the instructions given to begin working on the problem were both manipulated in this experiment. Interacting and nominal teams worked on a case study problem to promote environmentally sustainable living, and were given instructions during

their initial work on the problem either to proceed as normal or to restate the problem in as many ways as possible. Self-reports were elicited periodically in order to evaluate several process variables. Solutions were evaluated for quality and originality using a process similar to Experiment 1A/1B.

4.2 Experiment 2

4.2.1 Method

 4.2.1.1 Participants. Participants were n = 160 undergraduate students from introductory psychology courses. Volunteers received credit towards a class requirement for participation in the study. Volunteers signed up using an online research participation client.

 4.2.1.2 Design. This experiment utilized a 2x2 between-groups factorial design. Three-member teams were either able to work together (interacting structure) or independently (nominal structure) on the problem. In addition, teams were either provided with explicit instructions to restate the problem during their initial solving efforts (instruction condition), or were given no such instructions (no-instruction condition; Table 4.1).

Table 4.1 Design for Experiment 2.

REPRESENTATION CONSTRUCTION

		instruction	**no-instruction**
TEAM STRUCTURE	**nominal**	nominal instruction	nominal no-instruction
	interacting	interacting instruction	interacting no-instruction

4.2.1.3 Materials. Participants were asked to solve a problem to encourage environmentally sustainable living at both the individual and societal levels. This topic was used because it is a current topic in the public consciousness, was familiar and engaging for most students, and was solvable by a variety of different approaches. These inferences were corroborated by data from seven pilot sessions. The problem statement provided a few sentences of background information, and a bulleted list of requirements, much like the peanut sheller problem used in Experiment 1A (Appendix G).

After reading the problem and providing a self-report of their best solution, both nominal and interacting teams were given instructions from one of two representation construction conditions (Figure 4.1). In the "instruction" condition, team members were asked to restate the problem in as many different ways as possible during the first work session. In the "no-instruction" condition, teams were told to continue working on the problem during the first work session. For all other work sessions, teams in both conditions were instructed to continue to work on the problem (Appendix H).

4.2.1.4 Procedure. The procedure was similar to that used in Experiment 1A and 1B. Each session was run with a maximum of one team at a time, and therefore each team was randomly assigned to condition by block. Second, teams were assigned to an extra-nominal condition in the event that any of the three volunteers did not appear for their assigned session, similar to Experiment 1A and 1B. The result was n = 40 teams split evenly across each of the four cells of the main experimental design, with n = 40 individuals assigned to the extra-nominal condition. Similar to Experiment 1A and 1B, data from extra-nominal participants was used to train the LSA model and calibrate raters, but was not used in any inferential analyses.

70

Participants were given 28 minutes to work on the design problem, split up into four smaller seven-minute epochs, with one additional seven-minute epoch to write up their final design solution (Figure 4.1). The representation construction manipulation was provided via instructions during the first work session. Subsequent work session instructions were similar for both instruction and no-instruction conditions, and invoked participants to continue to work on the problem (Appendix H). Between epochs of work on the problem, participants provided independent self-reports of their current best solution to the problem using a computerized form. This form also asked each participant to evaluate the quality of their current best solution rating on a 5-point scale ranging from "Very Poor" to "Very Good" (Appendix I)

After working on the problem, teams wrote up their solution(s) to the design problem. Each participant provided a description of their own (nominal condition) or their team's (interacting condition) final problem solution. This was done to reduce the confounding present between team interaction structure condition and the LSA similarity-to-solution metric (Sections 3.2.2.2 & 3.3.2.2). All participants were instructed to include as part of their solution
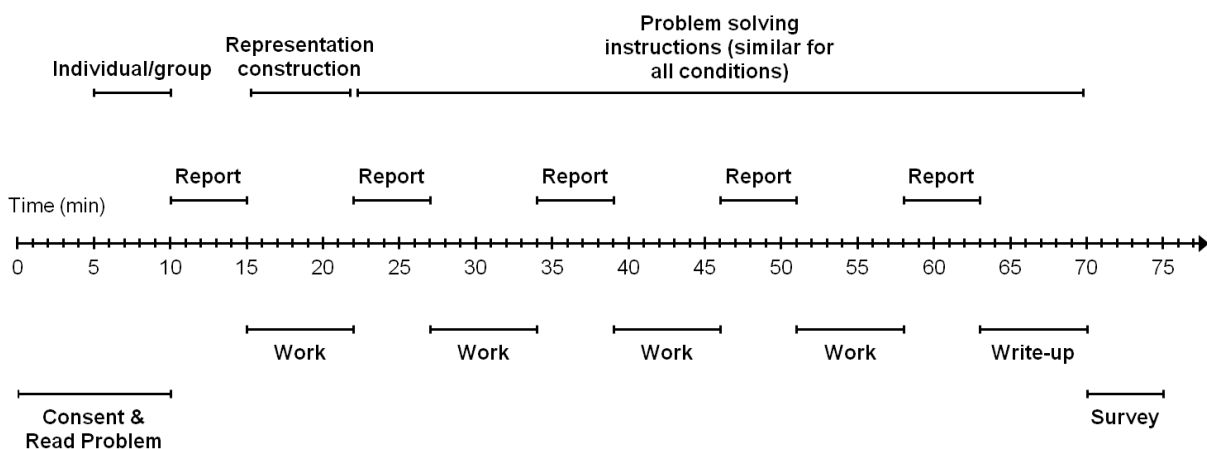
Figure 4.1 Procedure for Experiment 2.

71

Table 4.2 Correlation Matrix Comparing Condition, Consensus Rater Similarity, and LSA Cosine Similarity of Team Solutions.

All correlations significant at the p < 0.001 level.

	condition	*consensus similarity*	*LSA similarity*
Condition	---	0.884	0.308
Consensus similarity		---	0.306
LSA similarity			---

description: a sketch of the design, labels of major elements, and a few sentences describing how

the solution works. Teams then completed a brief questionnaire to supply demographic

information (e.g., age, major) and a rating of their familiarity with the other members of their

team.

In order to make sure that interacting teams were describing the same solution, three

trained raters were asked to evaluate whether each pair of final problem solution descriptions

within a team described either the same solution or two distinct solutions. Correlations were run

between the consensus of these ratings and what would be expected from the condition

participants were in by dummy-coding team interaction structure as "0" if nominal or "1" if

interacting, and dummy-coding consensus ratings as "0" if distinct or "1" if same. Consensus

rater evaluations and condition correlated highly with each other (Table 4.2), suggesting that

members of interacting teams were indeed describing a common final problem solution.

Correlations were also analyzed between both consensus similarity ratings and team

interaction structure condition with LSA cosine similarity of pairs of final problem solution

transcripts, using a corpus of all self-reports and final problem solution transcripts to build the LSA vector space. These correlations are consistent with human similarity judgments based on free recall frequency (Nelson, McEvoy, & Schreiber, 2004) and judge ratings of news articles on a continuous five-point scale (Pincombe, 2004).

4.2.2 Results

4.2.2.1 *Manipulation check.* In order to ensure that participants were behaving consistently with the representation construction instruction manipulation and test H_2, three trained raters counted the number of distinct restatements of the problem and number of distinct solutions detailed in the notes participants provided during the first work session. Raters were trained by familiarization with definitions for restatements and solutions, as well as several examples of restatements or solutions that were excerpted directly or paraphrased from the dataset (Table 4.3).

Raters then completed a calibration set of 10% of the data. Interrater agreement on counts of both restatements and solutions in first work session notes was acceptable ($\alpha = 0.874$ & 0.895, respectively), and raters then evaluated the entire set. Interrater agreement was poor across all three raters for the entire set on both restatements and solutions because one rater had poor agreement with the other two (rater #3, Table 4.4). It should be noted that the pattern of results does not change substantially from removing this rater from subsequent analyses.

Restatement and solution counts were averaged across the two raters who agreed highly with each other and analyzed using separate factorial between-subjects ANOVAs, entering team interaction structure and representation construction instructions as separate between-subjects

Table 4.3 Definitions, Examples, and Intraclass Correlations for Manipulation Check Training Set and Full Dataset.

Criterion	definition	example	α training set	α full set all raters	α full set without rater #3
Restatement	A distinct full or partial answer to the problem is elaborated.	*"Show kids minimizing their food waste. Receive positive reinforcement from peers &adults."*	0.874	0.516	0.853
Solution	A paraphrase or reinterpretation of all or part of the problem statement.	*"Is there a way to live a life that does not involve sacrifice our future environment?"*	0.895	0.677	0.849

Table 4.4 Pairwise Intraclass Correlations Between Raters for Distinct Restatements and Solutions in First Work Session Notes.

	restatements			solutions		
	rater #1	rater #2	rater #3	rater #1	rater #2	rater #3
rater #1	---	0.853	-0.335	---	0.849	0.437
rater #2		---	-0.327		---	0.445
rater #3			---			---

factors. The analysis identified a main effect of representation construction instructions on the number of restatements provided in notes of the first work session, $F(1,116) = 162.118$, $p < 0.001$, $\eta^2 = 0.583$, where teams who were given instructions to restate the problem produced more restatements compared to the no-instruction group. There was no effect of team interaction structure, and no statistical interaction between team interaction structure and representation construction instructions on the number of restatements provided (Figure 4.2).
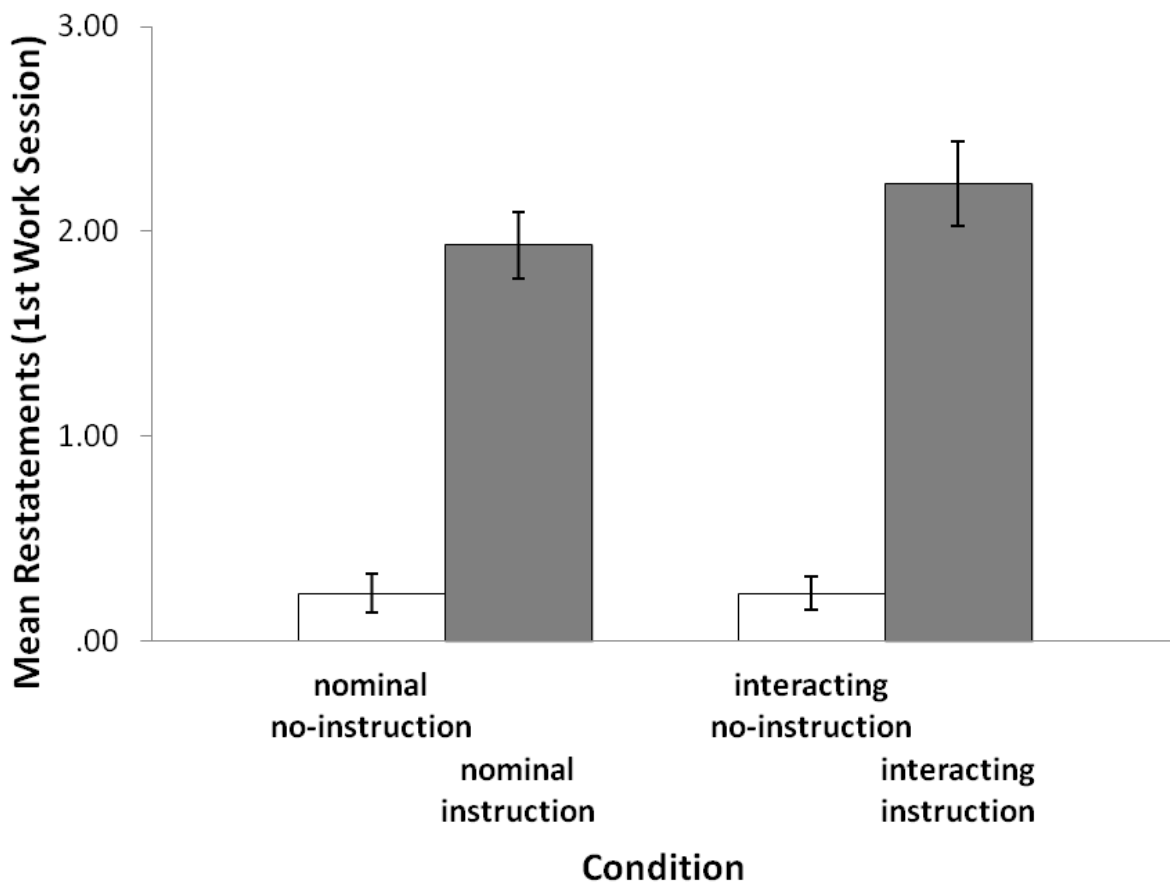
Figure 4.2 Mean Number of Distinct Restatements in Notes of First Work Session.

Bar grouping indicates team interaction structure (nominal left, interacting right), bar color indicates representation construction condition (no-instructions white, instructions grey). Error bars indicate 1SE.

Figure 4.3 Mean Number of Distinct Solutions in Notes of First Work Session.

Bar grouping indicates team interaction structure (nominal left, interacting right), bar color indicates representation construction condition (no-instructions white, instructions grey). Error bars indicate 1SE.

For the number of distinct solutions provided in first work session notes, the analysis identified a significant effect of representation construction instructions, $F(1,116) = 64.852$, $p < 0.001$, $\eta^2 = 0.359$, as well as a marginally significant effect of team interaction structure, $F(1,116) = 3.083$, $p = 0.083$, $\eta^2 = 0.026$. These main effects were qualified by a significant statistical interaction between representation construction instructions and team interaction structure, $F(1,116) = 3.937$, $p = 0.050$, $\eta^2 = 0.033$. Both interacting and nominal teams who were

given representation construction instructions provided very few solutions in their notes during the first work session. In contrast, no-instruction teams provided many more solutions, with interacting teams giving substantially more solutions in their first work session notes than nominal teams (Figure 4.3).

Combined, these results provide support for H_2. Not only do representation construction instructions lead to more restatements of the problem in both nominal and interacting teams, but these instructions also prevent participants from focusing solely on specific solutions to the problem during the first work session. This preventative effect may be greater for interacting teams than nominal teams, as interacting no-instruction teams produced more distinct solutions to the problem in their notes than nominal no-instruction teams.

4.2.2.2 Team coherence. In order to test the hypotheses that both team interaction structure (H_1) and representation construction instructions (H_3) would influence team coherence, a repeated-measures factorial ANOVA was conducted on team coherence scores for each team (Appendix J). This analysis demonstrated that interacting teams tended to have higher coherence over the course of the experiment than nominal teams (Figure J.1), and that coherence for interacting teams increased over time while coherence for nominal teams stayed flat (Figure J.2). This trend of increasing coherence for interacting teams is consistent with Experiment 1A and 1B, and the higher overall coherence for interacting teams is in support of H_1.

4.2.2.3 Similarity-to-solution. In order to better understand the relationship between team interaction structure, representation construction instructions, and changes in individual mental models over time, a repeated-measures factorial ANOVA was conducted on similarity-to-solution scores for each team with team interaction structure and representation construction instruction condition entered as separate between-subjects effects, with time entered as a within-

subjects effect. This analysis identified a main effect of Time, $F(4,464) = 29.947$, $p < 0.001$, η^2 $= 0.205$, as well as a main effect of team interaction structure, $F(1,116) = 52.003$, $p < 0.001$, $\eta^2 = 0.310$. Similarity-to-solution scores increased over time, and were higher for nominal teams than for interacting teams (Figure 4.4 & Figure 4.5).
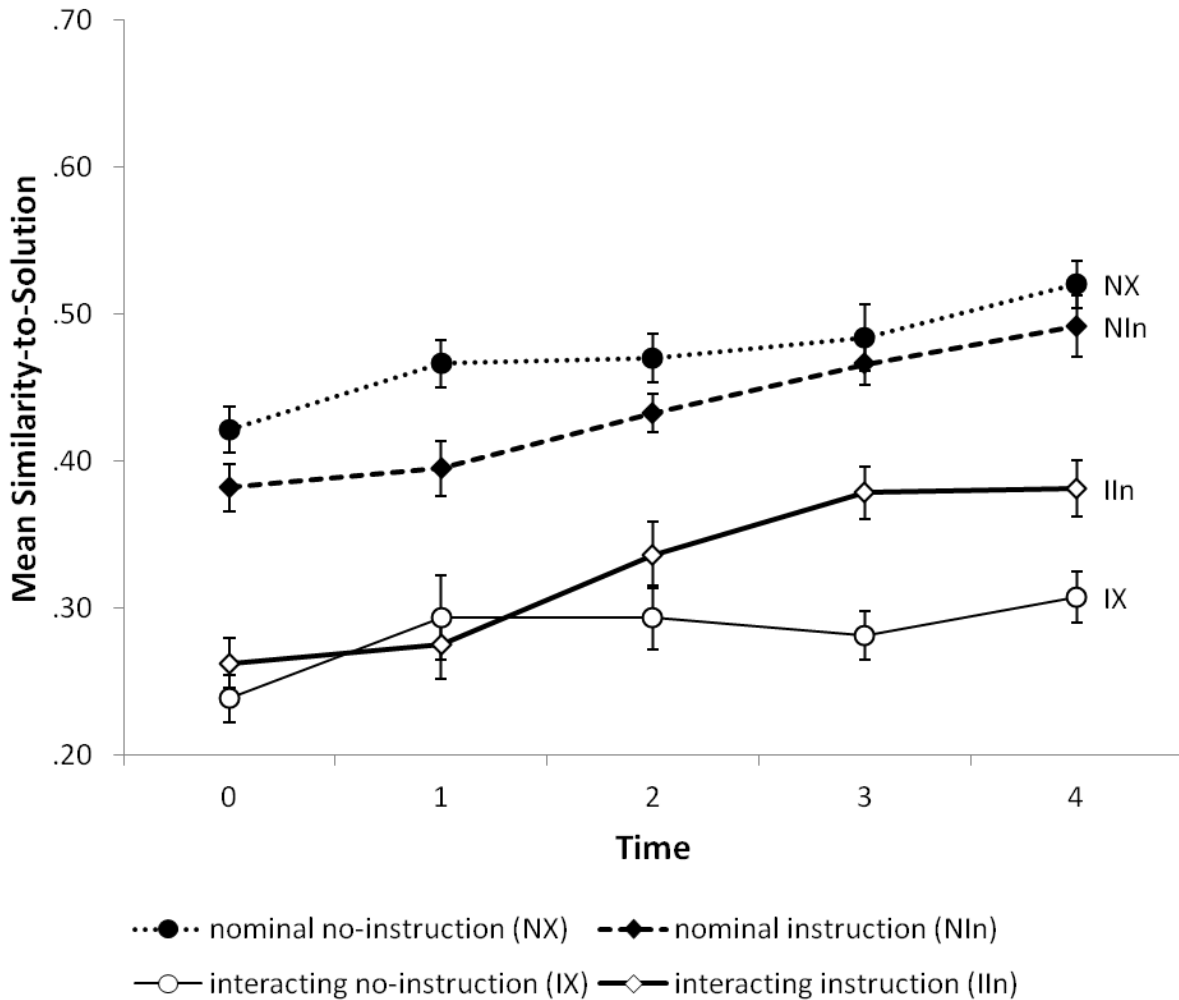
Figure 4.4 Experiment 2 Similarity-to-Solution Over Time.

