Effects of Individual versus Group Incentives on Group Problem Solving

by

Lin Chen

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Applied Science
in
Management Sciences

Waterloo, Ontario, Canada, 2010

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Organizations today face complex problems requiring individuals to work in groups to develop insightful solutions efficiently through coordination, sharing, and integration of distributed knowledge. However, very little research has investigated group problem solving, specifically in terms of incentives and problem structure.

This research uses laboratory experiments to investigate the effects of individual versus group goal conflict on collaborative behaviour and performance in group problem solving process. The experiments study 4-person problem solving groups, in which the group solution emerges through coordination and information sharing. The design of the experiment is a 3 by 3 design with two factors, incentive and task structure. Experiments manipulated the relative weights of individual and group rewards using three ratios (0:100, 50:50, 100:0). Three experimental tasks differing in structure were used to investigate the incentive conflict effect on different kind of problems; in particular, problems containing detours and requiring restructuring. One-hundred and sixty-four undergraduate students participated in this study.

The group problem solving process is viewed as a process towards increased structural balance based on Heider's balance theory. This method captures both incremental search and cognitive restructuring during the problem solving process.

Results report the effects of group versus individual goal conflict on group performance and behaviour. Results show that incentive influenced group performance and behaviour by affecting strategies groups used to approach the problem. Individual incentive encouraged the group to focus on the solution state while group incentives encouraged random exploration, and this difference is most significant under the complex problem structure. Results also show that task structure influenced group performance and behaviour by varying the amount of incremental search and restructuring required to solve the problem. Individual incentive weakened difference on performances among three problem structures, while group incentive amplified differences on performance and behaviour among three problem structures.

Acknowledgements

I owe my sincere appreciation to my supervisor, Prof. Rob Duimering, for guiding me through the journey of research and being a true friend in life. I learnt to challenge authorities and think independently, and most of all, found passion in my research.

I owe a big thank you to my partner and my love, Jonathan Aycan, for believing in me and always encouraging me to be positive when confronted with challenges. I must also thank my parents for their unconditional love and support, and for understanding the life path I chose even though that meant I would be far away from home and on the other side of the earth.

I would like to thank my friend, Kristine Meier, who has guided me in the journey of discovering who I am and becoming a stronger person. I also thank my friends, Mina, Paola, Ada, Jennifer, Mohammed, and others whose names are not here for being a wonderful part of my life as a Master's student.

I am also indebted to my thesis committee members, Prof. Frank Safayeni and Prof. Selcuk Onay, for their feedback after reading the initial draft of my thesis.

Last but not least, I would like to thank all the students who participated in the experiments in this thesis, and Prof. Cliff Blake for allowing access to them.

Table of Contents

List of Figures

List of Tables

Chapter 1
Introduction

In today's world, technology has complicated the roles of individuals. The information required to solve problems in complex systems are usually distributed among different sectors such that it is beyond one individual's capacity to solve these complicated problems alone. In many situations people gather together as a group and work on a task; their performance may thus be influenced by external forces such as performance related rewards, with their pay dependent on their own performance, group performance, or both. Existing literature has focused on individual problem solving and research on groups has focused on group decision making models, where the solution to a problem is selected among a set of potential alternate solutions; however, very little research has investigated group problem solving on problems without a known solution—specifically in terms of incentives and problem structure.

This study investigates the relationship between incentives (individual vs. group) and problem-solving performance and behaviours under three different problem structures. An experimental method based on Heider's balance theory (Heider, 1946) from previous research was adopted and revised. In this method, each group member processed part of the information of the whole problem. In order to solve the problem, they communicated with each other verbally and exchanged information physically. Effects of individual and group goal conflict on group performance and behaviour on three different problems were studied in this thesis, which is presented as listed below:

- The current chapter, Chapter 1, introduces the study and reviews prior research on group problem solving and incentives.
- Chapter 2 explains the interest, predictions, and theory behind the current research.
- Chapter 3 explains the methodology, including tasks used in this study which were adopted from previous research and revised to better fit the purpose of this study and experimental stimuli design.
- Chapter 4 presents and analyses the results.
- Chapter 5 discusses the findings of this study and limitations, and suggests future research directions.

1.1 General problem solving

This section provides a review of existing research on general problem solving. According to Lewin's theory (Lewin, 1936), the process of solving a problem can be view as a process of moving from the initial state to the goal state. Lewin's approach to problem solving is presented in the theory of "life space", where the output (behaviours of an individual) is the function of the person and the input (forces and the psychological environment at a certain stage). The process of solving a problem is thus the process of creating a path between the initial state and the goal state.

The classic Information Processing System theory viewed problem solving as a process of exploring and searching in a constructed problem space (Newell, 1972). By exploring and searching in a constructed problem space, the goal state would be eventually found. The variables that influenced this problem space were said to be environmental: the physical task environment, goals obtained by the individual, and the problem itself. However, Gestalt psychologists (Duncker, 1945; Köhler, 1969; Wertheimer, 1982/45) emphasized that a problem should be viewed as a whole instead of separate parts where one can only solve a problem with a complete view of the problem representation. They proposed that a sudden restructuring of the problem space (described as the "Aha!" moment) revealed the goal state. Their argument was demonstrated through "insight problems" where the problem was usually solved in one step associated with reconstitution of the structure. The difference between Information Processing System and the Gestalt view was that the former assumed a constructed problem space while the latter focused on how individuals created and restructured their representation of the problem itself to obtain insightful solutions. Recent research showed that both incremental search and cycles of cognitive restructuring exist in complex real problems (Adejumo, Duimering, & Zhong, 2008).

1.2 Group problem solving

This section provides a background on group problem solving. Group problem solving is different from individual problem solving due to the different magnitude and dimension of interactions. Researchers have shown that groups produce higher performance and increased productivity compared to individuals in some situations. An example would be the letters-to-numbers mathematical problems (Laughlin, Bonner, & Miner, 2002) where the group performance was superior to the best individual's performance. However, research on group problem solving

performance has been limited in terms of the versatility of tasks studied due to the difficulty of building effective and clear methods in a social context.

When investigated in a social context, research has focused on the "group" influence of behaviour—such as social influence, co-action effects, negotiation behaviours, etc. (Nijstad, 2009)—and the problem is usually a scenario in which solving the problem amounts to selecting a solution from a set of available options. Such research models the process of group decision making based on social decision schemes (Davis, 1973; Parksa & Kerrb, 1999), where a group of individuals have disparate sets of personal preferences, and the solution is a function of elements individuals agree on from a discrete set of choice options. However, in many real problem-solving situations there are no pre-existing solutions available, and group members must share information and create and explore the problem space together in order to gradually move to the solution state. Not much research has empirically investigated how a group of individuals actively works together on interdependent tasks where each individual obtains unique information and has to share information, integrate knowledge and develop an insightful answer to a problem. Moreover, although it is generally acknowledged that both incremental search and restructuring are involved in problem solving, little research has successfully integrated both approaches at the same time.

Recent research conducted by Duimering (Adejumo et al., 2008) and his student Abimbola (Abimbola, 2006) has developed an appropriate method that integrates both Gestalt approaches and information processing approaches, and allows both cognitive restructuring and incremental search to be investigated at the same time in the study of the group problem solving process. They investigated the effects of certain properties of problem structure on group problem solving performance and behaviour. They described a person's cognitive representation of a situation as a graph with relationships among cognitions based on balance theory. Balance was achieved through the incremental restructuring of this picture, and was used to track changes during the problem solving process. Their results showed that groups performed differently under different task structures, and problem solving process was proposed as periods of incremental search punctuated by intermittent periods of restructuring. They have examined structure effects on four-person group problem solving, and detour and restructuring were two major structures we are interested in this thesis.

In this thesis, detour means a path deviated from a direct or obvious way leading to the only solution of a problem; that being said, detour is a path leading to the solution, just not directly. In a detour problem, one must move around an interposed barrier to reach the goal. Detour exist in a task structure with a blind alley, an attractive temporary solution which seems to lead in the right direction towards the solution but actually leads to a behaviour that prevents reaching the solution. Thus the group needs to take the detour path to achieve the right solution. Bavelas (1973) has described this problem in the hen and food story: food is placed in front of a hen, but separated by a transparent fence which blocks the shortest path between the hen and the food. The hen has to first walk along the fence until the end of it and turn back on the other side of the fence to get the food. The shortest path is viewed as a blind alley in this situation. Psychologically, the blind alley creates difficulty because the problem solver has to move in a direction that is away from the goal subjectively, while this is the only way to reach the goal objectively. The recognition that the attractive solution is actually a blind alley can be viewed as a cognitive restructuring process. Once the blind alley is overcome, one would be able to work around the barrier and solve the problem.