Each line represents a different team interaction structure by representation construction instruction condition combination. White markers and solid lines indicate interacting teams; black markers and dashed lines indicate nominal teams. Teams given representation construction instructions have square markers and heavy lines; no-instruction teams have round markers and light lines. Error bars indicate 1SE.
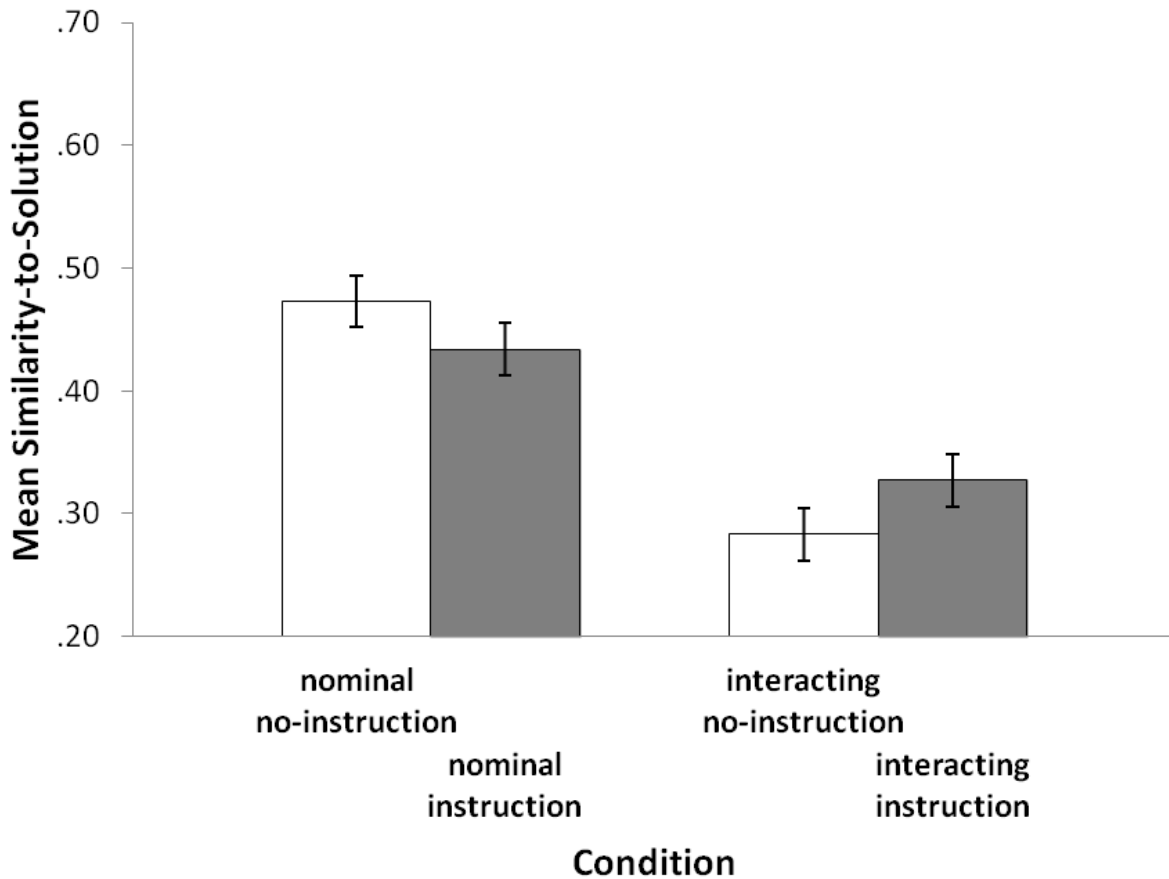
Figure 4.5 Experiment 2 Mean Similarity-to-Solution.

Bar grouping indicates team interaction structure (nominal left, interacting right), bar color indicates representation construction condition (no-instructions white, instructions grey). Error bars indicate 1SE.
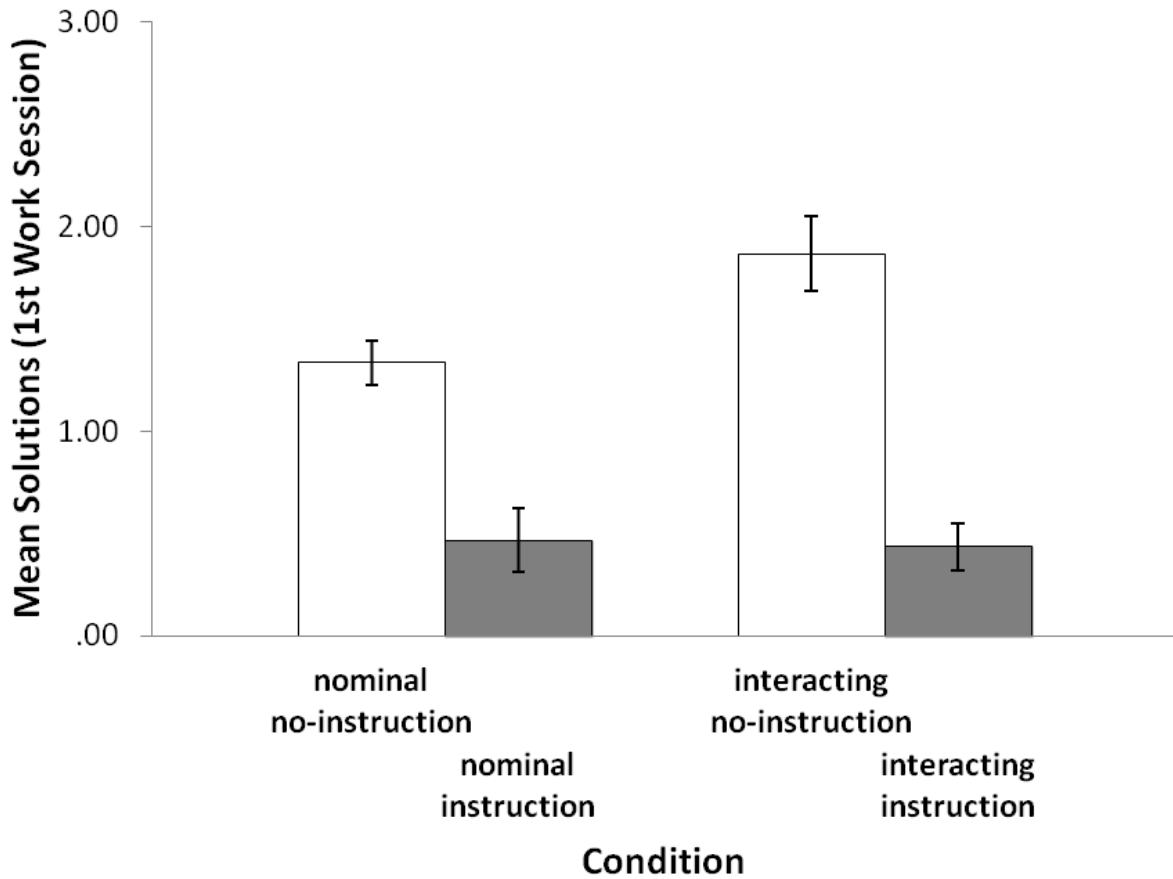
These main effects were qualified by two statistical interactions. The first was a statistical interaction between team interaction structure and representation construction instructions, $F(1,116) = 4.059$, $p = 0.046$, $\eta^2 = 0.034$. For nominal teams, those given representation construction instructions tended to have less similarity between their self-reports of the problem solving process and their final problem solutions. The opposite trend is true for interacting teams, where interacting teams with representation construction instructions had

higher similarity-to-solution scores compared to interacting no-instruction teams. This first

interaction is explained in part by a second statistical interaction between Time and

representation construction instruction condition, $F(4,464) = 5.166$, $p < 0.001$, $\eta^2 = 0.043$.

Teams given representation construction instructions tend to have similarity-to-solution values

that increase at a faster rate over time than no-instruction teams. While this trend appears greater

for interacting teams than for nominal teams, the three-way Time by team interaction structure

by representation construction instructions statistical interaction is non-significant.

In order to check for differences between both interacting teams and nominal teams as a

function of instructional condition, post-hoc contrast with Sidak correction were conducted

separately for interacting and nominal teams at each point in time. These tests identified a

significant difference as a function of representation construction instructions for nominal teams

at t_1, $F(1,116) = 5.027$, $p = 0.027$, $\eta^2 = 0.042$, where nominal no-instruction teams had higher

similarity-to-solution scores at t_1 than nominal instruction teams. This suggests that nominal

instruction teams may be considering a greater diversity of ways to represent the problem than

nominal no-instruction teams.

Differences between interacting teams on similarity-to-solution occurred much later in

the experiment, and show that interacting instruction teams are significantly higher than

interacting no-instruction teams at t_3, $F(1,116) = 7.727$, $p = 0.006$, $\eta^2 = 0.062$, and marginally

significantly higher at t_4, $F(1,116) = 7.727$, $p = 0.006$, $\eta^2 = 0.030$. Representation construction

instructions did not produce differences among interacting teams in how they are representing

the problem in the early part of the experiment, consistent with the model of team representation

construction outlined previously (Figure 1.1). Differences in similarity-to-solution scores in the

latter part of the experiment for interacting teams may suggest that representation construction

instructions encourage interacting teams to commit to a particular representation of a problem earlier during the problem solving process.

Two approaches were taken in order to evaluate the claim that representation construction instructions promote earlier commitment to a particular problem representation in interacting teams. First, a metric called *final path* was developed in order to evaluate qualitative differences in solutions reported by individuals in self-reports as a function of experimental condition. Second, counts of the number of distinct solutions provided in each self-report were evaluated in a manner similar to that of Experiment 1B (Section 3.3.2.5).

4.2.2.4 Final path. A final path metric was developed using a procedure similar to that used earlier to qualitatively compare final problem solution documents provided by members of the same team (Section 4.2.1.4 & Table 4.2). This metric was inspired in part by research with puzzle problem tasks where participants rapidly approach the final correct answer after some exploration of the problem (Kotovsky, et al., 1985; Reber & Kotovsky, 1997). In the current research, final path is meant to describe approach behavior to what becomes the participant's final solution, irrespective of the correctness of that solution. Three trained raters compared contiguous self-reports from each individual, making a judgment of "same" if the pair of self-reports described similar solutions, or "distinct" if the solutions described were qualitatively different. Rater agreement was fair across the population of contrasts evaluated (Fleiss' $\kappa = 0.603$). The raters' consensus response was used to determine whether contiguous self-reports described the same or distinct solutions. As with comparisons of team's final problem solution documents, the consensus response from comparisons of contiguous self-reports correlated with LSA cosine similarity at levels similar to those reported in the literature (Nelson, et al., 2004; Pincombe, 2004).

Table 4.5 Correlation Between Consensus Rater Similarity and LSA Cosine Similarity for Contiguous Self-Reports.

All correlations significant at the p < 0.001 level.

Contrast	r
t_0 & t_1	0.345
t_1 & t_2	0.352
t_2 & t_3	0.400
t_3 & t_4	0.400

To develop an aggregate metric of final path that represents the amount of time spent detailing a representation of the problem that eventually became the final problem solution, the number of rater consensus "same" responses was counted for each participant, summing back from the last self-report contrast (at t_3 & t_4) until a consensus rater response of "distinct" was reported. The resulting final path metric takes on a value that ranges from 0 to 4, with higher values indicating that more time was spent exploring or detailing a representation that the individual applied to what became their own or their team's final problem solution.

Final path scores were entered into a 2x2 between-subjects ANOVA, with team interaction structure and representation construction instructions entered as separate factors. This analysis identified a significant main effect of team interaction structure, $F(1,116) = 22.198$, $p < 0.001$, $\eta^2 = 0.161$, where nominal teams had significantly higher final path scores than interacting teams. The high Final Path scores for nominal teams corroborates the similarity-to-solution data, and suggests that nominal teams tend to commit to a particular problem

representation early in the solving process, possibly fixating on a problem representation they formulated on reading the problem initially. The analysis also identified a marginally significant effect of representation construction instructions, $F(1,116) = 3.459$, $p = 0.065$, $\eta^2 = 0.029$, which was qualified by a significant team interaction structure by representation construction instructions statistical interaction, $F(1,116) = 4.981$, $p = 0.028$, $\eta^2 = 0.041$. Post-hoc analyses with Sidak correction show that nominal teams had comparable final path scores across the representation construction instruction condition, $F(1,116) = 0.069$, $p = 0.793$, $\eta^2 = 0.001$, while interacting instruction teams had much higher final path scores than interacting no-instruction teams, $F(1,116) = 8.370$, $p = 0.005$, $\eta^2 = 0.067$ (Figure 4.6). The similarity of final path scores among nominal teams suggests that both nominal instruction and nominal no-instruction team members may have fixated on a particular representation of the problem they identified early in the solving process. In the case of nominal instruction teams, this fixation occurred after considering and selecting among a few alternatives for representing the problem, likely sometime shortly after considering different ways of restating the problem as indicated by differences between nominal instruction and no-instruction teams in similarity-to-solution scores at t_1 (Figure 4.4) and the number of distinct solutions in self-reports at t_2 (Figure 4.7). In contrast, nominal no-instruction teams were likely to fixate on one of the first solutions they identified when working on the problem. The statistical interaction of final path scores corroborates the idea from differences in interacting team similarity-to-solution scores (Figure 4.4) that interacting instruction teams were committing earlier to a particular problem representation than interacting no-instruction teams.

Table 4.6 Experiment 2 Intraclass Correlation Coefficients Training and Full Datasets for Self-Report Solution Counts.

Dataset	α
1st training set	0.689
2nd training set	0.829
full set	0.774
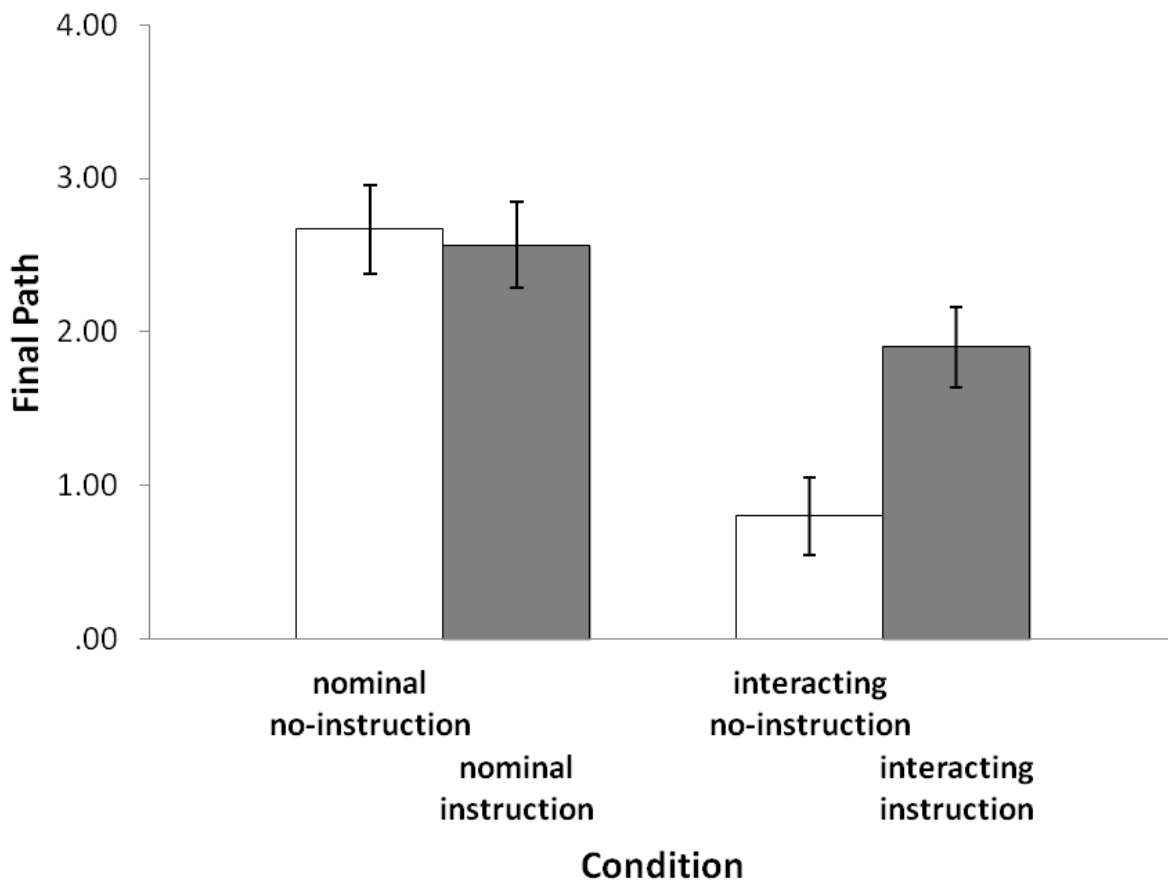
Figure 4.6 Experiment 2 Mean Final Path.

Bar grouping indicates team interaction structure (nominal left, interacting right), bar color indicates representation construction condition (no-instructions white, instructions grey). Error bars indicate 1SE.

4.2.2.5 Self-report solution counts. In order to find evidence of differences in representation construction processes as a function of instruction, three trained raters counted the number of distinct solutions provided in each individual self-report in a manner similar to that done in Experiment 1B (Section 3.3.2.5). Raters were trained by familiarization with the self-report data, as well as the definition of a distinct solution (Table 3.8). Raters then evaluated a calibration set of 10% of self-reports to assess interrater agreement. Agreement on this training set was just below the training criterion of $\alpha = 0.700$, so raters were asked to discuss cases in which they disagreed and then reevaluate the training set. Agreement on this second training set was acceptable, and raters proceeded to evaluate the full dataset (Table 4.6). Evaluations of the number of distinct solutions were averaged across raters to produce an aggregate measure of the number of distinct solutions in each self-report for each participant.

Mean solution counts were analyzed using a 2x2 mixed ANOVA, with team interaction structure and representation construction instruction condition entered as separate between-subjects factors, while Time was entered as a within-subjects factor. The analysis identified a significant main effect of Time, $F(4,464) = 6.977$, $p < 0.001$, $\eta^2 = 0.057$, as well as a significant main effect of team interaction structure, $F(1,116) = 8.052$, $p = 0.005$, $\eta^2 = 0.065$. As in Experiment 1B (Figure 3.11), the number of distinct solutions in self-reports tended to increase over time (Figure 4.7), and was greater for interacting teams compared to nominal teams (Figure 4.8). These main effects were qualified by both a Time by team interaction structure statistical interaction, $F(4,464) = 3.877$, $p = 0.004$, $\eta^2 = 0.032$, and a Time by representation construction instruction condition statistical interaction, $F(4,464) = 4.988$, $p = 0.001$, $\eta^2 = 0.041$. Similar to Experiment 1B, the number of solutions in each self-report for members of nominal teams stayed relatively flat. In contrast, the number of solutions increased for interacting teams once they had

an opportunity to work with other team members, though differences did exist. A post-hoc

contrast with Sidak correction shows a significant difference in number of solutions between

interacting instruction and interacting no-instruction teams at t_1, $F(1,116) = 14.076$, $p < 0.001$, η^2

$= 0.108$. This can be explained by fact that interacting instruction teams had spent the previous
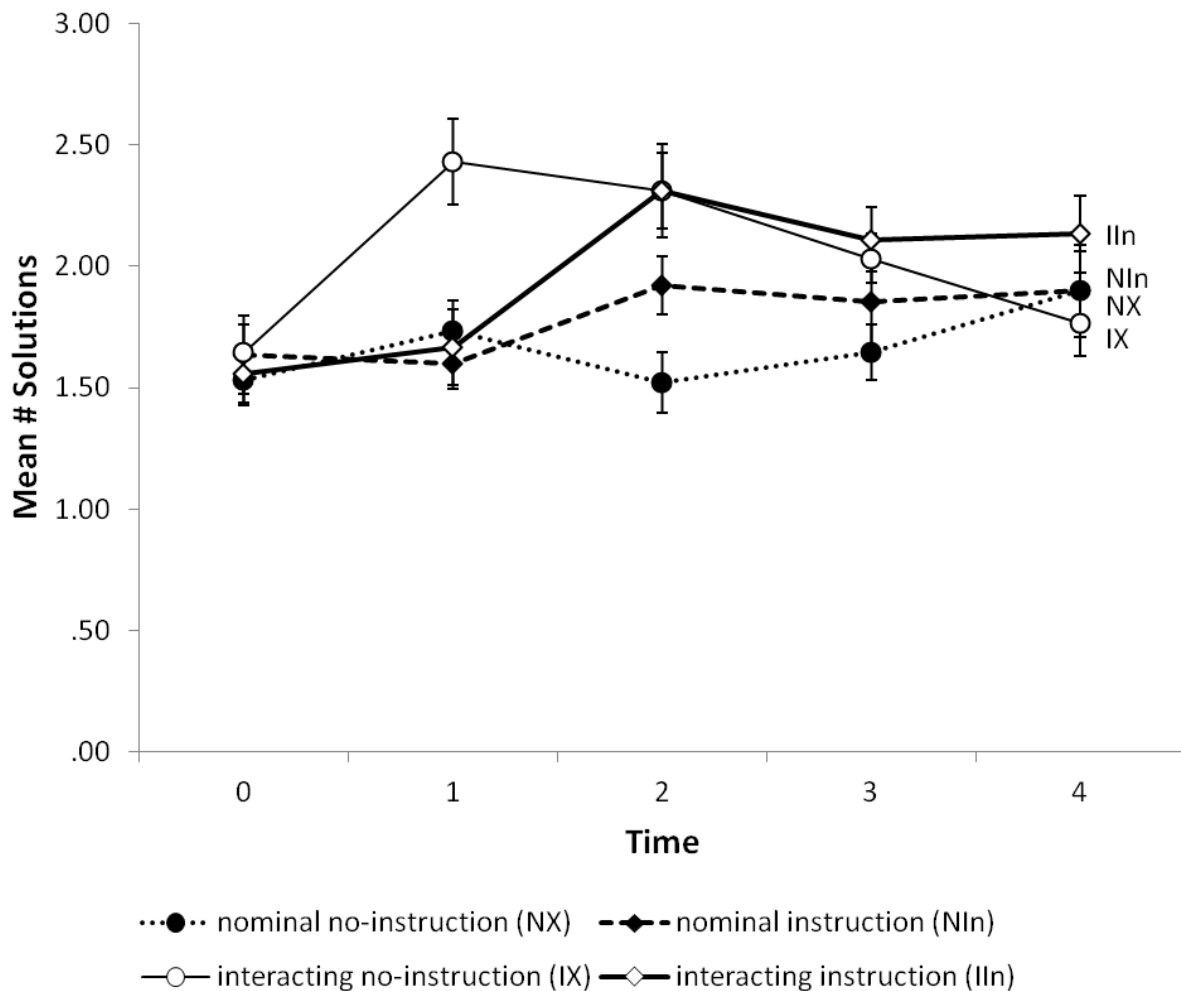
Figure 4.7 Experiment 2 Number of Solutions Over Time.

Each line represents a different team interaction structure by representation construction instruction condition combination. White markers and solid lines indicate interacting teams; black markers and dashed lines indicate nominal teams. Teams given representation construction instructions have square markers and heavy lines; no-instruction teams have round markers and light lines. Error bars indicate 1SE.

epoch restating the problem with other team members, and had not had the opportunity to

consider alternative solutions provided by their teammates. In contrast, interacting no-

instruction teams were able to share solution alternatives with other team members right away,

and therefore did have an opportunity to consider solutions to the problem other than those they

thought about after initially reading the problem. Counts of restatements and solutions in first
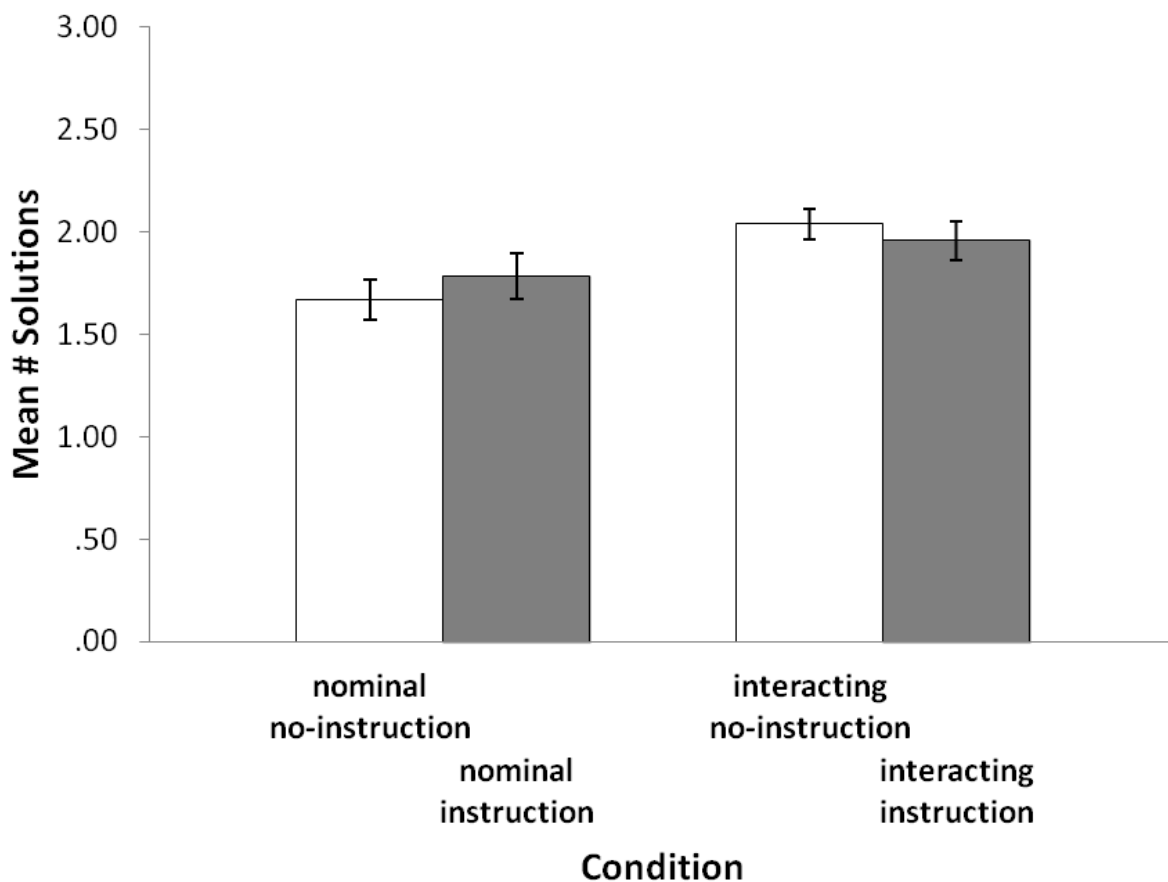
Figure 4.8 Experiment 2 Mean Number of Solutions.

Bar grouping indicates team interaction structure (nominal left, interacting right), bar color indicates representation construction condition (no-instructions white, instructions grey). Error bars indicate 1SE.

work session notes corroborate this idea that interacting instruction team members were exposed to very few solution alternatives compared to interacting no-instruction team members during the first work session (Figure 4.2 & Figure 4.3).

Nominal teams also showed different patterns in self-report solution counts as a function of representation construction instructions. A post-hoc contrast with Sidak correction identified a marginally significant difference between members of nominal instruction and nominal no-instruction teams in the number of solutions provided at t_2, $F(1,116) = 3.517$, $p = 0.063$, $\eta^2 = 0.029$. Nominal instruction team members provided more solutions to the problem that they considered "best" at this time than nominal no-interacting teams. However, whereas this increase for instruction team members at t_2 brought the number of solutions in line with no-instruction team members in the interacting condition, for nominal instruction team members the increase at t_2 did not align their behavior to that of their nominal no-instruction counterparts. The similarity-to-solution data (Figure 4.4) suggests that this difference between nominal instruction and nominal no-instruction team members in self-report solution counts may be attributed to nominal no-instruction participants fixating on a particular problem solution. Nominal no-instruction team members have higher similarity-to-solution scores than nominal instruction team members at t_1, and with an additional work epoch to think about a particular solution in greater detail, nominal no-instruction team members may have stopped considering other solutions as potential candidates.

The solution count results over time provide some support for H_1 and H_2. Consistent with H_1, members of interacting teams considered more solution alternatives as "best" over time than members of nominal teams, and this is consistent with the hypothesis that the solutions generated by team members act as a cue to help teams consider the different ways in which a

problem could be represented (Figure 1.1). While this increase for interacting teams in the number of self-reported solutions was delayed for those given representation construction instructions, the differences between interacting instruction and interacting no-instruction team members became negligible once the interacting instruction team members had an opportunity to share solution alternatives with their teammates.

Consistent with H_2, representation construction instructions were associated with differences in solution counts for members of nominal teams, but not interacting teams. Interacting team members did show different patterns of behavior as a function of representation construction instructions, however the end result was an alignment in terms of the solution counts in self-reports between interacting instruction and interacting no-instruction team members. Nominal instruction team members, in contrast, outperformed their nominal no-instruction counterparts after having the opportunity to work on the problem as normal following restatement generation during the first work session. A comparison between solution count and similarity-to-solution data for nominal teams suggests that nominal no-instruction teams may be fixating on a particular representation of the problem sooner than nominal instruction teams.

4.2.2.6 Final problem solution quality. Final problem solutions were evaluated using a process similar to that in Experiment 1B (Section 3.3.2.3 & 3.3.2.4). In order to evaluate the quality of solutions in support of H_1 and H3, all final problem solutions were reviewed to identify performance criteria (1) that were relevant to objectives stated in the problem statement which participants addressed in their solutions, and (2) which varied in the extent to which any particular solution may have addressed that criteria across the population of final problem solutions provided. Ten criteria were identified and used to develop a decision matrix, where a benchmark exemplar was enumerated for each level of the decision matrix for each criterion to

89

correspond to the relative range of performance on each criterion in the dataset (Appendix K), analogous to the approach taken in Experiment 1A & 1B (Chapter 3).

Three raters were trained by familiarization with the problem statement (Appendix G), set of final problem solutions, and each of the decision matrix criteria. In order to check for interrater agreement, raters completed a training set of 10% of final problem solutions and intraclass correlations were calculated setting a training criterion of $\alpha = 0.70$ across all ratings. Over the set of judgments agreement was below the training criterion ($\alpha = 0.670$; Table 4.7), so raters convened to discuss discrepancies in their judgments and then reevaluated the training set. Agreement over all judgments on this second evaluation of the training set was acceptable ($\alpha = 0.753$) and raters proceeded to evaluate the full dataset. Agreement across the full set of judgments was above the relaxed ICC criterion of $\alpha = 0.6$ used previously (Chapter 3), and was applied to this data because of the similarly imprecise nature of evaluations on some criteria and the incomplete nature of some final design descriptions.

An initial evaluation of differences in quality as a function of condition was conducted using a 2x2 between-subjects ANOVA, entering the mean score across all quality criteria as the dependent variable (Aggregate Quality; Section 3.2.2.3). This analysis did not identify a statistically significant main effect or interaction. However, there was a trend toward higher Aggregate Quality scores for all cells of the design except for the nominal no-instruction condition (Figure 4.9). To explore these trends in greater detail and reduce statistical noise contributed from unstable criteria with poor interrater reliability, a second Aggregate Quality score was computed which included only those criteria where rater agreement was greater than $\alpha = 0.6$ (Table 4.7). An ANOVA using this score as the dependent variable (herein Reliable Aggregate Quality) detected a marginally significant main effect of both team interaction

structure, $F(1,116) = 3.664$, $p = 0.058$, $\eta^2 = 0.031$, and representation construction instruction condition, $F(1,116) = 3.268$, $p = 0.073$, $\eta^2 = 0.027$. Both interacting teams and teams given instructions to restate the problem tended to produce final problem solutions that were of higher quality than nominal teams in the no-instruction condition (Figure 4.10). Although a statistically

Table 4.7 Experiment 2 Intraclass Correlations for Final Quality Evaluation.

	α		
Criterion	*1st training set*	*2nd training set*	*full set*
Overall Quality	0.740	0.703	0.605
Advantages	-0.275	0.583	0.518
Disadvantages	0.629	0.629	0.464
Cost	0.721	0.678	0.694
Feasibility	0.690	0.739	0.804
Changes Behavior	0.565	0.794	0.693
Stickiness / Adoption	0.475	0.304	-0.104
Onset	0.634	0.616	0.799
Longevity / Duration	0.420	0.775	0.579
Public Sentiment	0.796	0.805	0.661
TOTAL	0.670	0.753	0.644
TOTAL after removing ($\alpha < 0.6$):	---	---	0.724
- Advantages			
- Disadvantages			
- Stickiness / Adoption			
- Longevity / Duration			

Figure 4.9 Experiment 2 Mean Aggregate Quality Across All Criteria.

Bar grouping indicates team interaction structure (nominal left, interacting right), bar color indicates representation construction condition (no-instructions white, instructions grey). Error bars indicate 1SE.

significant interaction was not detected, it appears from the figure that interacting teams did not

appear to receive an additional benefit as a function of receiving representation construction

instructions. A series of post-hoc contrasts with Sidak correction were conducted in order to

explore these simple effects. The analysis identified significant differences between quality

scores for nominal no-instruction teams and both nominal instruction teams, $F(1,116) = 5.522$, p

$= 0.020$, $\eta^2 = 0.045$, as well as interacting no-instruction teams, $F(1,116) = 5.881$, $p = 0.017$, $\eta^2 =$

0.048. In both cases Reliable Aggregate Quality scores were lower for nominal no-instruction

teams. In contrast, significant differences were not detected between interacting instruction

teams and either interacting no-instruction or nominal instruction cells of the design. This

suggests that either working in an interacting team *or* using instructions to restate the problem

produce higher quality solutions, but that the benefits from each of these treatments are not

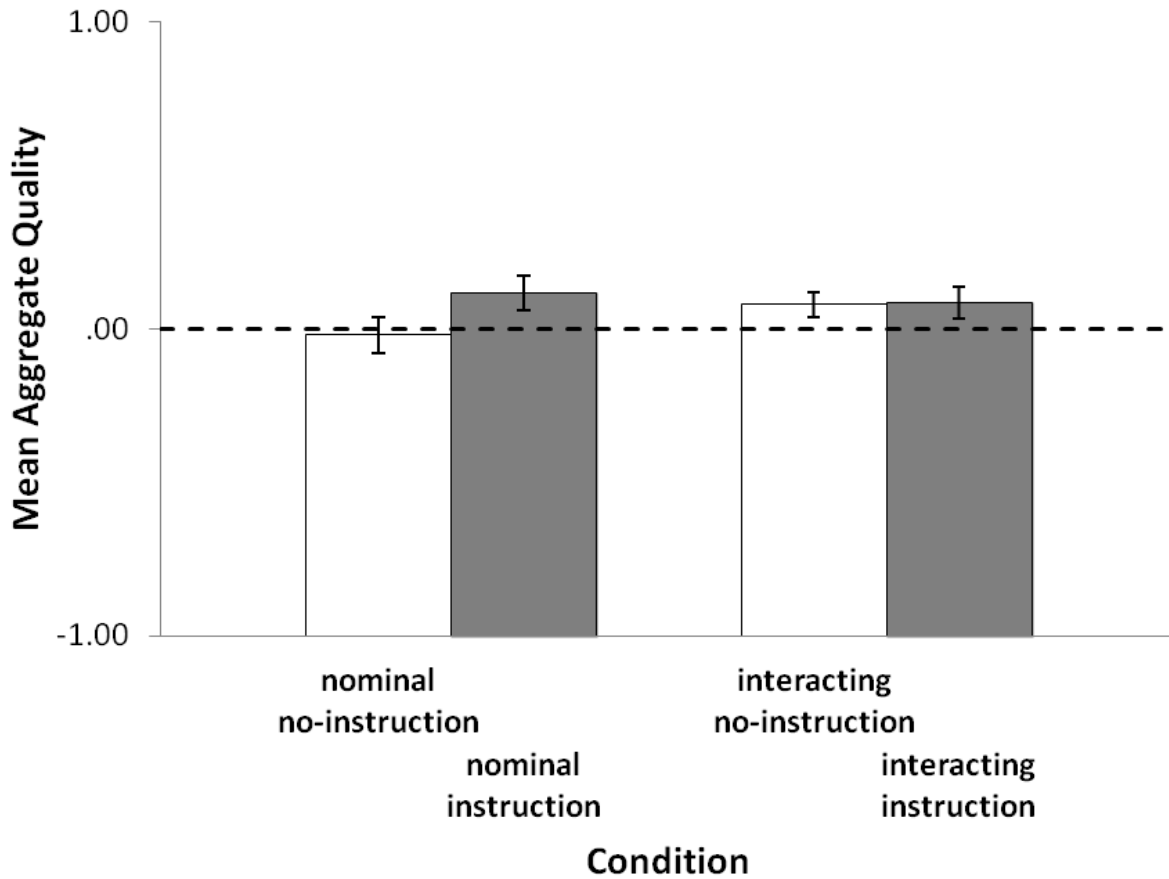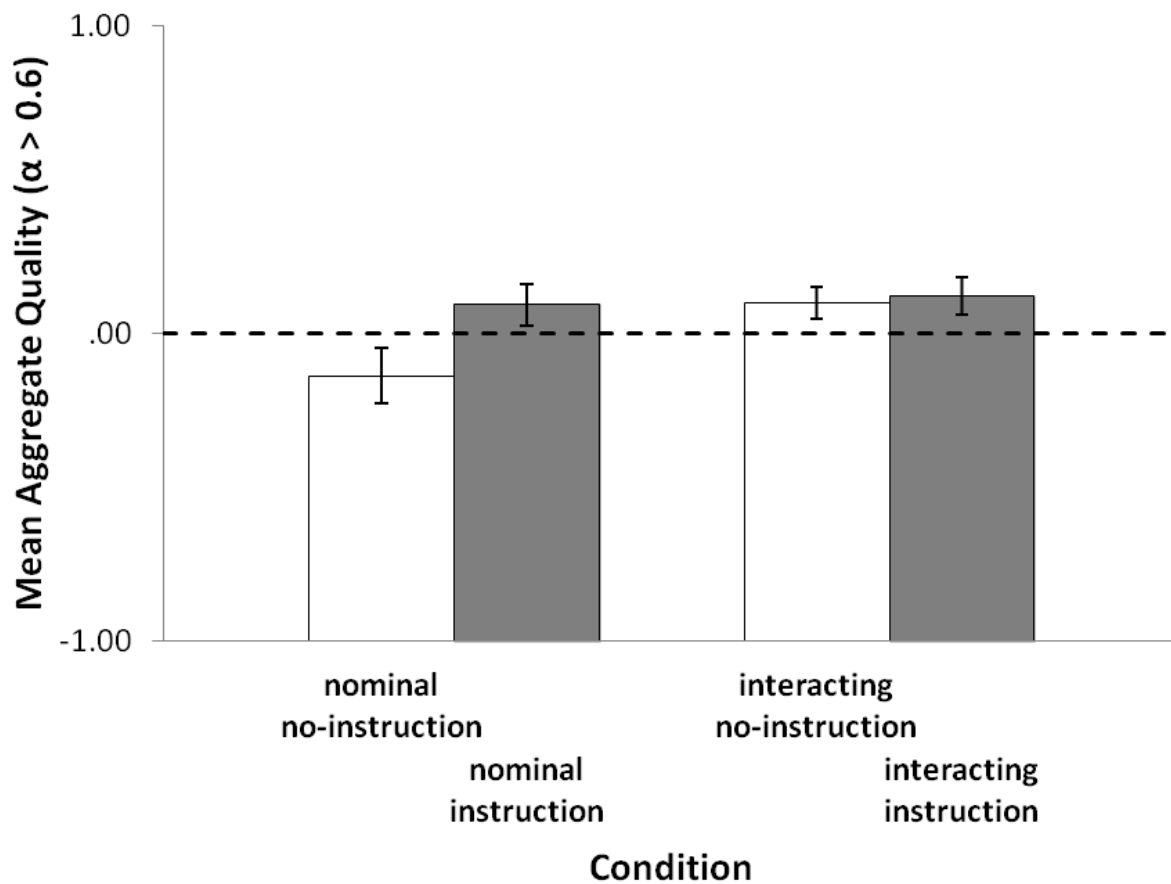*Figure 4.10 Experiment 2 Mean Reliable Aggregate Quality.*