Restructuring is a means to reconstitute the perceived problem (Wertheimer, 1982/45). It requires the problem solver to represent the problem in a different way in order to solve the problem. Figure 1 shows an example of restructuring: the 9 dot problem (Kershaw & Ohlsson, 2001).

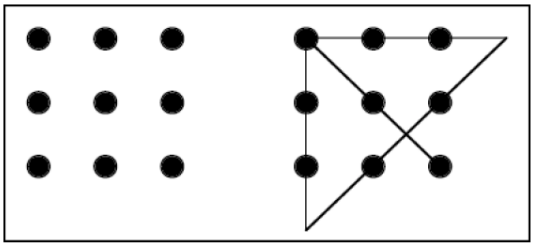


Figure 1: The nine-dot problem

In the nine-dot problem, as shown in Figure 1, one's task is to connect the nine dots in a square area with four straight lines without lifting the pen from the paper or retracing the lines. One needs to realize that the boundaries of the lines are not restricted by the boundaries of the dots (Abimbola, 2006). Once one realizes this, it becomes easy for one to solve the problem. Thus by restructuring the presentation of the problem, one solves the problem.

In conclusion, detour situation is a situation where the problem solver has to take a hidden path to the solution within the search space of a problem, while restructuring involves searching behaviour outside of the problem space. Detour and restructuring can exist at the same time, which constitute an interesting situation. The blind alley, an attractive temporary solution, attracts problem solvers at first sight, and works as a barrier to the right solution. In order to go out of the blind alley, the group needs to restructure the problem so that they view the barrier as a blind alley. However, even when able to overcome the blind alley, they still cannot see the right solution unless they restructure the rest of the problem too. Meanwhile, if they do not restructure the rest of the problem and see the right solution, they would not realize that they are in a blind alley. A question rises in this process, what kind of external force would help the group overcome the detour and encourage restructuring? This question has not been previously addressed, and thus becomes the interest of this study.

1.3 Effect of incentives on group problem solving

This section provides a background on the effect of group versus individual incentives on group task performance in existing literature. According to Lewin's theory (Lewin, 1936), psychological forces perceived by the individual influence the dynamic of the problem solving process. Individuals in a group are usually driven by incentives to solve the problem. How a problem is perceived by an individual at every stage of life space is influenced by incentives. Groups are composed of individuals working on a problem, thus groups are influenced by incentives too. However, no existing theory has successfully addressed the dynamic of group problem solving process, especially in terms of the effect of incentives. The concept of "group" separates incentive into two categories: "group incentive" and "individual incentive". Some groups award individuals by the success of the group, and some award differently among individuals according to the result, or success, of each individual. How each type of incentive influences the group performance and behaviour is of our interest.

Previous researchers have investigated the effect of incentives on group performance on interdependent tasks in industry, in particular, vehicle assembly lines. Fisher (1994) studied a traditional assembly line where the group's performance depended on key workers' effort directly, thus individual incentives was suggested to motivate key workers. Locke and Latham (1990) predicted that individual plus group incentive together result in better performance than either

individual or group incentive alone. Mitchell and Silver (1990) investigated the effects of individual and group goal setting and concluded that a mixed incentive system works best when the accomplishment of individual goals is aligned with the accomplishment of group goal. They also pointed out that when group and individual goals conflict, dysfunctions can result. However, Libby and Thorne (2009) found that when group members could provide useful information about how to perform the task better, group incentives resulted in higher group performance than individual or mixed incentives. One possible explanation is that group incentives encourage knowledge sharing, positive interaction, as well as lower competition within the group.

Experiments completed in the above literature belong to two types: one is empirical industry observation where conditions are hard to control and thus results are hard to validate and vary significantly; another is lab experiments such as building towers using Lego blocks (Goldberg & Maccoby, 1965; Mitchell & Silver, 1990; Young, 1993) or similar tasks such as constructing a sentence together (Guthrie & Hollensbe, 2004). Note that these experiments are not really problem solving tasks since they do not require incremental search within a problem space to solve a problem, there is little risk of entering a blind alley, and conceptual restructuring is not required to perform the task; instead, all one needs to do is to repeat one action such as to accumulate blocks. The path to the goal is thus quite transparent to every member.

Existing research has not studied situations where group and individual goals conflict where individuals in a group are not aware of a clear path leading to the solution state and have to work together to solve the problem as they exchange information, integrate knowledge, and build insightful thoughts. This study aims to investigate the effect of incentives on group problem solving performance and behaviour, in particular, on three different types of problem structure that have been proved to influence group problem solving behaviour (Abimbola, 2006).

Chapter 2
Hypotheses

As discussed earlier, recent research (Abimbola, 2006; Adejumo et al., 2008) has studied a situation where detour and restructuring exist at the same time in the process of group problem solving. A question rises in this process: what kind of external force would help the group overcome the detour and encourage restructuring? This thesis manipulated different incentives by varying the instruction explaining the association between performance and bonus. We expect to observe differences of group behaviour and performance in the process of solving a problem.

Previous research has shown that restructuring requires relatively high effort and enhances the detour effect when both exist (Adejumo et al., 2008). Thus the structure of detour provides a way to create individual and group goal conflict if we introduce individual incentive that draw individual towards an obvious temporary solution to an individual, but actually is a blind alley which works as a barrier to the solution for the group. The design of the task thus requires one to recognize the barrier as a barrier and move away from the direction of it, thus helping the group get the right solution through a detour. The fact that restructuring is required to see the detour that leads to the right solution added difficulty to this process.

Based on the above discussion, detour and restructuring were chosen to construct the experimental tasks in this study. Three conditions in terms of problem structure were chosen: control condition, the basic problem structure without any detour or restructuring; detours condition, where detours were added to the basic structure; and the detours and restructuring condition, where detours and restructuring were involved in the problem. The detour was designed so that if a group member obtains a temporary solution during the process of problem solving, it prevents the whole group from obtaining the right solution.

There are three conditions of incentives: individual incentive, mixed incentive, and group incentive. When an individual is given the individual incentive, it is expected that it will be hard for one to break up the obtained temporary solution. The individual incentive drives one toward the direction of detour, the most obvious solution to an individual, and once an individual successfully obtains the detour it serves as a "blind alley" for the whole group to reach the solution. In this case, individual

incentive encourages one to hold on the obvious solution and affects the group information processing. In this situation, the group would pay more attention to the remaining information than the obvious solution, while the obvious solution contains information the group needs to restructure the right solution. In return, the difficulty associated with the restructuring of the right solution makes this obvious temporary solution more appealing, thus should enhance the behaviour of "holding to an obvious temporary solution." When an individual is given the group incentive, we speculate that it should be relatively easier for one to come out of the blind alley because one is not losing anything by giving up the obvious temporary solution—the group incentive associate the bonus with the result of a group instead of an individual. Thus group incentive is expected to drive one away from the blind alley and towards the direction of the solution.

In summary, detour by definition attracts the group away from the direction of the goal, and restructuring enhances this tendency. Individual incentive should trap the group by encouraging an individual keep the obvious solution, and group incentive should encourage an individual to give up the obvious solution and emphasize on group success. This idea is well aligned with the goal mechanism, which says that goals drive attention toward the direction of goal-relevant activities and away from goal-irrelevant activities (Locke & Latham, 2002).

In terms of performance and behaviours, groups under group incentive are expected to show more group activities. As the incentive switches from group to individual, it becomes harder for a group to come out of the blind alley, and we expect longer time to solve the problem and less group level activities.

We also expect that as the complexity of problem structure increases, there would be weaker performance and more struggles in the blind alley as well as signs that indicate searching for new paths. Different incentives might not influence the basic problem structure where there is no detour and the answer is relatively straightforward; however, the effect of incentives is expected to be stronger in the structure of detours, as discussed earlier, where individual incentive makes the blind alley—the sub-optimal solution to the group but optimal to the individual—appealing to group members while group incentive would not do so. In the condition with both detours and restructuring, the effect of incentives is expected to be even stronger as the existence of restructuring enhances the attraction of detour, so it is expected to be even harder for groups under individual

incentives to give up detours. Based on previous discussion, a summary of the hypotheses is as follows.

Hypothesis 1 focuses on the main effect of the problem structure.

> Hypothesis 1: As the complexity of problem structure increases, the time to solve the problem and the complexity of search behaviour both increase.

We do not propose a hypothesis of main effect of incentives independent of structure; however, we do predict interaction effects of incentives and problem structure on group performance and complexity of search behaviour.

Hypothesis 2 is about interaction effect of the incentive and structure.

> Hypothesis 2a: Under the basic task structure, individual versus group incentives will affect neither the time to solve the problem nor the complexity of search behaviour.

This is because the solution to the problem under basic structure is straightforward and direct. Based on previous discussion, incentives should affect group performance and behaviour under the detour and restructuring task structure conditions.