Bar grouping indicates team interaction structure (nominal left, interacting right), bar color indicates representation construction condition (no-instructions white, instructions grey). Error bars indicate 1SE.

additive. Said another way, restatement instructions enabled individuals to perform more like

interacting teams.

In order to identify which criteria were most influential in producing quality differences

across conditions, a multivariate analysis of variance (MANOVA) was conducted, entering each

Reliable Aggregate Quality criteria as a separate dependent variable. This analysis identified a

significant omnibus main effect of representation construction instructions, $F(6,111) = 2.191$, $p =$

0.049, $\eta^2 = 0.106$, with teams who received instructions performing better on most criteria than

those in the no-instruction condition. The omnibus MANOVA also identified a marginally

significant main effect of team interaction structure, $F(6,111) = 2.045$, $p = 0.065$, $\eta^2 = 0.100$.

Interacting teams performed better on most criteria than nominal teams (Appendix M & Figure

M.1).

Two discriminant functions were constructed to identify the combination of dependent

variables used in the MANOVA to discriminate teams from each other on the basis of

representation construction instruction condition and team interaction structure. With respect to

instruction condition, a function was able to be constructed that discriminated instruction teams

from no-instruction teams, canonical $R^2 = 0.325$, $\Lambda = 0.894$, $X^2(6) = 12.849$, $p = 0.046$.

Correlations between dependent variables and the discriminant function not only corroborate

univariate ANOVA results in showing that teams which were given instructions to restate the

problem developed solutions that were less expensive and were more likely to be perceived

favorably by the public, but these solutions were also more likely to have a more immediate

impact than those by teams who were not given these instructions. In contrast, instruction teams

were less likely to produce solutions that changed individuals' behavior (Table 4.8).

Table 4.8 Standardized Discriminant Function Coefficients for Instruction Condition and Correlations Between Discriminant Function Coefficients and Outcome for Experiment 2.

Negative discriminant function values indicate "no-instruction" representation construction instruction condition, positive values indicate "instruction."

Criterion	β	r
Overall Quality	0.108	0.153
Cost	1.509	0.538
Feasibility	-1.910	0.234
Changes Behavior	-0.141	-0.306
Onset	0.511	0.327
Public Sentiment	0.754	0.542

Table 4.9 Standardized Discriminant Function Coefficients for Team Interaction Structure and Correlations Between Discriminant Function Coefficients and Outcome for Experiment 2.

Negative discriminant function values indicate "interacting" team structure, positive values indicate "nominal" team structure.

Criterion	β	r
Overall Quality	-0.194	-0.484
Cost	-0.810	0.541
Feasibility	0.146	0.775
Changes Behavior	-0.269	-0.640
Onset	1.158	0.872
Public Sentiment	0.118	0.347

A function was also able to be constructed which discriminated interacting teams from nominal teams, though its ability to do so was only marginally significant, canonical $R^2 = 0.312$, $\Lambda = 0.902$, $X^2(6) = 11.810$, $p = 0.066$. Correlations between dependent variables and the discriminant function show that interacting teams were more likely to provide solutions that were

likely to have a more immediate impact, were more feasible to implement, and cost less than those solutions provided by nominal teams. In contrast, nominal teams were more likely to propose solutions that were higher in Overall Quality, and were more likely to change behavior than interacting teams' solutions (Table 4.9). Overall Quality placed some emphasis on the logical coherence of proposed solutions. Members of nominal teams tended to have longer and better elaborated segments of text as part of their solution descriptions that explained how the parts of their solutions fit together. In contrast, interacting teams were less explicit when describing interactions between different pieces of their plan.

Overall, final problem solution quality results provide some support for H_1 and H_3. Although interacting or utilizing instructions to restate the problem in different ways did not have a beneficial impact on all criteria, engaging in one or the other did have a positive impact on more criteria overall than it did a negative impact. Furthermore, Reliable Aggregate Quality results suggest that the benefit of interaction and representation construction instructions are not additive; that is to say that interacting and using instructions to restate the problem provided no additional benefit to teams who were utilizing either in isolation.

4.2.2.7 Final problem solution originality. The originality of final problem solutions was also evaluated, using an approach similar to that of Experiment 1B (Section 3.3.2.4), to more closely link Experiment 2 with the brainstorming literature. In order to explore the impact of team interaction structure and representation construction instructions on solution originality, raters were asked to score designs on the basis of their originality. The decision matrix scale for evaluating originality (Appendix L) was developed and evaluated for interrater reliability in the same manner and at the same time as quality metrics for this dataset. Interrater reliability on originality for both training sets ($\alpha = 0.827$ & 0.837, respectively) and the full dataset ($\alpha = 0.713$)

was acceptable, so originality scores were averaged across raters to produce a single originality

score for each final problem solution. These scores were then entered into a 2x2 between-

subjects ANOVA with team interaction structure and representation construction instruction

condition entered as separate factors. This analysis identified a main effect of team interaction

structure, $F(1,20) = 25.368$, $p < 0.001$, $\eta^2 = 0.179$, where interacting teams produced less original
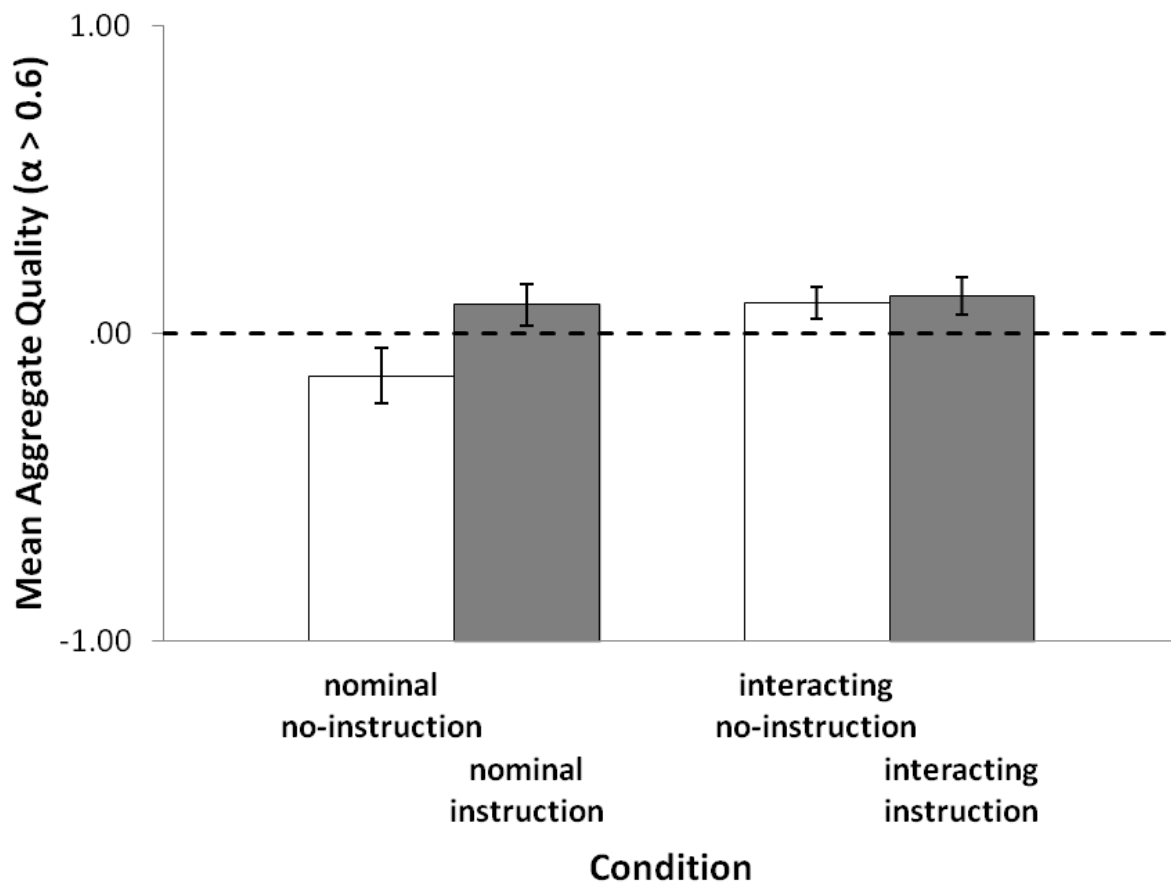
Figure 4.11 Experiment 2 Mean Originality Ratings.

Bar grouping indicates team interaction structure (nominal left, interacting right), bar color indicates representation construction condition (no-instructions white, instructions grey). Error bars indicate 1SE.

solutions than nominal teams (Figure 4.11). No other main effects or interactions were detected.

This trend toward higher originality for nominal teams is consistent with the results of

Experiment 1B, and corroborates the notion that members of interacting teams are concerned

about how their ideas are being interpr*e*ted by their teammates. Members of nominal teams do

n*o*t need to negotiate their solutions with other team members, and may be more likely not only

to propose more novel solution configurations, but also to stick with these less conventional

solutions. Qualitatively, members of nominal teams were more likely to propose outlandish

solutions to the problem that were highly impractical, including starting a military-style coup to

institute a pro-environment government, developing futuristic telecommuting capabilities that

decreases the need for people to commute to school/work, and moving people to space to

decrease their impact on the earth.

    4.2.2.8 Self-perceived quality. Participants were asked to rate the quality of the

solution(s) contained in their self-reports throughout the problem solving process on a one to five

scale ranging from "very poor" to "very good" (Appendix O). These results show that self-

perceived quality tended to increase over time for all conditions. In addition, teams given

instructions to restate the problem tended to have a grea*t*er degree of increase in self-perceived

quality over t*i*me than no-instruction teams, although instruction teams did not have greater self-

perceived quality scores at any point during the experiment.

    4.2.2.9 Mediation analyses. A series of correlations were computed to investigate the

possibility of a mediating relationship between process variables which capture the extent to

which individuals are engaged in representation construction, team coherence, and the quality of

final problem solutions (Section 4.2.2.6). If representation construction processes mediate the

relationship between team coherence and solution quality, there should first be a correlational

relationship between (1) team coherence and representation construction, (2) team coherence and solution quality, and (3) representation construction and solution quality. Change in team coherence ($t_4 - t_0$) was used as the measurement of team coherence for these correlational analyses. Since the representation construction process variables and final problem solution quality are all measured at the individual level, the change in team coherence scores for a team was assigned to each individual within the team for purposes of this analysis. Several variables were used to estimate the extent to which individuals were engaged in representation construction. These included both the number of restatements and the number of solutions provided in first work session notes (Section 4.2.2.1: Manipulation Check), and the number of solutions described in self-reports at t_1 and t_2 as the greatest variance in solution counts as a function of condition occurred at these times and they are also close in time to when the condition manipulations were administered. Reliable Aggregate Quality was entered as the quality variable for purpose of this analysis. Correlations between all of these factors and the two experimental factors were also calculated.

A Bonferroni correction was applied to reduce Type 1 error, with the resulting value to accept a correlation as significant set at $p = 0.0018$ (Table 4.10). These analyses did not identify significant relationships between any of the pairs of variables necessary to infer a mediating relationship. As a result, no mediating relationship is shown to exist between team coherence, representation construction processes, and the quality of final problem solutions. While support was found elsewhere in Experiment 2 for H_1 in that team interaction influences team coherence, representation construction, and solution quality separately, it does not appear that team coherence or representation construction influence one another to impact solution quality and thus they are making independent contributions to the quality effects seen in Experiment 2.

99

Table 4.10 Correlations Between Team Coherence, Representation Construction Process Variables, Solution Quality, and Experimental Factors.

* indicates significance at p < 0.0018.

	Team coherence change	Representation construction instructions	Team interaction structure	1st work session restatements	1st work session solutions	t_1 self-report solutions	t_2 self-report solutions	Reliable Aggregate Quality
Team coherence change	---	-0.001	0.566*	-0.064	0.102	0.105	0.067	0.209
Representation construction instructions		---	0.000	0.761*	-0.588*	-0.266	0.114	0.162
Team interaction structure			---	0.062	0.128	0.226	0.336*	0.171
1st work session restatements				---	-0.598*	-0.069	0.324*	0.075
1st work session solutions					---	0.219*	-0.082	-0.152
t_1 self-report solutions						---	0.353*	0.036
t_2 self-report solutions							---	0.017
Reliable Aggregate Quality								---

4.3 Summary and Conclusions

The results of Experiment 2 provide some support for all three hypotheses. Related to H_1, interacting teams were shown to have greater team coherence and solution quality than nominal teams, but did not engage in more representation construction behavior. Consistent with H_2, instructions to restate the problem did have an impact on measures of representation construction, including an increase in the number of restatements and a decrease in the number of solutions proposed during the first work session, as well as differences as a function of the representation construction instruction condition for teams who interacted during Experiment 2. Partial support was found for H_3. While representation construction instructions did influence representation construction behavior and also solution quality, data exploration in support of mediation analysis did not identify an influence of instructions or behaviors on team coherence.

Instructions to restate the problem were found to be equally effective for individuals working both independently and collaboratively to encourage them to think about different ways of conceptualizing the problem. Instructions to restate the problem were associated with a greater number of actual problem restatements in notes written during that first work session when that manipulation was deployed, fewer solution statements in those notes, and consideration of fewer solutions to the problem as "best" until immediately following that restatement exercise. Similarity-to-solution and final path indices suggest that both nominal instruction and nominal no-instruction teams tended to fixate on a representation identified early in the solving process. However, for nominal instruction teams, this fixation occurred after individuals took a moment to consider alternative representations of the problem, leading them to select a representation that ultimately benefited them in terms of the quality of the solutions that nominal instruction team members ultimately produced.

101

Instructions among teams who were interacting were also associated with an earlier commitment to a particular solution representation as measured by similarity-to-solution and final path indices. One reason for this earlier commitment may be that interacting instruction teams were encouraged to adopt a strategic approach, focusing their efforts on the problem in an increasingly narrow way by starting with a goal they identified through restating the problem, and then clarifying how that goal is to be achieved over time. In contrast, interacting no-instruction teams may prefer a solution-exemplar approach where they randomly generate unique solutions and identify a goal only once these solution alternatives conflict, then randomly generate more unique solution alternatives in hopes that one will address the identified goal. Teams who received instructions not only produced higher quality problem solutions, but analysis of self-reports to evaluate the quality of interim solutions (Appendix N) suggests the interim products they proposed also tended to be of higher quality. Instructions were not only associated with actual solution quality, but were also associated with increasing confidence of team members in the quality of their problem solutions over time (Appendix O).

Team interaction was again shown to influence team coherence. Although interaction was associated with higher quality solutions, the team coherence produced by interaction was not significantly correlated with quality of final problem solutions. Similarity-to-solution and final path data suggest that interacting teams spent less time compared to nominal teams committed to a particular problem representation and/or solution. Interacting teams tended to entertain more "best" solutions as they worked on the problem, as evidenced by higher solution counts in self-reports relative to nominal teams.

Instructions to restate the problem did not have an influence on team coherence, even when considering only interacting teams where these instructions might influence the dialogue

among team members in a different way compared to teams who received no such instructions. Differences in final path and similarity-to-solution between interacting instruction and interacting no-instruction teams suggests that interacting instruction teams tended to commit to a particular class of solution to the problem sooner than interacting no-instruction teams, but earlier commitment did not translate into greater team coherence or higher quality solutions compared to interacting no-instruction teams. It appears that although both interaction and instructions produced improvements in interim and final problem solution quality, the benefits of these interventions were not additive. This provides support to the notion that these distinct interventions invoke comparable problem solving processes, such as identifying one or a small number of specific goals for solving the problem early in the problem solving process. Instructions to restate the problem may invoke a focus on the goals for solving among nominal teams as well, and may have helped to produce performance on-par with interacting teams among individuals in the nominal instruction condition.