> Hypothesis 2b: Under the task structures of detour and detour plus restructuring, as incentives switch from group to individual, the time to solve the problem increases.

However, we don't have any prediction of complexity of search behaviour in terms of incentives. Under the group incentive condition, if the group is stuck in the blind alley, it is expected that they would together search their way out. Under individual condition, if a group is stuck in a blind alley, one or more individuals might wish to keep the group stuck in the blind alley because of the benefits they will get individually. In other words, the sub-optimal solution to the group could be the optimal solution to an individual in the group, thus individuals are attracted to obtain the sub-optimal solution. The complexity of search behaviour of other group members at this moment is hard to predict. There could be increasing complexity of search behaviours as other group members struggle to find a way

out, while there could also be decreasing complexity of search behaviours if other group members give up and get stuck; therefore, there is no specific prediction of search behaviours.

> Hypothesis 2c: As incentives switch from group to individual, the effect of structural complexity on time to solve the problem and complexity of search behaviour will increase.

Chapter 3
Pre-Experiment

The interest of this thesis is to study the effect of individual versus group incentive conflict on group problem solving behaviour with the existence of detour and restructuring. A 3 by 3 design was used by varying the problem structure in three conditions: 1) basic problem structure; 2) detour; 3) detour and restructuring, and the ratio of individual versus group incentive in three levels: a) 100:0; b) 50:50; c) 0:100. However, before study could be run, the stimulus needed to be designed.

We adopted a card sort problem developed by previous researchers (Adejumo et al., 2008) which allowed the study of both incremental search and restructuring in the process of problem solving. Sixteen cards, each with two pictorial items, were used in each task. The cards were distributed randomly among a group of four participants, and their task was to exchange cards with each other so that at the end each person would obtain four cards of the same kind, using only one pictorial item in each card to represent the card.

While the previous study addressed some problem solving issues, a limitation of that experiment was that the choice of pictorial items used in experimental stimulus was relatively ad hoc. The relationship among pictorial items used for card categorization was not controlled, causing variability in results. As category homogeneity, the degree to which objects are similar to each other, influence group problem solving behaviour in this experiment (Abimbola, 2006), it is necessary to control category homogeneity among different stimuli. Thus this Pre-Experiment was developed for the purpose of controlling category homogeneity among different stimuli that would be used in the experiment.

The Pre-Experiment serves three primary purposes:

1. To understand similarity relationships among pictures that will be used in the experiment.
2. To control homogeneity on all experimental materials (pictorial items) across stimuli that will be used in the experiment.
3. To help design detours and restructuring items in the experiment.

3.1 Participants

Nine graduate students from the Department of Management Sciences participated in this study. This is the first and only time they were involved in this study. Participation was voluntary and there was no compensation of any kind. In this experiment, each participant worked alone to categorize 138 cards of pictorial items into 23 categories.

3.2 Materials

There were 138 cards in total. Each card had one colour picture of an object that was easy to identify on a white background. These cards belong to 23 categories with 6 cards in each category, as shown in Table 1. In each category there were five cards selected to be at the same level of abstraction while the sixth card belong to a higher level of abstraction. For example, the category of furniture consists of five cards of different chairs and one card of a coat rack. And the category of weather consisted of five cards of umbrellas and one card of a snowflake. All 138 cards were randomly spread on a table, facing up, with some of them overlapping with each other. Table 1 summarizes the twenty-three categories tested in the Pre-Experiment as well as the composition of each category.

Table 1: Twenty-three categories of cards and its composition used in Pre-Experiment

Category	Basic level items (5)	Restructuring item (1)
Alphanumeric	Numbers	Letter D
Animal	Birds	Lizard
Baking	Bread	Cake
Building	Bridges	Mayan Pyramid
Clothes	Tops	slippers
Competition	Athletic competitions	Two children playing chess
Costume	Hats	Mask
Drink	Coffee mugs	A glass of martini
Emergency	Police	Firefighter
Furniture	Chairs	Coat Rack
Instruments	Guitars	Piano
Medical	Medical Equipment	Nurse
Plant	Trees	Flower
Shape	Rectangles	14 sided polygon
Sports	Hockey equipment	Football
Stretching	People doing Yoga	A girl playing skipping rope
Technology	Computer accessories	Walkie Talkie
Transportation	Vehicles	Helicopter
Utensils	Forks and Knives	Bowl and chopstick
Vase	Vases	A pot of gold
Vegetable	Fruit	Corn
Weather	Umbrellas	Snowflake
Whisk	hand mixers	Electronic mixer

Restructuring items were used in this categorization problem to create two different levels of abstraction for each of the 23 categories (Rosch, 1978). Rosch suggested that categories have different levels of abstraction; for example, if one is asked to categorize three birds, one will refer to the categorization as "birds" because this is the most accurate and efficient way to categorize. However, if one is shown a lizard and asked if the lizard belongs to the category with the three birds, one might say no because a lizard is not a bird. On the other hand, if one is then given three pictures of cars and is asked to classify all items into two categories, one will likely classify the lizard with the birds instead of the cars, and name this category "animal" instead of "bird". In this example, the "bird" category is what Rosch called basic level category while the "animal" is a super-ordinate category at a higher level of abstraction. This shift from categorization at the basic level to

categorization at the super-ordinate level is an example of conceptual restructuring and is used to implement restructuring in problem solving in the experiment.

Figure 2 provides an example of cards used in the category of "Alphanumeric". The first five items are all single digit numbers; observation shows that one usually categorizes them as "numbers" very quickly with high confidence. However, the sixth item is a letter D, which is not a number. It is usually left on the table for a period of time before one could categorize it with the five cards of numbers. It can still be categorized with numbers because it is closer to "numbers" than to any other card available on the table, such as a guitar or an umbrella. Knowing that there must be six cards in a category, these restructuring items can usually be categorized correctly at the end, although the category is no longer "numbers", thus one has to reconstitute this category and name it "alphanumeric". This is a cognitive restructuring process, and thus we call these items "restructuring item".

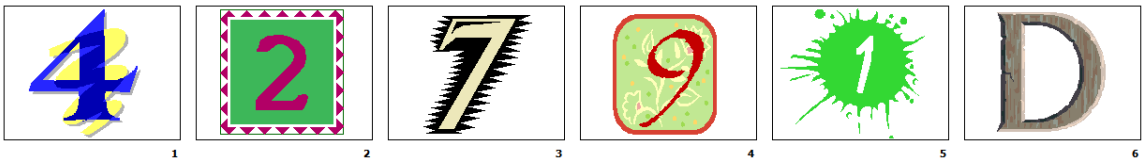


Figure 2: An example of a category with five basic items and one restructuring item

3.3 Procedure

Nine graduate students from the Department of Management Sciences at the University of Waterloo were asked to participate in a short study which would take between 30 minutes and 60 minutes. Each student was then scheduled for a time to come to the Uncertainty Lab at the University of Waterloo. The room was equipped with five digital cameras and four of them were used to record the experiment. Cameras used were from different angles, one from the top of the table, one facing the participant, and two from two corners of the room. Three suspended microphones connected to a multiplexer were used to record sound and voice. All participants were asked if they allow use of recording equipment prior to the start of the experiment and signed consent agreements. The participant was told that there were 138 cards randomly placed on the table, and their task was to sort all the cards into 23 categories with 6 cards in each category by placing cards from the same category into 23 cells marked by blue tape on the table. The participant was seated on the long side of the table and they were allowed to stand up and move along the long side. Every participant worked

alone in this experiment, and was asked to "think aloud" by verbally saying out loud what he/she was thinking during the process.

Figure 3 shows the setting of the Pre-Experiment by capturing a screen shot of one video of the experiments close to the end of the experiment. All cards were categorized into 23 marked cells on the surface of a table.

Figure 3: Setting of the Pre-Experiment

A training session was given before each participant was asked to perform the task. In the training session, the experimenter showed each participant six cards of pens and put them into one cell, and said: "I put these six cards together because I think they belong to the category of pens." The participant then started the actual task.

The experimenter used four cameras to record the whole process, focusing on the surface of the table while the participant's face expression was also visible. After each participant finished categorizing

all the cards, he was presented with a simple questionnaire asking him to rank the perceived similarity of each category of cards from 1 to 23, where 1 represents the most similar category and 23 represents the most dissimilar category

The results of Pre-Experiment contribute to the design of the experiment: the categories that were shown to be easy to identify were adopted as detour categories, and the hard-to-identify categories were adopted as restructuring categories, with the rest adopted as the categories served as the distracting items. Details of this process will be explained in the results section.

3.4 Measures

The purpose of this Pre-Experiment is to control the level of similarity among different categories. A few measures were used to indicate the similarity of each category:

1. Sequence of start: The order of starting a category. For example, if a participant first started collecting numbers, then the sequence of start for this category was 1.

All cards are randomly spread on the table, so each card has equal chance of being picked up and put into the category cell on the table by the participant if they are equally attractive. However, if certain categories are always recognized early or late, that means these categories vary in the degree of attractiveness. The earlier one category was started, the easier it indicates that category being recognized. Note that only the sequence that a category of cards started being collected was counted, whether that category was completed early or not was not indicated by this measure.

2. Number of steps to finish a category: This indicates the easiness for the restructuring item being recognized as part of a category.

Each time when a participant moved a card to the blue taped cell, it was considered as a step. Number of steps one made from the first card a participant put down in one category to the last card was counted. This measurement counted how long it took for each participant to complete one category. Because the restructuring item was usually the last card being put into a category, this measure indicates the easiness for the restructuring item being recognized as part of the category. The more steps one took to finish a category, the harder it was to categorize the restructuring item and restructure the category from basic to super-ordinate level.

3. Ratio of number of steps to finish a category to the total number of steps to finish all 23 categories for each participant: Steps taken to finish a category, divided by total steps one made from during the whole experiment.

There were two strategies used by participants and these two strategies cause difference on measure 2. One strategy was to finish each category completely before starting a new category, which result in small number of steps to finish a category for most categories. Another strategy was to work on many incomplete categories at the same time, and then allocate the remaining cards (usually restructuring items) to the most appropriate categories at the end of the experiment. Measure 3 took into consideration differences caused by these twp strategies, thus was believed to be more accurate indicating the easiness for the restructuring item being recognized as part of a categorization. This measure was highly correlated to measure 2.

3.5 Results

The results of the Pre-Experiment regarding to similarity of each category were shown in Table 2. The results of this experiment contribute to the design of the experiment. The easiest-to-identify categories were adopted as detour categories, the hard-to-identify categories were adopted as restructuring categories, and the rest were adopted as distraction categories. The result for each measure is explained as follow:

1. Column B of Table 2 shows the average of the sequence of start of each category among all 9 participants. Usually participants look at all the cards and start from cards and categories that are the most obvious to them. This measurement gives insight into how easily participants recognize a category without considering the restructuring item. As there were 23 categories, the sequence of start varies from 1 to 23.

2. Column C of Table 2 shows the average number of the steps needed to finish each category among all 9 participants. Each time when a participant moved a card to the blue area, it was considered as a step. Numbers of steps from the first card a participant placed in one category block to the sixth card placed in this category block were counted. This measurement counted how long it took participants to complete each category, which indicates the easiness of each category for each participant. Because the restructuring item was usually the last card to be

placed into a category, this measurement provides insight into how difficult it is to restructure the category at a super-ordinate level of abstraction by grouping the super-ordinate and basic level items together.

3. Column D shows the average of the ratio of steps to finish each category among all 9 participants. It's the steps taken to finish a category, divided by the total steps one made from the first movement of card to the last movement of card during the experiment.

4. Column E estimates the overall degree of categorization difficulty for each category. The product of the average start sequence (B) and the ratio of steps to finish a category (D) show a degree of relative difficulty for each category to be recognized and performed. The product increases as the category becomes harder to form, which indicates that cards in this category become less similar to each other.

Table 2: Results of the Pre-Experiment

Category (A)	Average start sequence (B) n=9	Average steps to finish a category (C) n=9	Average of ratio (D) n=9	Categorization difficulty (E=B*D)
Alphanumeric	9.667	26.333	0.3295	3.184809
Animal	12.11	26.222	0.2784	3.372078
Bread	10.78	22	0.265	2.855809
Building	12.56	26.222	0.2953	3.708141
Clothes	16.67	29.444	0.3194	5.322782
Competition	14.67	36.333	0.3605	5.286684
Costume	15.22	33.556	0.3706	5.64119
Drink	11	19.778	0.2323	2.555754
Emergency	6.222	34.222	0.3827	2.38121
Furniture	14.89	36.444	0.4245	6.320789
Instruments	8.111	16	0.1913	1.551905
Medical	12.44	22	0.2357	2.932689
Plant	15	17	0.1973	2.959715
Shape	13.89	7.6667	0.0832	1.156024
Sports	12.89	34.111	0.3642	4.694189
Stretching	13.78	32.444	0.4129	5.688386
Technology	11.22	24.444	0.3018	3.38698
Transportation	9.889	28	0.318	3.144657
Utensils	9.222	22.111	0.2214	2.04186
Vase	16	26.333	0.277	4.431678
Vegetable	7	31.222	0.3751	2.625407
Weather	11.22	37.333	0.3988	4.475175
Whisk	11.56	13	0.1428	1.650393

According to the categorization difficulty, categories with low categorization difficulty should be selected as detours and categories with high categorization difficulty should be selected as restructuring categories.

Categorization mistakes made by participants were also recorded. For example, if a picture of bread was put into the category of utensil, but later corrected, this action indicates that there was some overlap (i.e. perceptual similarities) among the items in these two categories. The number of overlapping cards between categories is recorded in Appendix E. This result was used to avoid including overlapping categories in the same stimulus in the experiment.

Table 3 shows the design of three stimuli in the experiment based on above discussion. Categories with low categorization difficulty were selected as detours and categories with high categorization difficulty were selected as restructuring categories.

Table 3: Three stimuli used in the experiment and the relative difficulty

		Category	Categorization Difficulty
Stimulus 1	Detour category 1	Instruments	1.552
	Detour category 2	Utensils	2.042
	Restructuring category 1	Alphanumeric	3.185
	Restructuring category 2	Weather	4.475
	Restructuring category 3	Costume	5.641
	Restructuring category 4	Competition	5.287
Stimulus 2	Detour category 1	Shape	1.156
	Detour category 2	Whisk	1.650
	Restructuring category 1	Animal	3.372
	Restructuring category 2	Building	3.708
	Restructuring category 3	Clothes	5.323
	Restructuring category 4	Stretching	5.688
Stimulus 3	Detour category 1	Drink	2.556
	Detour category 2	Plant	2.960
	Restructuring category 1	Technology	3.387
	Restructuring category 2	Vase	4.432
	Restructuring category 3	Sports	4.694
	Restructuring category 4	Furniture	6.321

Chapter 4
Experiment

This experiment investigates the effect of individual and group incentives on group performance and behaviour when solving three types of problems. A 3(incentives) by 3 (structures) factorial design was used where the problem structure varied in three conditions: 1) basic problem structure; 2) detour; 3) detour and restructuring, and the ratio of individual versus group incentive in three levels: a) 100:0; b) 50:50; c) 0:100.

By varying different ratios of individual and group incentive, situations are simulated where the individual's goal could conflict with the group's goal. This conflict between individual and group goal was created by giving individual incentives that encourage them to pursue individual success, the achievement of which might prohibit the group from taking the path leading to group success.

As discussed earlier, a card sort problem from a previous study (Adejumo et al., 2008) was adopted and revised here. And results from the Pre-Experiment were used in the design of stimulus.

4.1 Participants

One hundred and sixty-four students from MSCI 311, a third year undergraduate Management Sciences class at the University of Waterloo, signed up for this experiment to receive extra credit for the course. Each group consisted of four participants who either signed up together for the same time slot or were randomly assigned to a time slot based on availability. Participants who were randomly assigned were informed that they would work with three other students; however, they were not told any information about their group members until they saw each other on the experiment day. All experiments took place in the Uncertainty Lab in Management Sciences at the University of Waterloo.

4.2 Stimulus set

Table 4 shows the 9 experimental conditions in this study. The study is a three by three design with three conditions for incentives and three conditions for task structure. Every group was given a fixed ratio of incentive and none of the students were aware that the ratio of incentive was a factor being manipulated. Every group participated in three different tasks with order of Task 1 (pure card sort),

Task 2 (two detours), and Task 3 (two detours plus restructuring), so that tasks increased in structure complexity through the experiment.

Table 4: The composition of nine stimuli in this study

		Incentive condition		
		a) Individual (100:0)	b) Mixed (50:50)	c) Group (0:100)
Task Structure	1) pure card sort	1a	1b	1c
	2) detour	2a	2b	2c
	3) detour + restructuring	3a	3b	3c

Participants were not aware of the difference of structure among three tasks and every group participated in three tasks in the order of 1), 2), 3). The 16 cards of each task were designed based on the results from Pre-Experiment. Analysis from Pre-Experiment provided similarities among 138 cards from 23 categories. Three different versions of stimulus were used across three tasks to each group to make participants feel that they were facing a new problem each time. The following sections explain the structures of three experimental tasks in Experiment II and how they were designed.