Chapter 5. General Discussion, Conclusions, and Contributions

These two experiments find support for the idea that representational processes which occur early in problem solving and encourage individuals to consider alternative representations of a problem have a powerful influence on both individual and team performance. Experiment 1A and 1B found that early interaction induced representation construction and subsequent performance. Teams whose members worked together early on during solving tended to perform better than those who did not. Experiment 2 demonstrated that individuals given instructions to restate a problem produce solutions of comparable quality to interacting teams. Self-report solution counts in Experiment 1B & 2 suggest the reason why interaction or instructions to restate the problem both produce a quality boost is because both of these interventions encourage team members to entertain more options for solving the problem as viable, and may reflect deliberation over the best goal or goals to be solving for to achieve success.

5.1 General discussion

This series of experiments finds support for the thesis in that representation construction and activities which promote it leads to changes in team and individual task mental models and the quality of resulting solutions. Related to H_1, interacting teams had greater team coherence and higher solution quality than nominal teams. Higher quality solutions for interacting teams came at the expense of solution originality, as nominal teams tended to have more original solutions than interacting teams. This parallels the results of brainstorming research, which finds that nominal teams tend to produce more ideas than interacting ones because nominal team members do not have to take turns in conversation with other teammates on the problem, nor do

104

nominal team members need to be concerned about their solutions being received poorly by the group. The number of "best" solutions in self-reports, when taken as a proxy for representation construction, suggests that team interaction does facilitate representation construction processes by priming team members with alternative representations to a problem that they may not have considered on their own. As a result, interacting teams were less likely to fixate on one particular problem representation. Interacting teams produced interim self-reports that were less similar to each other and to the final problem solution than nominal teams, who were more likely to become fixated on a specific problem representation.

Instructions to restate the problem did facilitate representation construction in both interacting and nominal teams in Experiment 2, consistent with H_2. Members of both interacting and nominal teams given these instructions produced more restatements to the problem in their notes during the first work session. These instructions helped nominal teams consider a greater number of acceptable alternative solutions than nominal teams not given such instructions, while they assisted interacting teams in committing to a particular representation of the problem earlier than interacting teams without instruction.

Related to H_3, representation construction as induced through instructions and assessed by counts of solutions considered "best" throughout the solving process positively influenced the quality of final problem solutions in Experiment 2. However, a mediation model could not be constructed that linked these changes in representation construction to changes in team coherence. Although H_3 proposed that instructions would encourage team members to deliberate goals and constraints, eventually deciding on a set with which to proceed, and thus boosting team coherence, this result could not be disentangled from the strong main effect of team members working together on team coherence. The effect of interaction on team coherence may therefore

mask any effect that agreeing on goals for solving might have on team coherence. A challenge for future research is to either identify methods that make team coherence metrics like that used here more sensitive to changes in the content of communications, or conversely to develop systems that make it easy to identify when a set of goals and constraints have been agreed upon from either naturalistic speech data or written self-reports of the problem solving process.

5.2 Conclusion

These two experiments find that representation construction and activities which promote representation construction early in the solving process, can boost problem solving performance. Most importantly, activities that promote representation construction can be used to encourage individual problem solvers to produce solutions to problems that are of comparable quality to teams working together. This has clear potential applications for real-world problem solving. An assumption when assembling many problem solving teams, either implicit or explicit, is that teams are more likely to produce high-quality work compared to individuals. This research suggests that, if provided proper instruction and/or heuristics to approach a problem, individuals are capable of achieving comparably high-quality results. Care should be taken when extending these results to contexts where the problems are larger, or more complex, or require a greater diversity of knowledge than those used in this dissertation. In addition, future work should attempt to replicate these results using industry teams to understand how representation construction behaviors occur "in the wild" as team members are forced to split their time across projects, or begin working on problems asynchronously with respect to other team members.

In addition, this work provides data supporting the idea that correlations between team coherence and solution quality seen in Dong et al. (2004) and others (Bearman, Paletz, Orasanu,

106

& Thomas, 2010; Fu, et al., 2010) may be reinterpreted as a causal relationship between an activity that promotes representation construction (e.g., team interaction) and solution quality, consistent with the proposed team representation construction model at the beginning of this dissertation (Figure 1.1). Future work should aim to clarify this relationship between task mental model coherence and performance by developing more sensitive measures of the content of team communication, particularly early on in the problem solving process. Though data here suggests that individuals working in isolation have a tendency to become fixated on a single problem representation early in the solving process, future research should provide further analysis on the role that fixation prevention, through means such as team interaction or problem restatement, might play in improving the quality of solutions. The negotiation of a common problem representation among teams who interact early or individuals working in isolation while restating the problem during the course of solving may be consistent with either team coherence or fixation prevention explanations.

One important consideration is that the teams in these experiments were intentionally constructed to be new teams that had not previously worked together and were relatively flat in their leadership structure. Teams with different work histories or leadership structures may engage in representation construction through interaction or develop coherent task mental model representations using processes that differ as a function of team interaction structure from those studied here. For instance, hierarchically structured teams with one clear leader may agree on a common problem representation because the leader has decided on a representation to pursue that other team members adopt, whereas members of flat teams (like those in the current experiments) may have a better opportunity to discuss attributes of the problem and negotiate with each other which of those attributes are relevant. Potential differences in team structure and

107

processes that give rise to coherent team mental models should be considered when disentangling the relative contribution of representation construction and/or team coherence to performance on problem solving tasks.

This research demonstrated a clear benefit of early interventions for problem solving tasks of relatively brief duration. However, such relatively short projects are often the exception rather than the rule in business and industry. Future work should attempt to extend these laboratory findings in two directions. First, future research should extend the duration and types of problems pursued by both individuals and teams in this line of inquiry to more naturalistic contexts. Second, future work should focus on enumerating the strategies that may be employed to promote representation construction or early coherence on a common task mental model. This may include use of propitious examples when introducing problems (i.e., Fu, et al., 2010), or additional techniques beyond restating the problem which encourage development of a clear representation and set of goals for solving the problem (Cagan & Vogel, 2013; Reiter-Palmon & Robinson, 2009).

The experiments described in this dissertation also provide insights on the relationship between brainstorming and problem solving literatures with respect to team interaction structure. Whether teams or individuals are "best" for solving problems depends on your outcome criterion. Here, we found that interacting teams produce higher quality results, but nominal teams produce more original results. The conclusions are similar in brainstorming research, which finds that nominal teams produce more solutions, though interacting teams tend to produce better solutions (Section 2.4). Most studies of brainstorming implement a process of idea generation and selection that does not fully capture the iterative nature of idea generation and refinement seen in real world problem solving. The focus of most brainstorming studies is

only on the generation and identification of alternatives, and sometimes subsequent selection of a subset of ideas. Problem solving individuals and teams in the real world frequently combine and modify ideas while searching the problem space for alternatives. In addition, the process of generation then selection of ideas artificially breaks the process into two phases, whereas real problem solving individuals and teams are more likely to iterate through idea generation and selection throughout the course of problem solving. The above experiments find that early interventions can be beneficial for problem solving performance, without following methods explicitly outlined in the brainstorming literature (Baruah & Paulus, 2009). In addition, the above experiments provide a framework for future research in that they present a methodology for measuring process variables like team coherence and similarity of brainstormed ideas at much higher resolution than has been done to-date, and in a way that researchers have only been able to simulate via computational models up until now (Brown & Paulus, 1996; Brown, et al., 1998).

5.3 Contributions

This dissertation presents several contributions to the extant literature on problem solving, team processes, and mental model development. These contributions can be divided into two categories: theoretical and methodological.

5.3.1 Theoretical contributions

- Teams perform better than individuals working without heuristics like restating the problem because collaboration naturally encourages team members to consider alternative ways of representing the problem early in the solving process.

- Individuals who are encouraged to consider alternative ways of representing the problem produce solutions that are as good as those produced by teams.

- Team coherence on a task mental model does not produce higher quality solutions. Agreeing on a common approach or rough solution framework through deliberation is the key feature of task mental model overlap, not the degree of agreement on other task features.

- Whether teams or individuals are "better" in the context of a problem solving task depends on how you measure the outcome. This research shows that teams perform better than individuals when one high quality solution is desired. In contrast, brainstorming research finds that individuals are better than teams when many solutions are desired and quality is not a consideration (Section 2.4)

5.3.2 Methodological contributions

- LSA was successfully extended to make relative comparisons of self-reports broader problems and less specific participant expertise than previous research.

- Two process metrics were developed to quantify the similarity of ill-defined problem representations using textual data.
   - LSA similarity-to-solution, which provides a quantitative comparison of the degree to which a self-report of a current solution is similar to the final problem solution.

o   Final path, which provides an estimate of the point at which a solver begins to

use a representation of the problem that is qualitatively similar to what

eventually becomes his or her final answer to the problem.

5.4 Coda

This dissertation provides important insights into the process by which collaborating

teams produce solutions of high quality. The key factor is to encourage individuals to identify

different ways to think about the problem, especially in ways that are different from the

individual's first impression of a solution or approach to the problem. Working with other

people, who are likely to have different first impressions of the problem compared to the

individual, is only one way to accomplish this goal. Restating the problem in different ways was

another mechanism identified in this dissertation, and was shown to facilitate individual problem

solvers in performing as well as teams of solvers working together. These findings present

potential opportunities to those solving real-world problems to increase productivity by paying

attention to the problem solving tasks given to teams, and considering whether individuals given

the proper instruction may be able to address these tasks just as effectively.

REFERENCES

Amabile, T. M. (1996). *Creativity in context*. Boulder, CO: Westview Press.

Badke-Schaub, P., Neumann, A., Lauche, K., & Mohammed, S. (2007). Mental models in design teams: a valid approach to performance in design collaboration? *CoDesign, 3*(1), 5-20.

Baer, J. M. (1988). Long-term effects of creativity training with middle school students. *The Journal of Early Adolescence, 8*(2), 183-193.

Baruah, J., & Paulus, P. B. (2008). Effects of Training on Idea Generation in Groups. *Small Group Research, 39*(5), 523-541.

Baruah, J., & Paulus, P. B. (2009). Enhancing group creativity: The search for synergy. In E. Mannix, M. A. Neale & J. A. Goncalo (Eds.), *Creativity in Groups* (pp. 29-56). Bingley, UK: Emerald Group Publishing, Ltd.

Bearman, C., Paletz, S. B. F., Orasanu, J., & Thomas, M. J. W. (2010). The Breakdown of Coordinated Decision Making in Distributed Systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 52*(2), 173-188.

Brown, V. R., & Paulus, P. B. (1996). A simple dynamic model of social factors in group brainstorming. *Small Group Research, 27*(1), 91-114.

Brown, V. R., & Paulus, P. B. (2002). Making group brainstorming more effective: Recommendations from an associative memory perspective. *Current Directions in Psychological Science, 11*(6), 208-212.

Brown, V. R., Tumeo, M., Larey, T. S., & Paulus, P. B. (1998). Modeling cognitive interactions during group brainstorming. *Small Group Research, 29*(4), 495-526.

Cagan, J., & Vogel, C. M. (2013). *Creating breakthrough products : revealing the secrets that drive global innovation* (2nd ed.). Upper Saddle River, N.J.: FT Press.

Cannon-Bowers, J. A., Salas, E., & Converse, S. (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, England: Lawrence Erlbaum Associates Inc.

Cooke, N. J., Gorman, J. C., Myers, C. W., & Duran, J. L. (2013). Interactive Team Cognition. *Cognitive Science, 37*(2), 255-285.

Craik, K. J. W. (1943). *The Nature of Explanation*. Cambridge, UK: University Press.

DeChurch, L. A., & Mesmer-Magnus, J. R. (2010). The cognitive underpinnings of effective teamwork: A meta-analysis. *Journal of Applied Psychology, 95*(1), 32-53.

Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., & Harshman, R. (1990). Indexing by latent semantic analysis. *Journal of the American Society for Information Science, 41*(6), 391-407.

Diehl, M., & Stroebe, W. (1987). Productivity loss in brainstorming groups: Toward the solution of a riddle. *Journal of Personality and Social Psychology, 53*(3), 497-509.

Dong, A. (2005). The latent semantic approach to studying design team communication. *Design Studies, 26*(5), 445-461.

Dong, A., Hill, A. W., & Agogino, A. M. (2004). A document analysis method for characterizing design team performance. *Journal of Mechanical Design, 126*(3), 378-385.

Faure, C. (2004). Beyond brainstorming: Effects of different group procedures on selection of ideas and satisfaction with the process. *Journal of Creative Behavior, 38*(1), 13-34.

Fiore, S. M., Rosen, M. A., Smith-Jentsch, K. A., Salas, E., Letsky, M., & Warner, N. (2010). Toward an understanding of macrocognition in teams: Predicting processes in complex collaborative contexts. *Human Factors, 52*(2), 203-224.

Fu, K., Cagan, J., & Kotovsky, K. (2010). Design Team Convergence: The Influence of Example Solution Quality. *Journal of Mechanical Design, 132*(11), 111005.

Full Belly Project. (2010). Full Belly Project - Nut Sheller. Retrieved October 1, 2011, from http://www.thefullbellyproject.org/Products/UniversalNutSheller.aspx

Hayes, J. R. (1989). Cognitive processes in creativity. In J. A. Glover, R. R. Ronning & C. R. Reynolds (Eds.), *Handbook of Creativity* (pp. 135-145). New York: Plenum Press.

Hennessey, B. A., & Amabile, T. M. (1988). The conditions of creativity. In R. J. Sternberg (Ed.), *The nature of creativity: Contemporary psychological perspectives.* (pp. 11-38). New York, NY: Cambridge University Press.

Herman, A., & Reiter-Palmon, R. (2011). The effect of regulatory focus on idea generation and idea evaluation. *Psychology of Aesthetics, Creativity, and the Arts, 5*(1), 13-20.

Hernandez, N. V., Shah, J. J., & Smith, S. M. (2010). Understanding design ideation mechanisms through multilevel aligned empirical studies. *Design Studies, 31*(4), 382-410.

Isaak, M. I., & Just, M. A. (1995). Constraints on thinking in insight and invention. In R. J. Sternberg & J. E. Davidson (Eds.), *The Nature of Insight* (pp. 281-325). Cambridge, MA: MIT Press.

Jansson, D. G., & Smith, S. M. (1991). Design fixation. *Design Studies, 12*(1), 3-11.

Johnson-Laird, P. N. (2006). Models and heterogeneous reasoning. *Journal of Experimental & Theoretical Artificial Intelligence, 18*(2), 121-148.

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

Kim, M. H., Kim, Y. S., Lee, H. S., & Park, J. A. (2007). An underlying cognitive aspect of design creativity: Limited Commitment Mode control strategy. *Design Studies, 28*(6), 585-604.

Klimoski, R., & Mohammed, S. (1994). Team mental model: construct or metaphor? *Journal of Management, 20*(2), 403-437.

Kohn, N. W., & Smith, S. M. (2011). Collaborative fixation: Effects of others' ideas on brainstorming. *Applied Cognitive Psychology, 25*(3), 359-371.

Kotovsky, K., Hayes, J. R., & Simon, H. A. (1985). Why are some problems hard? Evidence from Tower of Hanoi. *Cognitive Psychology, 17*(2), 248-294.

Landauer, T. K. (1998). Learning and Representing Verbal Meaning: The Latent Semantic Analysis Theory. *Current Directions in Psychological Science, 7*(5), 161-164.

Landauer, T. K. (2007). LSA as a theory of meaning. In T. K. Landauer, D. S. McNamara, S. Dennis & W. Kintsch (Eds.), *Handbook of latent semantic analysis.* (pp. 3-34). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to Latent Semantic Analysis. *Discourse Processes, 25*(2-3), 259-284.

Leggett, K. L., Putman, V. L., Roland, E. J., & Paulus, P. B. (1996). *The effects of training on performance in group brainstorming*. Paper presented at the annual meeting of the Southwestern Psychological Association.

LePine, J. A. (2005). Adaptation of Teams in Response to Unforeseen Change: Effects of Goal Difficulty and Team Composition in Terms of Cognitive Ability and Goal Orientation. *Journal of Applied Psychology, 90*(6), 1153-1167.

Lim, B.-C., & Klein, K. J. (2006). Team mental models and team performance: a field study of the effects of team mental model similarity and accuracy. *Journal of Organizational Behavior, 27*(4), 403-418.

Linsey, J. S., Clauss, E. F., Kurtoglu, T., Murphy, J. T., Wood, K. L., & Markman, A. B. (2011). An Experimental Study of Group Idea Generation Techniques: Understanding the Roles of Idea Representation and Viewing Methods. *Journal of Mechanical Design, 133*(3).

Linsey, J. S., Green, M. G., Murphy, J. T., Wood, K. J., & Markman, A. B. (2005). *"Collaborating to Success": An experimental study of group idea generation techniques*. Paper presented at the ASME 2005 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference.

Lonergan, D. C., Scott, G. M., & Mumford, M. D. (2004). Evaluative aspects of creative thought: Effects of appraisal and revision standards. *Creativity Research Journal, 16*(2&3), 231-246.

Lund, K., & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behaviour Research Methods, Instruments, & Computers, 28*(2), 203-208.

Mathieu, J. E., Heffner, T. S., Goodwin, G. F., Salas, E., & Cannon-Bowers, J. A. (2000). The influence of shared mental models on team process and performance. *Journal of Applied Psychology, 85*(2), 273-283.

McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods, 1*(1), 30-46.

McNamara, T. P. (2005). *Semantic priming: Perspectives from memory and word recognition*. New York, NY US: Psychology Press.

Mohammed, S., Ferzandi, L., & Hamilton, K. (2010). Metaphor No More: A 15-Year Review of the Team Mental Model Construct. *Journal of Management, 36*(4), 876-910.

Mumford, M. D., Feldman, J. M., Hein, M. B., & Nagao, D. J. (2001). Tradeoffs between ideas and structure: Individual versus group performance in creative problem solving. *Journal of Creative Behavior, 35*(1), 1-23.

Mumford, M. D., Reiter-Palmon, R., & Redmond, M. R. (1994). Problem construction and cognition: Applying problem representations in ill-defined domains. In M. A. Runco (Ed.), *Problem finding, problem solving, and creativity.* (pp. 3-39). Westport, CT: Ablex Publishing.

Nelson, D., McEvoy, C., & Schreiber, T. (2004). The University of South Florida free association, rhyme, and word fragment norms. *Behavior Research Methods, Instruments, & Computers, 36*(3), 402-407.

Norman, D. A. (1983). Some observations on mental models. In D. Gentner & A. L. Stevens (Eds.), *Mental Models* (pp. 7-14). Hillsdale, NJ: Lawrence Earlbaum Associates.

Osborn, A. F. (1957). *Applied imagination*. Buffalo, NY: Creative Education Foundation.

Otto, K. N., & Wood, K. L. (2001). *Product design : techniques in reverse engineering and new product development*. Upper Saddle River, NJ: Prentice Hall.

Paulus, P. B., Dzindolet, M., & Kohn, N. W. (2011). Collaborative creativity - Group creativity and team innovation. In M. D. Mumford (Ed.), *Handbook of Organizational Creativity* (pp. 327-357). New York: Elsevier.