4.2.1 Problem structure 1: pure card sort

Figure 4 describes the coding used to represent the design of sixteen cards used in task 1. Each card used in the experimental task had two pictures on it. For the purpose of explaining the structure and logic behind the pictorial item, letters and numbers are used in this section to represent each pictorial item. The number at the centre of each card represents the number of each card. Besides this number at the centre of the card, there are two codes on each card: one on top, representing a pictorial item, and the other on at the bottom of each card, representing another pictorial item. Thus all pictorial items are represented in this figure. Take card No. 1 for example: there are two pictorial items in the card, A1, and M1. "M1" means this pictorial item is the "first" item in category "M", and "A1" means this pictorial item is the "first" item in category "A". As we see, there are only two cards in category "M" across all 16 cards, while there are four cards in category "A", which means category "A" could compose a right solution (four cards of the same kind of "A") and "M" cannot as there are only two cards in that category. Categories like "M" are called distracting items. M1 and M2 simply mean that these two pictorial items are very similar and they are two different pictorial items that belong to the same category of "M".

M1 1 A1	N1 2 A2	G1 3 A3	H1 4 A4
N2 5 B1	I1 6 B2	J1 7 B3	K1 8 B4
L1 9 C1	J2 10 C2	M2 11 C3	G2 12 C4
H2 13 D1	K2 14 D2	L2 15 D3	I2 16 D4

Figure 4: The design of sixteen cards used in task 1

As discussed earlier, we aimed to control similarity across all categories using results from the Pre-Experiment. If we use S(A,B,C,…) to represent the similarity among A, B, C, etc, then the design of problem structure 1 is based on the following criteria:

S(G1,G2) = S(H1,H2) = S(I1,I2) = S(J1,J2) = S(K1,K2) = S(L1,L2) = S(M1,M2) = S(N1,N2) = S1
S(A1,A2,A3,A4) = S(B1,B2,B3,B4) = S(C1,C2,C3,C4) = S(D1,D2,D3,D4) = S2
S(Xi,Yj) = S3 (where X, Y = A, B…, N; X≠Y; i, j = 1, 2, 3…)
S1 > S2 > S3

In words, the similarities between cards that belong to one of the distracting categories were designed to be at the same level (S1), and the similarities among cards that belong to one of the four categories that serve as the solutions were designed to be at the same level (S2), the similarities between any two cards that do not belong to the same category were design to be at level 3 (S3). Level 1 is higher than level 2 and higher than level 3.

This design ensures that the correct solution corresponds to the four rows of cards, representing category A, B, C, and D. Distracting items always come in pairs and they serve the purpose of distracting participants from concentrating on one category. So here G1, G2, H1, H2, I1, I2, K1, K2,

L1, L2, N1, N2, M1 and M2 are distracting items. So four sets of cards the group will obtain at the end should be: (1, 2, 3, 4); (5, 6, 7, 8); (9, 10, 11, 12); (13, 14, 15, 16).

4.2.2 Problem structure 2: detour

Structure 2 maintained the basic structure of task 1, but added "detour" as a new factor; specifically pictorial items that formed two "detour categories" replaced some of the distracting from structure 1. Figure 5 shows the structure of task 2. Items M1, M2, I1, I2, H1, H2, J1 and J2 were replaced with eight pictures belonging to two detour categories (E and F). To show them clearly, they are shaded in Figure 5.

E1 1 A1	N1 2 A2	G1 3 A3	F1 4 A4
N2 5 B1	E2 6 B2	F2 7 B3	K1 8 B4
L1 9 C1	F3 10 C2	E3 11 C3	G2 12 C4
F4 13 D1	K2 14 D2	L2 15 D3	E4 16 D4

Figure 5: The design of sixteen cards used in task 2

As before, if we use S(A,B,C,…) to represent similarity among A, B, C, etc, then the design of problem structure 2 is based on the following criteria:

S(G1,G2) = S(K1,K2) = S(L1,L2) =S(N1,N2) = S1

S(A1,A2,A3,A4) = S(B1,B2,B3,B4) = S(C1,C2,C3,C4) = S(D1,D2,D3,D4) = S2

S(Xi,Yj) = S3 (where X, Y = A, B…, N; X≠Y; i, j = 1, 2, 3…)

S(E1,E2,E3,E4) = S(F1,F2,F3,F4) = S4

S4 ≥ S1 > S2 > S3

The only way for everyone in a group to get a set of four cards of the same kind are the same as in Task 1: (1, 2, 3, 4); (5, 6, 7, 8); (9, 10, 11, 12); (13, 14, 15, 16).

However, the design gave detour categories (E and F) the highest level of similarity, to encourage group members to collect detour set E or F at some point in the problem solving process. Once collected, category E and F would become a blind alley that prevents the group from solving the problem completely. As discussed earlier, detours create difficulty because the problem solver follows a path that seems to take him closer to the solution while objectively this path is actually a blind alley. The only way to reach the solution is to follow a detour path that subjectively seems to move away from the solution, making this path less attractive to the problem solver.

4.2.3 Problem structure 3: detour and restructuring

Figure 6 shows the structure of task 3. Restructuring is a new factor in the structure of task 3. Cognitive restructuring was implemented using one item that looked very different from all other items in that category, requiring participants to "think broader" and form categorization at a super-ordinate level of abstraction, which was harder than thinking at the more natural basic level of abstraction used in task 1 and 2 (Rosch, 1978). A4, which was at the same level of abstraction as other items in "A" cateogry, was replaced with A'4, which was a pictorial item at a different level of abstraction within the category of A. A'4 is less similar to A1, A2 and A3 than A4, but still more similar to them than to the pictures in other categories in this task. With the same logic, B3 was replaced with B'3, C3 was replace with C'3 and D4 was replaced with D'4. Examples of pictorial items from this task condition are shown in the following section.

As before, if we use S(A,B,C,…) to represent similarity among A, B, C, etc, then the design of problem structure 3 is based on the following criteria:

$S(N1,N2) = S(G1,G2) = S(K1,K2) = S(L1,L2) = S1$
$S(A1,A2,A3) = S(B1,B2,B4) = S(C1,C2,C4) = S(D1,D2,D3) = S2$
$S(Xi,Yj) = S3$ (where X, Y = A, B…, N; $X \neq Y$; i, j = 1, 2, 3…)
$S(E1,E2,E3,E4) = S(F1,F2,F3,F4) = S4$
$S(Ai,A'4) = S(Bj,B'3) = S(Cj,C'3) = S(Di,D'4) = S5$ (where i= 1, 2, 3; j= 1, 2, 4)
$S4 \geq S1 > S2 > S5 > S3$

E1 1 A1	N1 2 A2	G1 3 A3	F1 4 A'4
N2 5 B1	E2 6 B2	F2 7 B'3	K1 8 B4
L1 9 C1	F3 10 C2	E3 11 C'3	G2 12 C4
F4 13 D1	K2 14 D2	L2 15 D3	E4 16 D'4

Figure 6: The design of sixteen cards used in task 3

The correct and only solution consists of: (1, 2, 3, 4), (5, 6, 7, 8), (9, 10, 11, 12), and (13, 14, 15, 16). The existence of A', B', C' and D' makes it harder to identify categories A, B, C, and D as the similarity between the restructuring item and other item in the same category is not lower than the similarity among items in the detour categories (E and F). Note that in this task design restructuring items were placed in the same cards as detours items. This might increased the relative attractiveness of detours compared to correct solution categories compared to task 2 more.

Three different versions of the stimulus were constructed for each of the task to decrease potential learning effects. The three versions were structurally identical; however, there were no duplicates of categories or pictorial items among three versions so that each group received a unique set of pictures and categories on each task. Each version of the stimulus was used equal number of times in all three tasks across all the groups to ensure randomness.

Figure 7 shows one of the three stimulus versions that were used for task 3. All other stimulus sets used for the various experiments can be found in Appendix B.

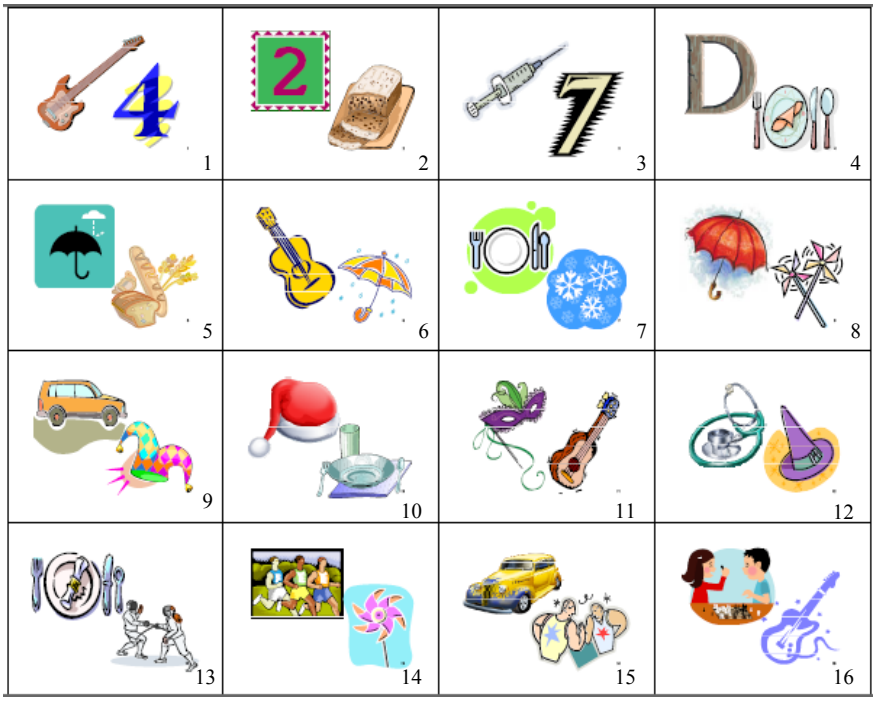


Figure 7: One of the three stimulus versions that were used for task 3

(Note: The numbers at right bottom corner on cards 1-16 in Figure 7 did not appear on the actual cards used in the experiment.)

The correct categories in the solution are categorizes of alphanumerical (cards 1-4), weather (cards 5-8), costume (9-12), and competition (13-16). The second item on each card was a distracting item that was not required to form categories.

Four dinner set pictures from card 4, 7, 10, and 13 formed a detour set; while four guitars from card 1, 6, 11, 16 formed another detour set. Once participants collected these two detour sets, they thought they were closer to the solution as two of them had "half of the solution". To separate cards from either of these two categories would make it like they were moving away from the solution. These two detours served as the blind alley and prevented participants from solving the whole problem effectively. The blind alley looked more like the path to the right solution when restructuring items exist. In this case, letter "D", snowflake, mask, and chess were four restructuring items. As it is cognitively simpler for participants to group 4, 2 and 7 as numbers than alphanumerical, they tend to view 4, 2, and 7 as a category of numbers and the D an odd one that does not fit into this category.

In the condition of task 2, there was no restructuring item. When the categories shown in Figure 7 were used for the task 2 stimulus, "D" was replaced with a number, snowflake was replaced with an umbrella, mask was replaced with a hat, and chess was replaced with an athletic competition. In the stimulus for task 1, the two detour categories were replaced with four pairs of random distracting items, leaving only four possible sets of four cards of a kind rather than six.

4.3 Procedure

When participants arrived in the lab, they were introduced to each other and given name tags, and then asked to sit at a round table of 1.5m diameter. Participants could see the whole surface of the table during the first training; however, there were two 30cm high barriers on the surface of the table during the second training and the actual tasks. This set up prevented the students from seeing others' table surface during the experiment so that each individual would not possess complete information of the whole problem. Figure 8 shows an overhead view of this experimental setup. This set up was consistent through all 41 groups in this experiment.

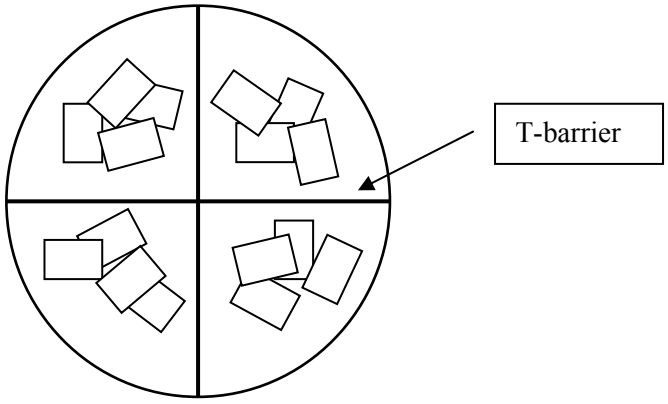


Figure 8: Overhead view of the experimental setup

Four video cameras (one directly above the table and three from other angles) and microphones in the lab were used to record all the experiments. Participants were made aware of the use of recording equipment and signed consent forms prior to the experiment.

Each participant was given four cards at the beginning of each task. Each card has two pictorial items on it. They can communicate with each other in the follow two ways:

1. Verbal communication: they can talk to each other at any point of time about anything during the experiment
2. Exchange cards with each other

The goal for them was to exchange cards so that at the end each person would obtain four cards of the same kind, using only one item on each card as the basis for categorization. Each group participated in three differently structured tasks, and students were told that for each task they could earn up to 1 bonus mark. Three ratios (0:100, 50:50, 100:0) of individual and group incentive were manipulated among 41 groups:

1. Pure individual incentive: Groups were given the instruction that each person would get the full bonus mark if at the end of a task they got four cards of the same kind, regardless of how well other group members performed.
2. Mixed incentive: Groups under this condition were told that each person would get half of the bonus mark per task if they got four cards of a kind and the other half of the bonus mark if every other group member got a set of four cards of a kind.
3. Group incentive: Groups under this condition were told that each person would get the bonus mark if everyone in the group got a set of four cards of a kind.

Although every group was treated consistently on how they would earn the bonus mark throughout the experiment, after the experiment, each participant eventually received the full bonus mark regardless of performance.

Each group completed two training tasks prior to exposure to the three experimental tasks. To simplify the problem, four identical items for each correct category were used in the first training, without the T-barriers on the table so that participants could see and understand the complete problem. In the second training, categories of identical items were again used, but this time barriers were used. To avoid potential learning, categories and pictorial items used in training were different from cards used in any of the stimuli used in the three experimental tasks.

The initial distribution of the cards for both task 1 and task 2 were always: Person A got (1, 5, 9, 13), Person B got (3, 8, 11, 16), Person C got (2, 6, 10, 14), and Person D got (4, 7, 12, 15). This

distribution ensured that each participant began with one item from each correct category. With this distribution, Person B and Person D in task 2 also received two cards containing the detour pictures to start, which might increase the attractiveness of detours. However, in task 3, the distribution was always: Person A got (3, 8, 11, 16), Person B got (1, 5, 9, 13), Person C got (4, 7, 12, 15), and Person D got (2, 6, 10, 14). Note the initial distribution of Person A and 2 were swapped as well that for Person C and Person D in task 3 compared to that in task 1 and task 2. This is to avoid learning effect as every group conducted both task 2 and then task 3. We didn't want any individual to apply what they learnt from a previous task (if any). With this distribution in task 3, Person A and Person C always received two cards containing the detour pictures to start, while in task 2 it was Person B and Person D. This distribution was held constant across all groups.

Participants were made aware of a limit of 15 minutes on each task, which was about 1.5 times longer than the longest time taken by groups in an earlier pilot study. Participants were told that when they solved the problem they should verbally tell the experimenter that they were done, otherwise, the task would end after 15 minutes. There were two large clocks on opposite walls in the lab such that every participant was able to see the time clearly without big physical movements. Two bright yellow stickers were used to mark the 15 minutes time range on each clock during each task, and the experimenter reminded everyone of the time after about 10 minutes on each task if they hadn't submitted the answer yet.

After each task, participants were asked to fill a short questionnaire regarding to the information of the task they just finished. At the end of the last task, they were asked to fill a short questionnaire regarding to the whole experiment experience. Questionnaires are in Appendix D.

4.4 Measures

4.4.1 Time

Time spent on each task was recorded during the experiment as a measurement of performance. Note that the groups were required to call the end verbally when they were done, so the time spent on a task means the time from beginning of a task to end of a task (when the groups told the experimenter that they were done). This required participants to agree on their solution and to be confident enough in their solution to call the end spontaneously.

A time limit of 15 minutes on each task was introduced prior to the experiment. The time limit might weaken the strength of using time as a measurement; however, most groups solved task 1 and task 2 in about 5 minutes, and task 3 in about 10 minutes. A time limit might also potentially encourage individuals take incentives more seriously and amplify their relative effects. For example, if individual incentive was introduced, one would get a bonus mark if one obtains four cards of a kind at the end of the task. In this case, the time limit might encourage one to keep a detour category because time could run out very quickly and if he agrees to break up the category, he might not get a full set of cards before the time limit. Without a time limit, one might be more willing to give up a detour category because there might be a higher chance of obtaining another complete set at some point. There were a few groups who couldn't solve task 3 after 15 minutes and the experimenter waited until they solved the problem. Only one group could not solve the problem even with extra time and the experimenter told them to stop after 25 minutes.

4.4.2 Number of card exchanges

The initial card distribution ensured that each participant began with one item from each correct category. The only way to reach the solution was to exchange cards. Each time a card was exchanged from one person to another, the number of card exchange increases by one. At any point of time, the number of card exchanges shows the number of cards that have been exchanged among the group members. Previous research used number of card exchanges as a performance measure and showed that tasks with greater difficulty usually were associated with more card exchanges (Abimbola, 2006).