Paulus, P. B., & Yang, H.-C. (2000). Idea Generation in Groups: A Basis for Creativity in Organizations. *Organizational Behavior and Human Decision Processes, 82*(1), 76-87.

Pincombe, B. (2004). *Comparison of Human and Latent Semantic Analysis (LSA) Judgments of Pairwise Document Similarities for a News Corpus*. Edinburgh South Australia, Australia: DSTO Information Sciences Laboratory.

Putman, V. L., & Paulus, P. B. (2009). Brainstorming, brainstorming rules and decision making. *Journal of Creative Behavior, 43*(1), 23-39.

Quesada, J. (2007). Spaces for problem solving. In T. K. Landauer, D. S. McNamara, S. Dennis & W. Kintsch (Eds.), *Handbook of Latent Semantic Analysis* (pp. 185-203). Mahwah, NJ: Lawrence Erlbaum Associates.

Reber, P. J., & Kotovsky, K. (1997). Implicit learning in problem solving: The role of working memory capacity. *Journal of Experimental Psychology: General, 126*(2), 178-203.

Redmond, M. R., Mumford, M. D., & Teach, R. (1993). Putting creativity to work: Effects of leader behavior on subordinate creativity. *Organizational Behavior and Human Decision Processes, 55*, 120-151.

Reiter-Palmon, R., Herman, A. E., & Yammarino, F. J. (2008). Creativity and cognitive processes: Multi-level linkages between individual and team cognition. In M. D. Mumford, S. T. Hunter & K. E. Bedell-Avers (Eds.), *Multi-level Issues in creativity and innovation* (Vol. 7, pp. 203-267). Bingley, U.K.: Emerald.

Reiter-Palmon, R., & Robinson, E. J. (2009). Problem Identification and Construction: What Do We Know, What Is the Future? *Psychology of Aesthetics Creativity and the Arts, 3*(1), 43-47.

Reiter-Palmon, R., Wigert, B., & de Vreede, T. (2011). Team creativity and innovation: The effect of group composition, social processes, and cognition. In M. D. Mumford (Ed.), *Handbook of Organizational Creativity*. New York: Academic Press.

Rietzschel, E. F., Nijstad, B. A., & Stroebe, W. (2006). Productivity is not enough: A comparison of interactive and nominal brainstorming groups on idea generation and selection. *Journal of Experimental Social Psychology, 42*(2), 244-251.

Salas, E., & Cannon-Bowers, J. A. (1997). Methods, tools, and strategies for team training. In M. Quinones & A. Ehrenstein (Eds.), *Training for a rapidly changing workplace* (pp. 249-279). Washington, DC: American Psychological Association.

Scott, G., Leritz, L. E., & Mumford, M. D. (2004). The effectiveness of creativity training: A quantitative review. *Creativity Research Journal, 16*(4), 361-388.

Shah, J. J., Vargas-Hernandez, N., Summers, J. D., & Kulkarni, S. (2001). Collaborative sketching (c-sketch)-an idea in generation technique for engineering design. *Journal of Creative Behavior, 35*(3), 168-198.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86*(2), 420-428.

Silveira, J. M. (1971). *Incubation: The Effect of Interruption Timing and Length on Problem Solution and Quality of Problem Processing.* University of Oregon, Eugene, OR.

Smith-Jentsch, K. A., Mathieu, J. E., & Kraiger, K. (2005). Investigating Linear and Interactive Effects of Shared Mental Models on Safety and Efficiency in a Field Setting. *Journal of Applied Psychology, 90*(3), 523-535.

Smith, C. M., Bushouse, E., & Lord, J. (2010). Individual and group performance on insight problems: The effects of experimentally induced fixation. *Group Processes & Intergroup Relations, 13*(1), 91-99.

Smith, S. M., & Blankenship, S. E. (1989). Incubation effects. *Bulletin of the Psychonomic Society, 27*(4), 311-314.

Smith, S. M., & Blankenship, S. E. (1991). Incubation and the persistence of fixation in problem solving. *American Journal of Psychology, 104*(1), 61-87.

Smith, S. M., Ward, T. B., & Schumacher, J. S. (1993). Constraining effects of examples in a creative generation task. *Memory & Cognition, 21*(6), 837-845.

Strait, M. J., Haynes, J. A., & Foltz, P. W. (2000). Applications of Latent Semantic Analysis to Lessons Learned Systems. In D. W. Aha & R. Weber (Eds.), *Intelligent Lessons Learned Systems: Papers from the 2000 Workshop*. Menlo Park, CA: AAAI Press.

Wallas, G. (1926). *The Art of Thought*. London: J. Cape.

Wood, M. D., Hocker, E. M., Hunter, S. T., & Ligon, G. S. (2011). Impact of strength- versus problem-focus in idea revision. In L. DellaPietra (Ed.), *Perspectives on Creativity* (Vol. 2). New York: Cambridge Scholars Publishing.

Woodworth, R. S., & Schlosberg, H. (1962). *Experimental psychology* (Rev. ed.). New York: Rinehart & Winston.

Youmans, R. J. (2011). The effects of physical prototyping and group work on the reduction of design fixation. *Design Studies, 32*(2), 115-138.

APPENDIX A

DESIGN TASK

In places like Haiti and certain West African countries, peanuts are a significant crop. Most peanut farmers shell their peanuts by hand, an inefficient and labor-intensive process. Our goal is to build a low-cost, easy to manufacture peanut sheller targeted at individuals and small cooperatives that will increase the productivity of the peanut farmers. Further, this peanut sheller should be manufacturable with materials that are readily available in the target communities. Our target throughput is approximately 50 kg (110 lbs) per hour.

Customer Needs

Must remove the shell with minimal damage to the peanuts.

Electrical outlets are not available as a power source.

A large quantity of peanuts must be quickly shelled.

Low cost.

Easy to manufacture.

APPENDIX B

SELF-REPORT PROMPTS

First prompt

In your own words, without consulting the instructions sheet or any other materials, please describe your current approach to a solution or your current best solution and how it works in three to five sentences.

Subsequent prompts

At this point in your solving, please describe your current best solution or solution approach and how it works in three to five sentences. Please use your own words, without consulting the instructions sheet or any other materials.

APPENDIX C

QUALITY PUGH CHART FOR EXPERIMENT 1A

Rating	much worse -2	worse -1	similar to datum 0	better +1	much better +2
Cost	Made of expensive materials & complicated components, especially as it concerns the design's structure.	Mostly expensive material/components, with a few inexpensive material/components.	Some inexpensive materials (brick, aluminum mesh) & some expensive materials (bicycle, steel plate). $100-120 total.	Many inexpensive material/components & only a few expensive material/components.	Made mostly of inexpensive materials, some of which are scavenged (e.g., sticks, rocks).
Removes shell w/o damaging peanut	Risk of damage to most nuts from cutting unshelled nuts in half or pulverizing nuts to remove shell.	Damage to many nuts.	Some damage from pressing or smashing unshelled peanuts, & may destroy some nuts.	Damage to a few nuts, & very few are destroyed.	Minimal damage. Machine is calibrated to nut size or nut fruit escapes shelling process once shelled.
Separates shell & nut	Nuts & cracked shells are not separated in the output.	Design tries to separate shells & nuts, but many shells are still mixed in with nuts	Mesh shaker table lets nuts fall to a collecting bin, catching shells in the process. Some shells may still be mixed with nuts.	A few shells are mixed in with nuts, but most nuts and shells are separated.	Nearly all of the shells are separated from the nut, e.g., using several sieve screens serially.
Amt. of peanuts shelled per hr	Less than 10 lbs (5 kg) per hr, similar, or slightly faster than, hand-shelling. Production is arrhythmic & regularly stopped.	~30 lbs (15 kg) per hr, with arrhythmic production and frequent stops.	~65 lbs (30 kg) per hr, assuming press cracks ~2lbs of unshelled nuts at once. Production stopped regularly to reset device.	~125 lbs (55 kg) per hour, some production stops to reset the device.	More than 200 lbs (90 kg) per hr, designed for few stops in production.
Energy source required to operate	Energy source that is expensive to purchase or transport, e.g., gasoline, electricity.	Energy source that is less expensive and/or locally available, e.g., fire from wood or straw.	Human power, inexpensive and locally available.	Animal power, quite inexpensive and locally available.	Renewable energy that is practically free and locally available, e.g., water or wind power.
Amt. of energy required	5+ at medium exertion, or non-human equivalent (e.g., ~200 watt electric motor).	3-4 people at medium exertion, OR 2 at greater than medium exertion, OR equivalent.	2 people at medium exertion: 1 for cracker, 1 for sorting assembly, or non-human equivalent.	1 person at medium exertion, OR 2 people at less than medium exertion, OR equivalent.	1 person at less than medium exertion, or non-human equivalent.
Feasibility	Very low feasibility. Most or all parts are machined or imported, & may be difficult to transport because of their size or weight. Training required for assembly/operation.	Many parts are machined or imported, & a few may be difficult to transport. Some training required for assembly/operation.	Some materials are cheap & can be locally sourced (e.g., brick); others (e.g., bicycle frame, bearings) are expensive &/or require importing. Assembly or operation training may be needed.	Many materials are locally sourced, while only a few are expensive or require importing. Some training may be needed for operation or assembly.	Most materials are readily available, & can be assembled & operated with minimal training.
Size/ Portability	Permanent or near-permanent structure. Hard to move/disassemble.	Can be disassembled & moved with difficulty.	Can be disassembled & moved with some time & effort.	Easily disassembled with little effort. More than one person needed to move.	Requires almost no disassembly & is easily transported by one person.
# of people needed to operate	Full attention of 4+ people required.	Full attention of 3 people required.	Full attention of 2 people required to operate.	Full attention of 1 person required.	Less than full attention of one person required.
Time to set up & build	2 weeks.	6-8 days	3-4 days.	1-2 days.	A few hours.

APPENDIX D

PARKING PROBLEM

Problem – Parking on Campus

Parking on college campuses is a difficult problem, and Carnegie Mellon is no exception. This problem presents a real challenge to campus planners. Develop a way to address the issue of parking availability at CMU.

APPENDIX E

QUALITY DECISION MATRIX FOR EXPERIMENT 1B

Rating	much worse -2	worse -1	average 0	better +1	much better +2
Overall Quality	Low quality relative to other plans. Not much consideration went into solution. It is not likely to succeed, and has multiple unintended consequences.	Below average quality relative to other plans. A few logical & thoughtful elements, but also a few unintended consequences.	Average quality relative to other plans. Solution has some thoughtful & logical elements. Could be effective, but may have unintended consequences.	Better than average quality relative to other plans. Solution has several thoughtful & logical elements. Would likely be effective, with only a couple unintended consequences.	High quality relative to other plans. Many thoughtful & logical elements. Would certainly be effective if implemented, with limited unintended consequences.
Parking availability	Much worse than other plans. Much worse than it is currently.	Worse than other plans. Worse availability than it is currently.	Similar to other plans. No different or slightly better than current availability.	Better than other plans. Better than current availability.	Much better than other plans. Much better than current availability.
Parking need	Much worse than other plans. The average person needs a car to get to campus much more than they do now.	Worse than other plans. The average person needs a car to get to campus more than they do now.	Similar to other plans. The need for the average person to drive a car to campus is similar to current need.	Better than other plans. The average person needs a car to get to campus less than they do now.	Much better than other plans. The average person needs a car to get to campus much less than they do now.
Cost to university	Costs much more than other plans, e.g., building an underground subway.	Costs more than other plans, e.g., building a new parking garage.	Similar cost to other plans, e.g., adding to an existing garage.	Costs less than other plans, e.g., developing a web app, like ParkPGH.	Costs much less than other plans, e.g., buying several bike racks.
Space requirements	Much more space than other plans, e.g., adding several new lots/garages.	More space than other plans, e.g., adding a new parking garage or lot.	Similar to other plans, e.g., adding levels above/below an existing structure.	Less space than other plans, e.g., expanding bike racks.	Much less than other plans, e.g., no change to current physical infrastructure.
Aesthetics	Much worse than other plans, solution creates a large eyesore on campus.	Worse than other plans, solution adds a feature on campus that is not aesthetically pleasing.	Similar to other plans, solution is aesthetically neutral.	Better than other plans, solution is aesthetically pleasing.	Much better than other plans, solution beautifies & compliments adjacent architecture & scenery.
Community satisfaction	Much worse than other plans, campus community hates the plan and discourages friends/colleagues from participating.	Worse than other plans, community does not like plan.	Similar to other plans, satisfaction is similar to current solution.	Better than other plans, campus community likes the plan.	Much better than other plans, campus community likes it encourages friends/colleagues to participate

APPENDIX F

ORIGINALITY DECISION MATRIX FOR EXPERIMENT 1B

Rating	much worse -2	worse -1	average 0	better +1	much better +2
Originality	Very unoriginal, similar to most other plans. Addresses the issue in a totally mundane way.	Worse than average originality, similar to many other plans. Addresses the issue in an unsurprising way.	Average originality, similar to some other plans with some unique elements. May address the issue in a surprising way.	Better than average originality, similar to a few other plans. Addresses the issue in a surprising way.	Very original, only 1-2 other plans are similar. Addressed the issue from a totally new angle.

APPENDIX G

SUSTAINABILITY PROBLEM

Instructions

Consider the problem below. Please read the description carefully. You will be given twenty-eight minutes total [*working individually / with your group*] to create a solution to this problem. You will be asked to **describe the solution** and the **steps necessary to accomplish** that solution. There will be four discrete seven minute solving sessions, and one session at the end to write up your solution. Between solving sessions, you will be asked to write a short paragraph on the computer. Your goal is to create **one optimal solution**. Your final solution and all written materials that you generate will be collected at the end of the session.

Use the provided blank pages of paper to record your solutions. An adequate solution should include a few sentences describing the solution and how it works so other people can understand it. Include a sketch of the solution and labels of major elements if helpful. Please feel free to record any thoughts or comments that you might have as you develop each solution. You will be able to refer back to these notes and the problem statement below as you work on the problem, but will be asked to set these aside when writing on the computer.

Problem – Encourage more environmentally sustainable living

Many people want to do more to decrease their environmental footprint as concerns about climate change intensify and become more widespread. One potentially important way to promote environmental sustainability is by encouraging individuals to make changes to the way they live and the environments they live within. General concern for one's own and others wellbeing provides people with some motivation to change behavior in more environmentally sustainable ways. However, people can often have difficulty identifying lifestyle changes that are more sustainable despite their best efforts, and often feel powerless to change their living environments and societal structures at large.

Our goal is to develop a solution, and the steps necessary to accomplishing that solution, which facilitates more environmentally sustainable living at the individual and societal levels. Most people should be able to adopt the solution once available, and it should change their behavior in a way that is more environmentally sustainable than current practice. In addition, the advantages and disadvantages of the solution should be considered.

Problem requirements

- Promote environmentally sustainable living

- Can be adopted by most people

- Changes behavior in a more environmentally sustainable way

- Maximizes number and impact of advantages

- Minimizes number and impact of disadvantages

APPENDIX H

WORK SESSION INSTRUCTIONS

SESSION	REPRESENTATION CONSTRUCTION CONDITION	
	instruction	**no-instruction**
1st	Spend the next seven minutes [*working independently / with your group*] on thinking about alternate ways of looking at the problem you saw earlier. Do this by stating the problem in as many different ways as you can. Write each restatement in the space below.	Spend the next seven minutes [*working independently / with your group*] on developing a solution to the problem you saw earlier. Be sure to consider how the solution would be accomplished. Show your work on the solution and its implementation in the space provided below.
2nd	Spend the next seven minutes [*working independently / with your group*] on developing a solution to the problem you saw earlier. Be sure to consider how the solution would be accomplished. Show your work on the solution and its implementation in the space provided below.	Spend the next seven minutes [*working independently / with your group*] on developing a solution to the problem you saw earlier. Be sure to consider how the solution would be accomplished. Show your work on the solution and its implementation in the space provided below.
3rd & 4th	Spend the next seven minutes [*working independently / with your group*] on developing a solution to the problem you saw earlier. Be sure to consider how the solution would be accomplished. Show your work on the solution and its implementation in the space provided below.	Spend the next seven minutes [*working independently / with your group*] on developing a solution to the problem you saw earlier. Be sure to consider how the solution would be accomplished. Show your work on the solution and its implementation in the space provided below.

SELF-REPORT FORM FOR EXPERIMENT 2

In your own words, without consulting the instructions sheet or any other materials, please describe your current best solution and how it works in three to five sentences in the space below. Try to be as complete as possible in your description.

```
Write your entry here...
```

Once you have provided your description, use the five-point scale below to rate the quality of your solution.

◯ Very Poor   ◯ Poor   ◯ Fair   ◯ Good   ◯ Very Good

[Submit]

APPENDIX J

TEAM COHERENCE RESULTS FOR EXPERIMENT 2

In order to test the hypotheses that both team interaction structure (H_1) and representation construction instructions (H_3) would influence team coherence, a repeated-measures factorial ANOVA was conducted on team coherence scores for each team with team interaction structure (nominal or interacting) and representation construction instructions condition (instruction or no-instruction) entered as separate between-groups effects, while Time was entered as a within-groups effect. This analysis identified a significant main effect of team interaction structure, $F(1,36) = 11.723$, $p = 0.002$, $\eta^2 = 0.246$, where interacting teams tended to have higher coherence over the course of the experiment than nominal teams (Figure J.1). This main effect was qualified by a statistical interaction between Time and team interaction structure, $F(4,144) = 7.514$, $p < 0.001$, $\eta^2 = 0.173$. Team coherence for interacting teams increased over time, while team coherence for nominal teams tended to stay relatively flat over time (Figure J.2). This trend of increasing coherence for interacting teams is consistent with Experiment 1A and 1B, and the higher overall coherence for interacting teams is in support of H_1.
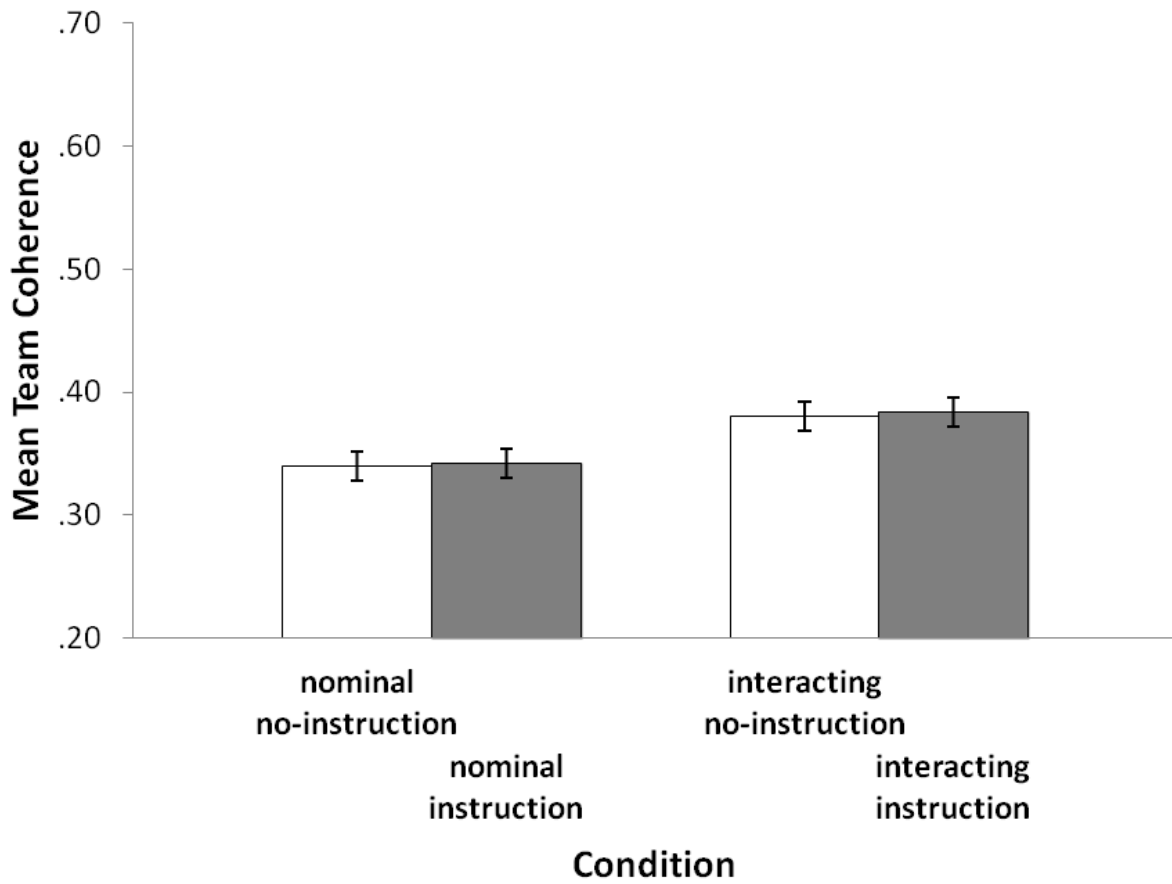
Figure J.1 Experiment 2 Mean Team Coherence.