With the initial card distribution used in all tasks in this study, the minimum number of card exchanges required to solve any task is 12. We expected to see difference on this measurement between different incentives as well as different structures.

Figure 9 shows a typical card exchange sequence, in which the group reached the solution after twelve card exchanges. Each card was represented using the coding scheme introduced earlier. Because the combination of letters for each card was unique, letters were enough to represent a card in this situation. For example, when we introduced the stimulus sets, we used A1E1 to represent card 1, while here we simplified A1E1 to AE.

	Task 2																			
	Task ID: Experiment 28_Jan 23_Individual_213_2:30 Design Set:3																			
Time	Person A					Person B					Person C					Person D				
0:23:42	AE	BH	CN	DF		AI	BM	CE	DE		AH	BE	CF	DM		AF	BF	CI	DN	
0:28:24	DE	BH	CN	DF		AI	BM	CE	AE		AH	BE	CF	DM		AF	BF	CI	DN	
0:28:25	DE	BH	CN	DF		AI	BM	CE	AE		AH	CI	CF	DM		AF	BF	BE	DN	
0:28:35	DE	BH	CN	DF		AI	AF	CE	AE		AH	CI	CF	DM		BM	BF	BE	DN	
0:28:38	DE	BH	DM	DF		AI	AF	CE	AE		AH	CI	CF	CN		BM	BF	BE	DN	
0:28:55	DE	DN	DM	DF		AI	AF	CE	AE		AH	CI	CF	CN		BM	BF	BE	BH	
0:28:59	DE	DN	DM	DF		AI	AF	AH	AE		CE	CI	CF	CN		BM	BF	BE	BH	
0:29:00	DE	DN	DM	DF		AI	AF	AH	AE		CE	CI	CF	CN		BM	BF	BE	BH	

Figure 9: Sample Card Exchange

The first two rows were information about the task, and row 3 was the heading. After row 3, each row represents the cards distribution among four participants at a point of time. For example, at time 0:23:42, the task was started and participant A had four cards: AE, BH, CN and DF. At time 0:23:24, Participant A and B exchanged card AE and DE, thus A now held the four cards: DE, BH, CN and DF. Because two cards were exchanged at this moment, the number of card exchanges was 2. The cards that were exchanged at each point of time are shaded in Figure 9. Every card exchange during each task was coded in this way and the total number of card exchanges was the number of cards that were shaded. In this example 12 cards were exchanged in total.

4.4.3 Number of groups that collected detour categories in structure 2 and structure 3

Detour categories exist in both structures 2 and 3. In some situation a group could reach the right solution without collecting any detour categories, while other times, a group collected detour categories and got was stuck in blind alleys. The number of groups that collected detour categories in each experimental condition was counted. For example, if 3 groups in total in condition (detour and individual incentive) ever collected detour category, then the number of groups that collected detour categories for this condition is 3.

4.4.4 Number of detour categories a group collected in structure 2 and structure 3

Two detour categories exist in both structures 2 and 3. If a group never collected a complete detour category set, then this number is 0. For each group who ever collected detour categories, the number

of detour category collected was counted, as well as times each detour category was collected. For example, if one group only collected one of the two detour categories and for only once, then the number of detour categories this group collected is 1. If one group collected both detour categories, and for once each, then this number is 2. If one group collected one detour category, but collected it twice (the same detour category was collected again after broke up), then this number is 2. This number could be higher than 2 for groups who collected more than one detour categories for multiple times.

4.4.5 Time spent holding detour categories in structure 2 and structure 3

If any group member ever successfully obtained four cards from a detour category, the length of time from when four cards from the same detour category were obtained by one person to when they were separated was recorded. In cases when the group collected the same detour categories more than one occasion, the sum of the time was used as the time this group spent holding detour categories.

4.4.6 LIB (line index of balance) related measurements

A method based on balance theory was used in previous research (Adejumo et al., 2008) and adopted in this research. Balance theory models a group's cognitive representation of a situation as a graph with relationships among cognitions, and balance is achieved through the incremental restructuring of this picture. Balance theory is based on Heider's theory, which views the relations of cognitions sentimental; for example, like or dislike (Heider, 1946). Heider's primary unit of analysis consisted of three parts, either three people or two people and one object. A balanced relationship occurs when the product of the sign relations among the three parts is positive. If the product is negative, imbalance would result in psychological tension that eventually would change the relationship among cognitions and result in a balance state. For example, if Person A likes Person B (+), and knows that Person B likes Object C (+), but Person A dislike Object C (-), then the A-B-C network is not balanced and changes needs to be made. Person A could make this change by either disliking B (-), starting to like C (+), or trying to persuade B to dislike C (-) to reach the balance of A-B-C.

A few modifications have been made on Heider's theory. Newcomb (1953) generalized Heider's theory to symmetry theory as well as considered the strength of each relationship. Cartwright and Harary (1956) extended Heider's theory to networks with more than three entities, and kept the binary relationship (either like or dislike, positive or negative). Their result suggested that a network was

balanced when all the entities in this network could be draw into two subgroup that were mutually hostile to each other, saying any two points that were positively related should always be in one subgroup, and any two entities that were negatively related should always be in two different subgroups. They also proposed measures, line index of balance (LIB), for the degree of balance for such a 2-balanced signed graph. Doreian and Mrvar (1996) generalized the LIB measure for K-balanced signed graph for cases with more than two subgroups. This change made the LIB measure more general in social network analysis where more than two subgroups could be formed. Thus LIB measures the degree of imbalance, tracks changes of balance during the problem solving process, and reflects the cognitive dissonance of each state during this process. The change of LIB from time to time indicated incremental search and restructuring involved in the problem solving process.

The problem structure is represented as a network, with nodes representing each pictorial item and links between nodes representing similarity relationships between items. We specify the link between any two similar items as a positive one, and the relationship between any two dissimilar items as a negative one. In other words, if two pictorial items belong to the same category, they are attracted to each other, and vice versa. Figure 10 shows an example of the LIB between two cards (card 1 and card 2). In Figure 10, card 1 and card 2 are to help readers distinguish them as two cards. There are two pictorial items on card 1: E1 and A1. There are two pictorial items on card 2: N1 and A2.

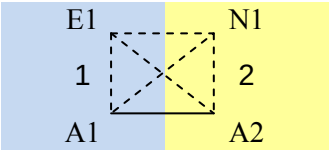


Figure 10: An example of LIB between two cards (card 1 and card 2)

We use a solid line to show a positive relationship (attracted to each other) between two pictorial items and a dash line to show a negative relationship (not attracted to each other). A1 and A2 belong to category "A", thus the link between them is assigned a positive sign (solid line). E1 and N1 are dissimilar and do not belong to the same category, thus the link between them is assigned a negative sign (dash line). The relationship between E1 and A1 is by default negative, because two pictorial items on the same card are always dissimilar.

In previous research (Adejumo et al., 2008), only one pictorial item of each card was counted as valid at one time, and the other item was ignored, thus only part of the LIB of the whole system (16 cards) was calculated. However, a problem with that approach was that it was hard to decide which picture should be counted or not. It was also hard to represent a situation where participants were aware of possibilities of categories for both pictorial items in one card. Thus in this study, in order to represent the complete system, all 32 items on the sixteen cards were taken into account when LIB was used.

LIB measures the degree of imbalance. The imbalance for the whole system consists of two parts. The first part is the number of imbalance among all cards held by each person (internal). There are usually four cards at a time for each person, and within four cards from one person, different pictorial items on different cards make one tend to separate these cards. So if we count the number of dash line within these four cards, it will be the total number of imbalance (LIB1) within one person. The greater this number is, the greater is the tendency to break them up. The second part is the number of imbalance between the four participants. For example, if Person A has a pictorial item A, and Person B has another pictorial item A, then these two As are similar to each other and create the tendency to be obtained by one person. This tendency leads to an exchange of cards between Person A and 2. Thus if we count the number of solid lines between four participants at every stage, this number become the total number of imbalance (LIB2) between four persons. The greater this number is the greater is the tendency to exchange cards. The sum of LIB1 and LIB2 is the total imbalance for the whole system.

For each group, the degree of imbalance was calculated after every card exchange step during the problem solving process for all three tasks. Based on the initial distribution of cards at the beginning of each task, each group started with a maximum imbalance number and correct solution corresponds to a minimum imbalance number. The LIB of a group after each card exchange was calculated and plotted on a figure where LIB served as the y axis and time served as the x axis. Thus a trajectory of LIB over time was obtained. This trajectory provides an objective view of the process of solving each task. It provides insight into the path taken by each group to reach the solution, including a potential means of identifying restructuring and incremental search phases of the problem solving process.

Based on the above discussion about LIB and its trajectory, a few related measures were developed as follows.