Bar grouping indicates team interaction structure (nominal left, interacting right), bar color indicates representation construction condition (no-instructions white, instructions grey). Error bars indicate 1SE.
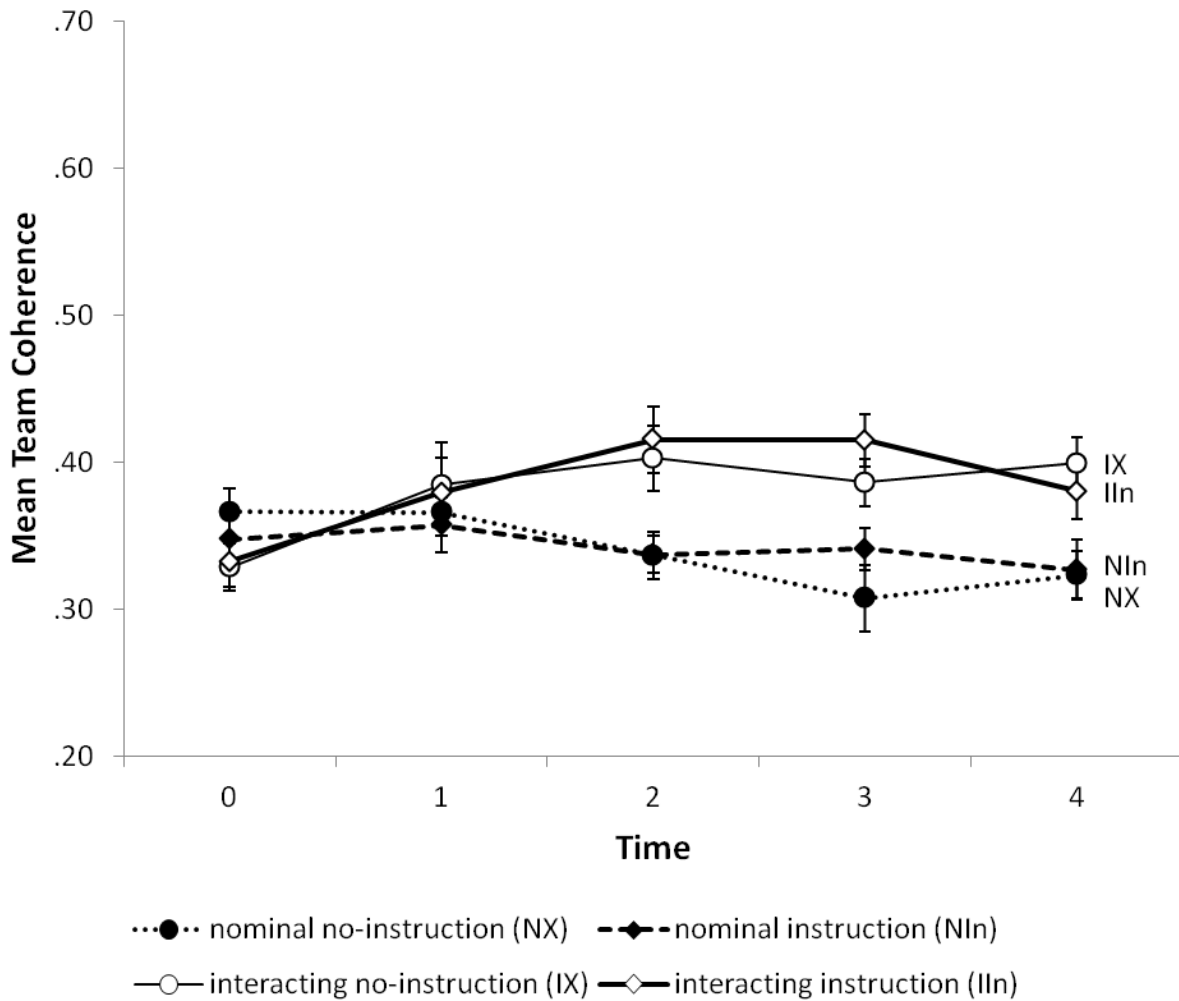
Figure J.2 Experiment 2 Team Coherence Over Time.

Each line represents a different team interaction structure by representation construction instruction condition combination. White markers and solid lines indicate interacting teams; black markers and dashed lines indicate nominal teams. Teams given representation construction instructions have square markers and heavy lines; no-instruction teams have round markers and light lines. Error bars indicate 1SE.

APPENDIX K

QUALITY DECISION MATRIX FOR EXPERIMENT 2

Rating	much worse -2	worse -1	average 0	better +1	much better +2
Overall Quality	Low quality relative to other plans.	Below average quality relative to other plans.	Average quality overall relative to other plans.	Better than average quality relative to other plans.	High quality relative to other plans.
Advantages	There are no benefits to plan compared to status quo.	There is only a slight benefit to plan compared to status quo.	There are a few or small benefits to plan compared to status quo.	There are several benefits and/or they are substantial compared to status quo.	There are many benefits and/or they are overwhelmingly beneficial.
Disadvantages	There are many disadvantages and/or they are quite problematic.	There are several disadvantages and/or they are substantial compared to status quo.	There are a few disadvantages to plan compared to status quo.	There is only a slight disadvantage to plan compared to status quo.	There are no disadvantages to plan compared to status quo.
Cost	Plan is quite expensive compared to others.	Plan is more expensive than other plans.	Plan costs about the same as others in the dataset.	Plan is less expensive than other plans.	Plan is quite cheap compared to other plans.
Feasibility	Plan is infeasible and extremely difficult to implement.	Plan is difficult to implement, potential roadblocks are numerous and/or substantial	Plan has some hurdles to implementation, but achievable.	Plan is feasible with only a few considerations or roadblocks.	Plan is quite feasible and can be implemented easily.
Changes behavior	Plan does not change individuals' behavior at all.	Plan does little to change individuals' behavior, other plans provide more change.	Plan causes some change in individuals' behavior, similar with other plans.	Plan causes substantial behavior change, more than other plans.	Plan creates paradigmatic change in individuals' behavior, more than most other plans.
Stickiness / adoption	People are much less likely to adopt plan compared to others.	People are less likely to adopt plan compared to other plans.	The likelihood plan will be adopted by people is similar to other plans.	People are more likely to adopt plan compared to other plans.	People are much more likely to adopt plan compared to others plans.
Onset	Plan takes much longer to implement than others, a few years or more.	Plan takes longer to implement than others, within a year or two.	The time to implement plan is similar to others in the dataset, within weeks or months.	Plan can be implemented more quickly than others, a few weeks.	Plan can be implemented much more quickly vs. others, a few days.
Longevity / duration	Plan's effects are ephemeral versus others.	Plan's effects would not last as long as other plans.	Effects of plan would last as long as other plans.	Plan's effects would last longer than other plans.	Plan's effects would last much longer than others.
Public sentiment	Public would hate this plan versus other plans.	Public would like this plan less than other plans.	Public sentiment is similar for others in the dataset.	Public would like this plan more than other plans.	Public would love this plan compared to other plans.

APPENDIX L

ORIGINALITY DECISION MATRIX FOR EXPERIMENT 2

Rating	much worse -2	worse -1	average 0	better +1	much better +2
Originality	Very unoriginal, similar to most other plans. Addresses the issue in a totally mundane way.	Less original than average, similar to many other plans. Addresses the issue in an unsurprising way.	Average originality, similar to some other plans. May address the issue in a surprising way.	Better than average originality, similar to a few other plans. Addresses the issue in a surprising way.	Very original, only 1-2 other plans are similar. Addressed the issue from a totally new angle.

APPENDIX M

UNIVARIATE ANOVAs OF AGGREGATE RELIABLE QUALITY CRITERIA

FOR EXPERIMENT 2

Follow-up univariate ANOVAs on each dependent variable identified a significant main effect of instructions on Cost, $F(1,116) = 4.104$, $p = 0.045$, $\eta^2 = 0.034$, and Public Sentiment, $F(1,116) = 4.135$, $p = 0.044$, $\eta^2 = 0.034$. Teams who received instructions to restate the problem tended to produce solutions that were less expensive and more likely to be accepted by the public than no-instruction teams (Figure M.1a-b). Follow-up univariate analyses identified a significant main effect of team interaction structure on Feasibility, $F(1,116) = 7.654$, $p = 0.007$, $\eta^2 = 0.062$, Changes Behavior, $F(1,116) = 5.194$, $p = 0.024$, $\eta^2 = 0.043$, and Onset, $F(1,116) = 9.818$, $p = 0.002$, $\eta^2 = 0.078$. Marginally significant effects of team interaction structure were also identified with respect to Overall Quality, $F(1,116) = 2.983$, $p = 0.087$, $\eta^2 = 0.025$, and Cost, $F(1,116) = 3.805$, $p = 0.053$, $\eta^2 = 0.032$ (Figure M.1a, c-f). Interacting teams tended to produce final problem solutions that were more feasible, more likely to have a more immediate impact, and less expensive than problem solutions proposed by nominal teams. Final problem solutions from interacting teams were also slightly lower in Overall Quality, and were less likely to change behavior compared to nominal team solutions.
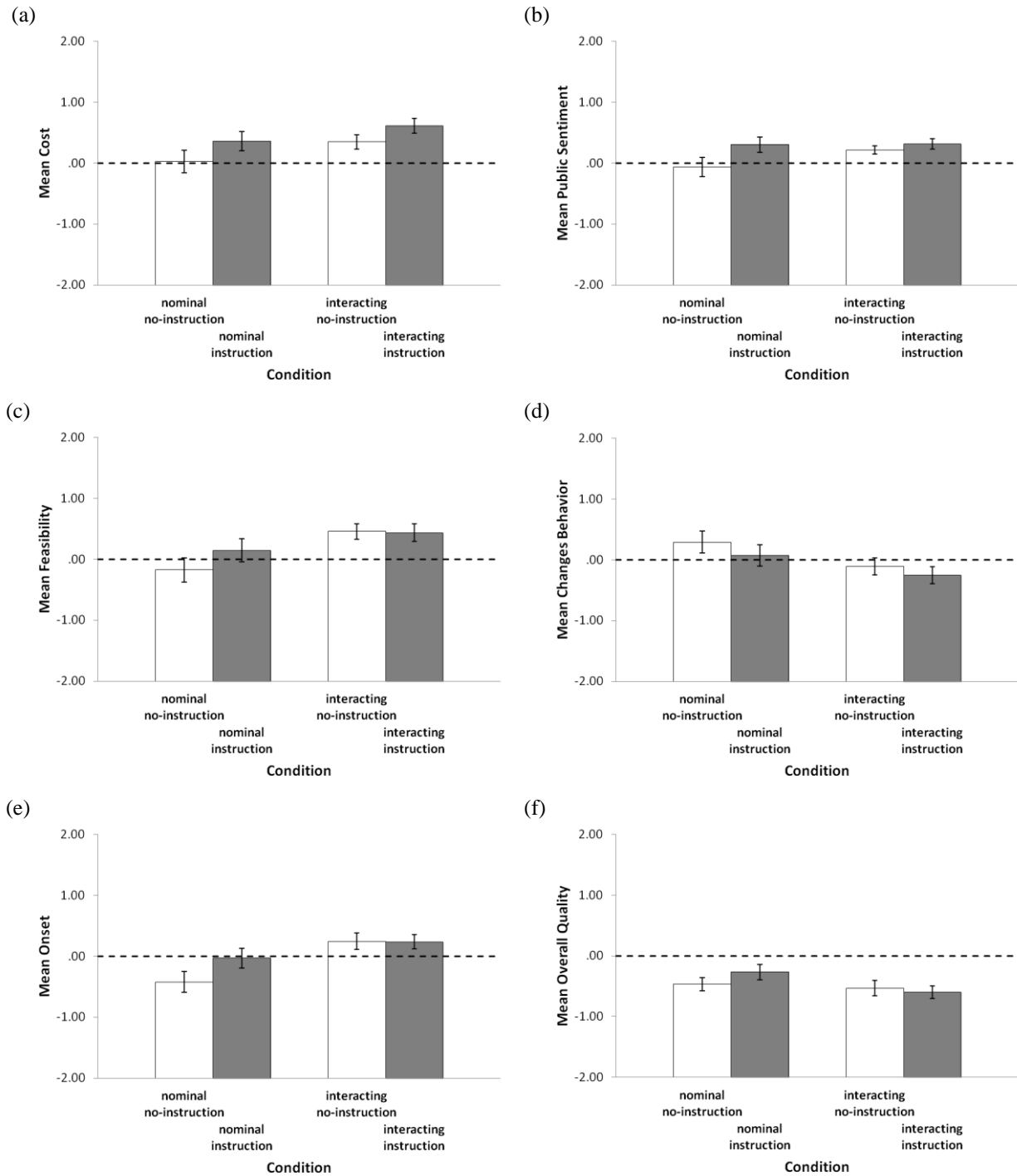
Figure M.1 Experiment 2 Univariate Effects of Reliable Aggregate Quality MANOVA.

Bar grouping indicates team interaction structure (nominal left, interacting right), bar color indicates representation construction condition (no-instructions white, instructions grey). Error bars indicate 1SE.

132

APPENDIX N

INTERIM SOLUTION EVALUATION FOR EXPERIMENT 2

Self-reports at t_0 and t_2 were evaluated using a similar procedure to that for final problem solutions to see if differences in either the quality or originality of solutions could be identified in the course of the problem solving process, before final problem solutions were actually proposed. These differences in solution quality as a function of team interaction or representation construction instructions may be taken as evidence for H_1 and H_3, respectively. In order to evaluate these interim products, the three trained raters who had conducted final problem solution evaluations applied the Overall Quality and Originality criteria to self-reports. They first completed a training set of 10% of all t_0 and t_2 self-reports to check for interrater agreement. Rater agreement on both Overall Quality and Originality was marginal with respect to the intraclass correlation $\alpha = 0.7$ training criterion used elsewhere in this dissertation. However, when two outlier cases of extreme disagreement were omitted from the training set, intraclass correlations were acceptable for both types of evaluation (Table N.1).

A training session was conducted where these extreme outlier cases were discussed along with several cases where agreement was good in the training set, and raters then proceeded to evaluating the full set of self-reports at t_0 and t_2. One additional criterion "Goodness," intended as a magnitude estimation task in terms of quality, was also added at this time for raters to evaluate. Whereas Overall Quality compares the quality of any particular solution to others in the dataset, Goodness evaluates quality in absolute terms with respect to solutions that may exist out in the real world. Goodness ranges on a scale from "1" (poor) to "10" (great).

Table N.1 Experiment 2 Intraclass Correlations for Interim Evaluations.

	α		
Criterion	*training set*	*training set without outliers*	*full set*
Overall Quality	0.651	0.711	0.253
Originality	0.714	0.712	0.636
Goodness	---	---	0.288

Agreement was quite poor in the full set on both Overall Quality and Goodness, but was acceptable for Originality (Table N.1). While Overall Quality and Goodness results will be presented here, they should be interpreted with caution as low interrater agreement values tend to be suggestive of an unstable measurement construct. Scores on each criterion were entered into separate 2x2x2 mixed ANOVAs, entering team interaction structure and representation construction instructions as separate between-subjects factors. Time was entered as a within-subjects factor with two levels (t_0 and t_2) for all criteria.

With respect to Overall Quality, the analysis identified a main effect of Time, where Overall Quality scores tended to increase slightly over time, $F(1,116) = 12.740$, $p = 0.001$, $\eta^2 = 0.099$. This main effect was qualified by both a statistical interaction of Time and team interaction structure, $F(1,116) = 7.030$, $p = 0.009$, $\eta^2 = 0.057$, as well as a statistical interaction of Time and representation construction instructions, $F(1,116) = 7.581$, $p = 0.007$, $\eta^2 = 0.061$. Overall Quality scores tended to increase from t_0 to t_2 for interacting teams as compared to nominal teams, as well as for teams given representation construction instructions compared to

the no-instruction condition (Figure N.1). No other main effects or statistical interactions were

identified.

The results with respect to Goodness scores were quite similar to those for Overall

Quality. The analysis identified a main effect of Time, $F(1,116) = 9.058$, $p = 0.003$, $\eta^2 = 0.072$,
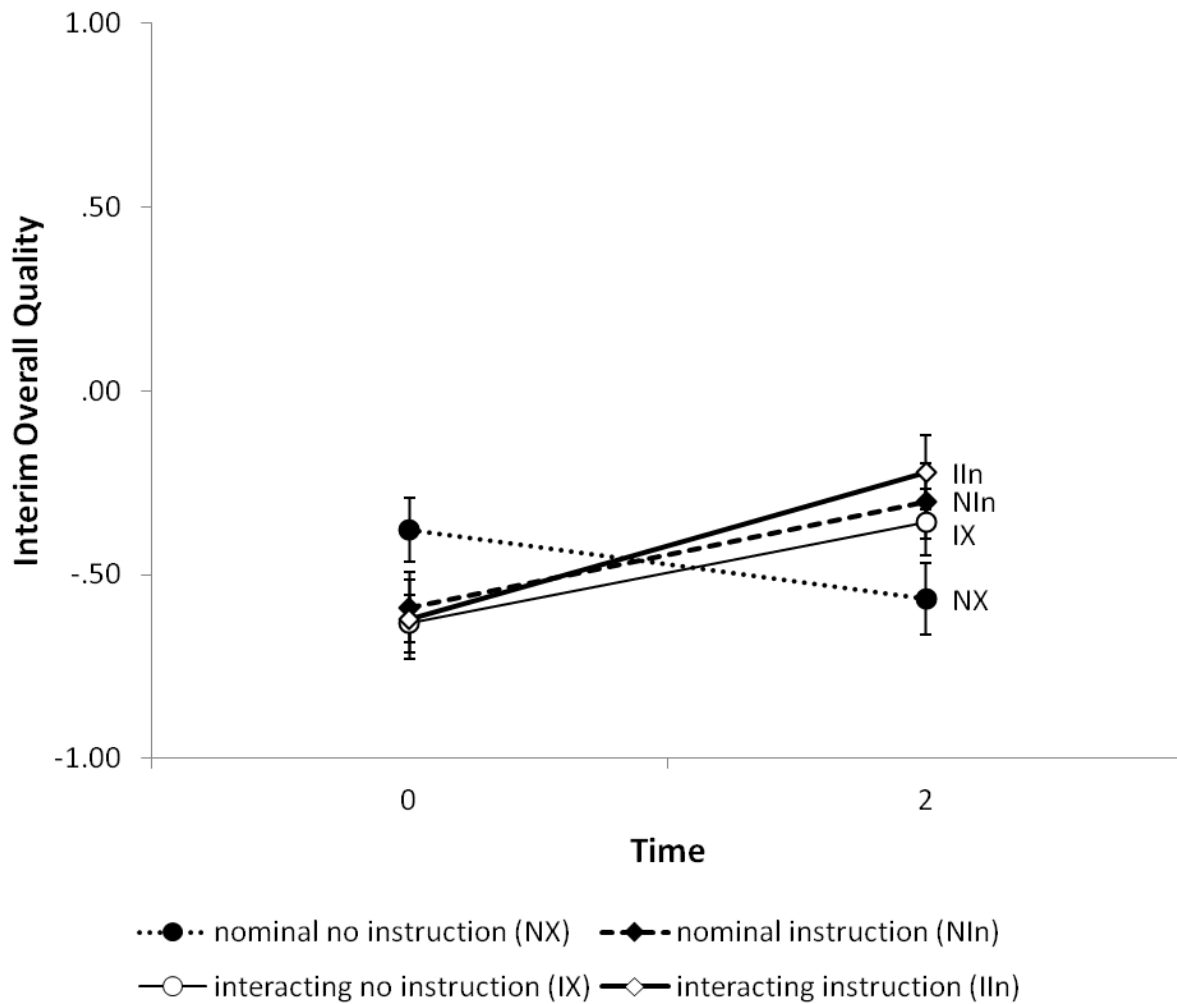

Figure N.1 Experiment 2 Interim Overall Quality.

Each line represents a different team interaction structure by representation construction instruction condition combination. White markers and solid lines indicate interacting teams; black markers and dashed lines indicate nominal teams. Teams given representation construction instructions have square markers and heavy lines; no-instruction teams have round markers and light lines. Error bars indicate 1SE.

where Goodness scores tended to increase over time. The analysis also identified separate statistical interactions between Time and team interaction structure, $F(1,116) = 9.058$, $p = 0.003$, $\eta^2 = 0.072$, as well as between Time and representation construction instruction condition, $F(1,116) = 6.269$, $p = 0.014$, $\eta^2 = 0.051$. As with Overall Quality scores, Goodness scores
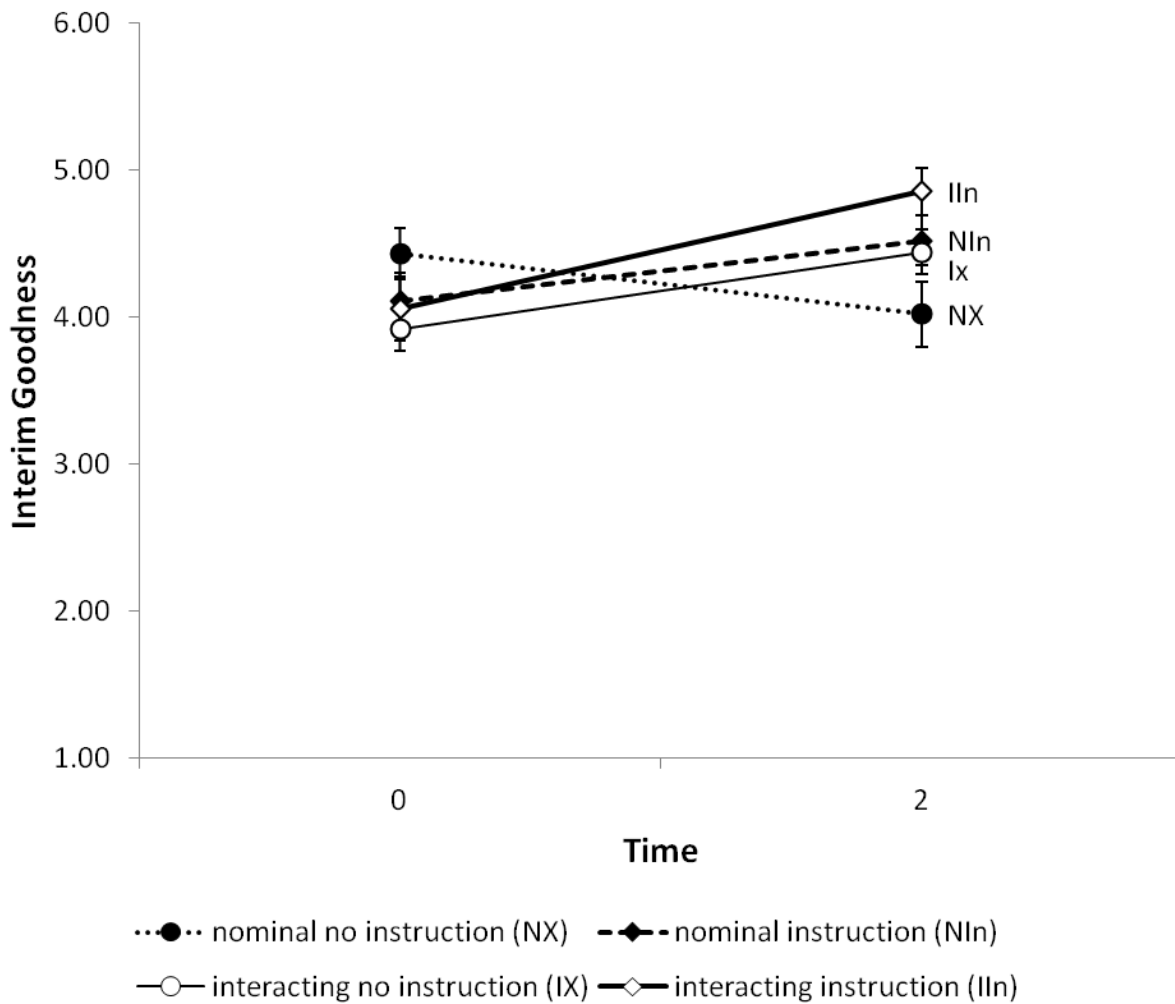

Figure N.2 Experiment 2 Interim Goodness.

Each line represents a different team interaction structure by representation construction instruction condition combination. White markers and solid lines indicate interacting teams; black markers and dashed lines indicate nominal teams. Teams given representation construction instructions have square markers and heavy lines; no-instruction teams have round markers and light lines. Error bars indicate 1SE.
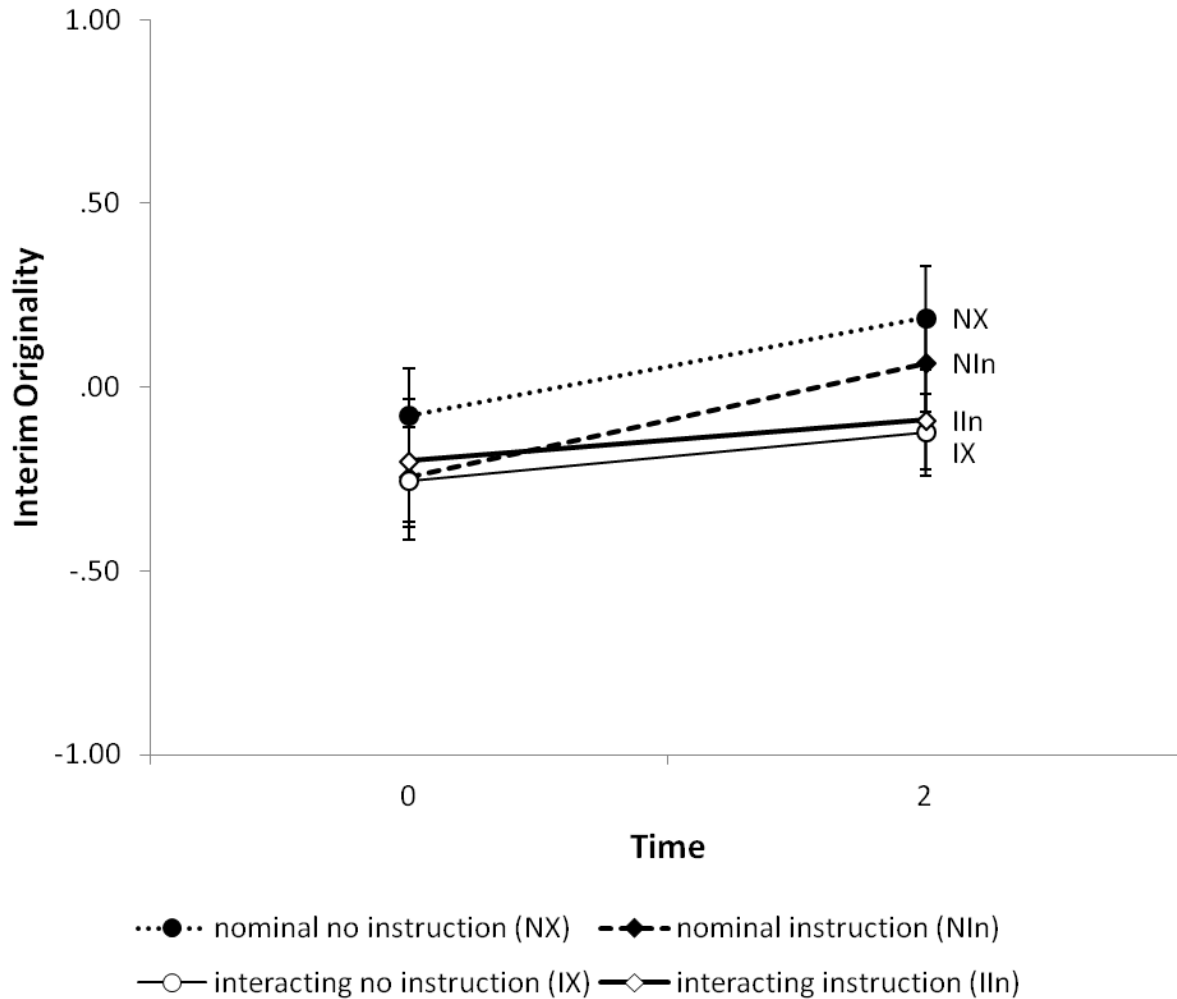
Figure N.3 Experiment 2 Interim Originality.

Each line represents a different team interaction structure by representation construction instruction condition combination. White markers and solid lines indicate interacting teams; black markers and dashed lines indicate nominal teams. Teams given representation construction instructions have square markers and heavy lines; no-instruction teams have round markers and light lines. Error bars indicate 1SE.

tended to increase more for interacting teams than for nominal teams over time, and for

instruction compared to no-instruction teams (Figure N.2). No other main effects or statistical

interactions were identified.

With respect to Originality, the analysis identified a main effect of Time, $F(1,116) = 5.238$, $p = 0.024$, $\eta^2 = 0.043$. No other main effects or statistical interactions were identified. Originality scores tended to increase over time, irrespective of team interaction structure or representation construction instruction condition (Figure N.3).

In sum, interim analyses of Overall Quality and Goodness are consistent with H_1 that team interaction should improve solution quality, and H_3 that representation construction instructions should improve solution quality. Both Overall Quality and Goodness increased from when participants read the problem to about halfway through the experiment, but only for participants who either were able to interact with others team members, or were given instructions to restate the problem. These results should be interpreted cautiously however, as the poor interrater agreement across the set of interim quality evaluations suggests that these measurement scales may not be stable.

In contrast to the results on interim quality assessments, evaluations of interim Originality seemed to suggest an increase in originality for all teams, irrespective of team interaction structure or representation construction instruction condition. Though there is a slight trend toward interacting teams having less original interim solutions compared to nominal teams, this difference does not reach statistical significance. This trend may reflect a process by which nominal team solutions are slowly becoming more original than interacting teams throughout the course of the experiment, but this difference does not become meaningful until a final problem solution is produced (e.g., Figure 4.11).

APPENDIX O

SELF-PERCEIVED QUALITY FOR EXPERIMENT 2

In one additional attempt to evaluate the interim quality of developing solutions, participants were asked to rate the quality of the solution(s) contained in their self-reports throughout the problem solving process on a one to five scale ranging from "very poor" to "very good." These ratings were entered into a 2x2 repeated-measures ANOVA, with team interaction structure and representation construction instruction condition entered as separate between-subjects effects. The analysis identified a significant main effect of Time, $F(4,464) = 30.335$, $p < 0.001$, $\eta^2 = 0.207$. Self-perceived quality of interim solutions tended to increase over time. This main effect was qualified by a significant Time by representation construction instruction statistical interaction, $F(4,464) = 6.277$, $p < 0.001$, $\eta^2 = 0.051$, as well as a marginally significant three-way Time by representation construction instruction by team interaction structure statistical interaction, $F(4,464) = 2.089$, $p = 0.081$, $\eta^2 = 0.018$. Teams given representation construction instructions tended to report greater increases in self-perceived quality of their solutions over time compared to no-instruction teams. However, instruction teams did not have significantly higher self-perceived quality scores over the course of the experiment. In addition, self-perceived quality for nominal teams tended to increase slowly over time, while interacting teams' self-perceived quality scores increased more abruptly over the first half of the experiment before leveling off. The pattern of results suggests that, while instruction teams are not more confident in the quality of their work than nominal teams, instructions do appear to help these teams develop confidence in their solutions.
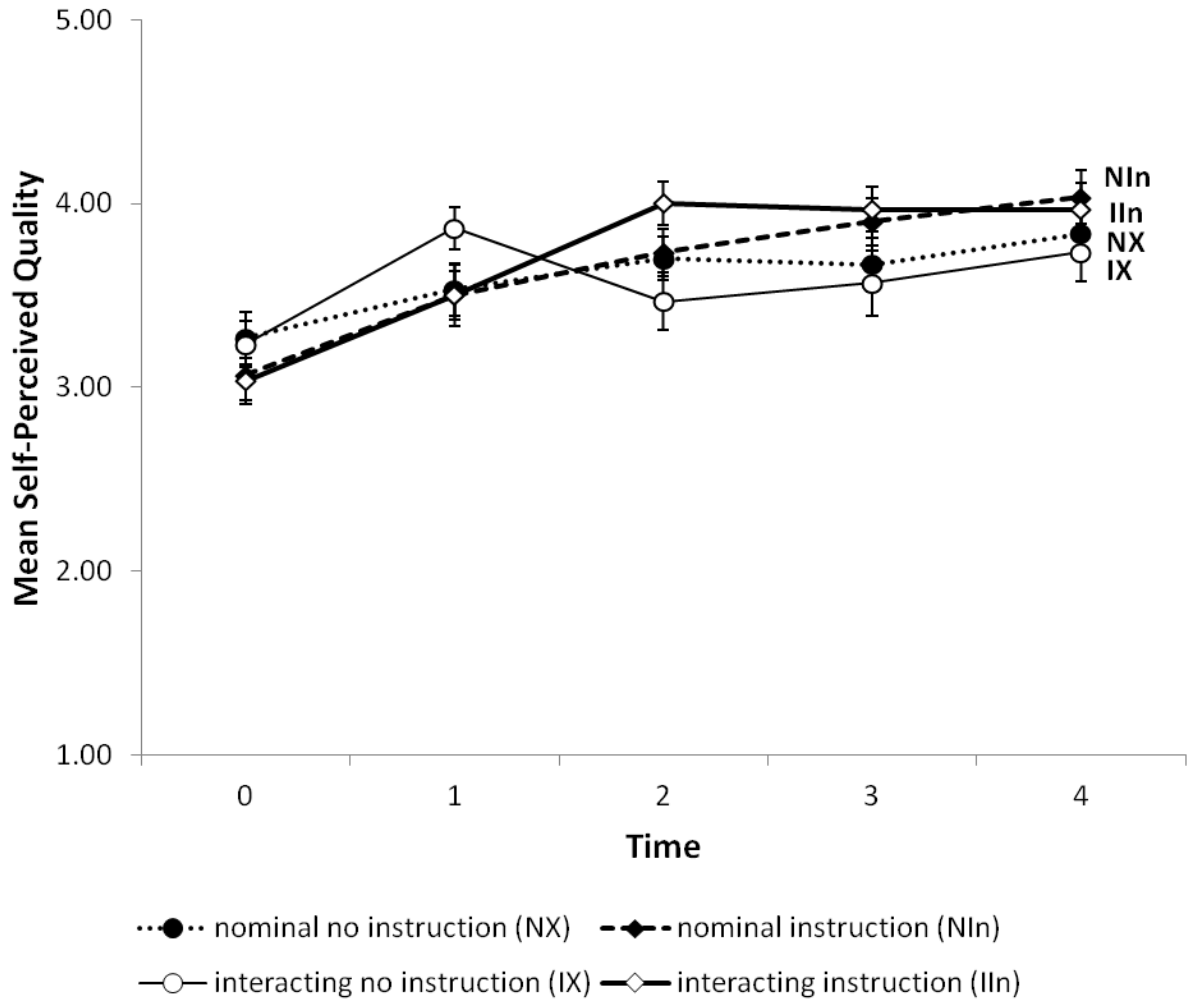
Figure O.1 Experiment 2 Self-Perceived Quality Over Time.

Each line represents a different team interaction structure by representation construction instruction condition combination. White markers and solid lines indicate interacting teams; black markers and dashed lines indicate nominal teams. Teams given representation construction instructions have square markers and heavy lines; no-instruction teams have round markers and light lines. Error bars indicate 1SE.