1. Number of LIB trajectory direction reversals:

A direction reversal occurs when the direction of imbalance changes from an increasing to a decreasing trend or from a decreasing to an increasing trend. The number of direction changes in a group's LIB imbalance trajectories was counted. When imbalance decreases, it indicates that the group is moving closer to the solution, and when imbalance increases, it indicates that the group is moving farther from the solution. Thus, there is a threshold in the detour condition, where decrease of imbalance at a certain point could mean that the group was trapped in a blind alley and in order to get out of the blind alley, they has to first increase the imbalance as they need to break up some categories they had already formed.

2. R squared value for polynomial regression trend lines:

A series of regression analyses were performed on the time series of LIB data. First order (linear) to fifth order polynomial regression analyses were used in the regression analyses, and the R squared value was recorded as a measurement. LIB imbalance number was modeled as a polynomial function of time and R squared was used as an indicator of how well the regression model fit the LIB data as the polynomial order varies. Previous research (Adejumo et al., 2008) showed that when the data can be fit into low order polynomial regression, it usually implies that the path the group took more or less led them to the solution directly. As higher order polynomial regression models are required to fit the data, it corresponds to more reversals in directions of the LIB trajectories, implying that groups follow more complex search paths and have greater difficulty reaching the correct solution.

Chapter 5
Results

This experiment investigates the effect of individual and group incentives on group performances and behaviours when solving three types of problems. Experiments manipulated the relative weights of individual and group rewards using three ratios (0:100, 50:50, 100:0). Three different experimental tasks were used to investigate such incentive conflict effect on different kind of problems. In particular, problems had detours and restructuring; they are: 1) basic card sort; 2) detour; 3) detour and restructuring.

5.1 Descriptive measures

5.1.1 Time

Time measures efficiency of solving each problem. Table 5 shows the results of average time and standard deviation across the 9 stimuli and Figure 11 show the graphic representation of the same results.

Table 5: Average time and standard deviation across nine experimental conditions

Task Structure	Incentives	Sample Size	Minimum (s)	Maximum (s)	Mean (s)	Std. Deviation (s)
1-basic card sort	a-individual	14	150	398	247.6	84.9
	b-mixed	13	111	535	284.9	118.8
	c-group	14	102	368	234.7	86.3
2- detour	a-individual	14	135	447	271.9	104.8
	b-mixed	13	104	964	393.4	276.0
	c-group	14	122	732	293.0	185.1
3-detour and restructuring	a-individual	14	312	1415	693.6	344.5
	b-mixed	13	514	1484	790.2	237.7
	c-group	14	298	1553	804.4	370.8

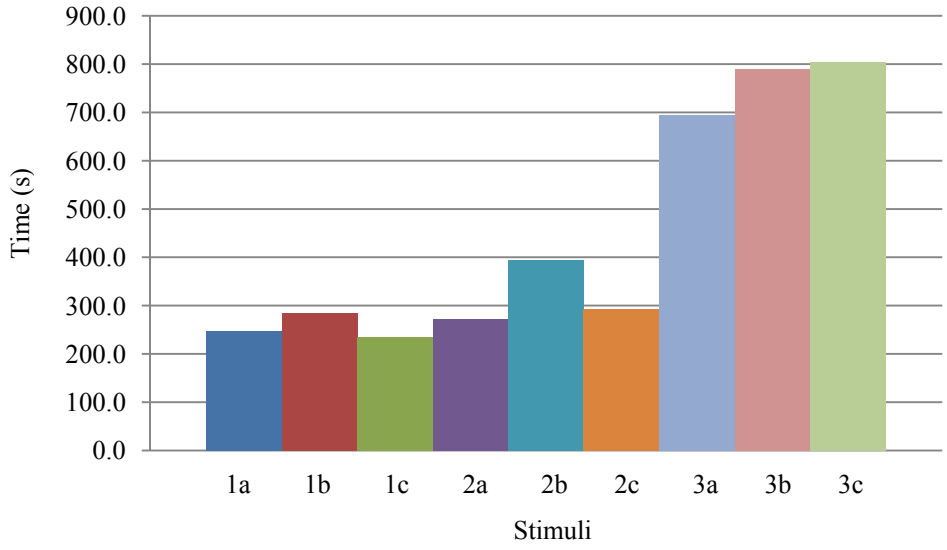


Figure 11: Average time across the nine experimental conditions

Different from expected, there was no big differences between the average time to solve task 1 and task 2. The time groups spent on task 1 is more than expected for a few reasons. First, each group was given two training sessions before the first task, where the correct answers were identical pictorial items. It might cause an impression that "four of a kind" means "four identical" cards among some participants; thus, in the first task, a significant amount of time was spent to figure out that "four of a kind" can also mean "four cards that are not identical but belong to the same kind." Second, groups tended to discuss and form a strategy in the first task, which they usually use across three tasks; this also cost time in task 1. Some groups were "lucky" in task 2 and didn't notice the existence of detour, thus they solved the problem without any struggle. A few groups solved the second task even quicker than the first task as they didn't spent time in task 2 to figure out what can "four of a kind" mean or explore the strategy for the group, which they had done in task 1.

5.1.2 Card Exchanges

Number of card exchanges provided an indication of the amount of incremental search used in information processing needed for each group to solve each task. Table 6 shows the results of average number of card exchanges and standard deviation across the 9 experimental conditions and Figure 12 show the graphic representation of the same results.

Table 6: Average number of card exchanges and standard deviation across nine experimental conditions

Task Structure	Incentives	N	Minimum	Maximum	Mean	Std. Deviation
1-basic card sort	a-individual	14	12	22	13.1	2.7
	b-mixed	13	12	22	14.5	3.8
	c-group	14	12	20	14.1	3.4
2- detour	a-individual	14	12	24	14.0	3.8
	b-mixed	13	12	34	16.2	6.1
	c-group	14	12	28	14.8	5.3
3-detour and restructuring	a-individual	14	12	40	17.9	7.9
	b-mixed	13	12	60	31.2	16.4
	c-group	14	19	70	36.1	17.7

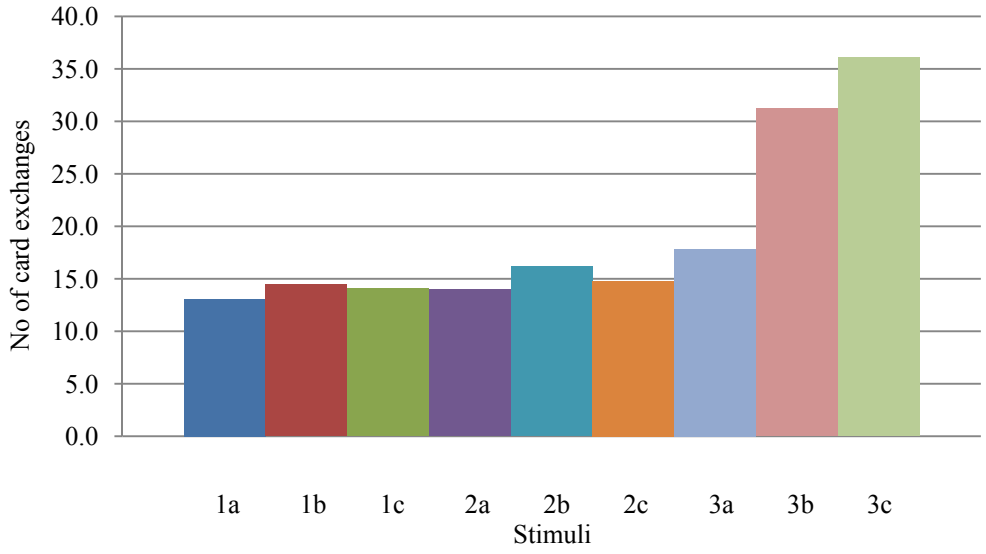


Figure 12: Average number of card exchanges across nine experimental conditions

5.1.3 Number of groups that collected detour categories in task 2 and task 3

The number of groups that collected detour categories in each incentive condition in structure 2 and 3 provides a view of the chance that a group collected the detour category under each condition. Table 7 shows the result of number of groups that collected detour categories as well as the percentage. As we see, less than one third of groups collected detour categories under structure 2 regardless of the

incentive condition. However, in structure 3, 43% of groups under incentive a (individual incentive) didn't even collect detour categories, while this percentage increase to 92% under incentive b (mixed incentive) and 100% under incentive c (group incentive). Figure 13 shows the same result in graphic presentation. Note that there is no data for the basic structure problems as there is no detour in basic structure problems.

Table 7: Number of groups that collected the detour categories

Stimulus	No. of groups that collected detour categories	Total groups	Percentage
2a	3	14	21%
2b	3	13	23%
2c	2	14	14%
3a	6	14	43%
3b	12	13	92%
3c	14	14	100%

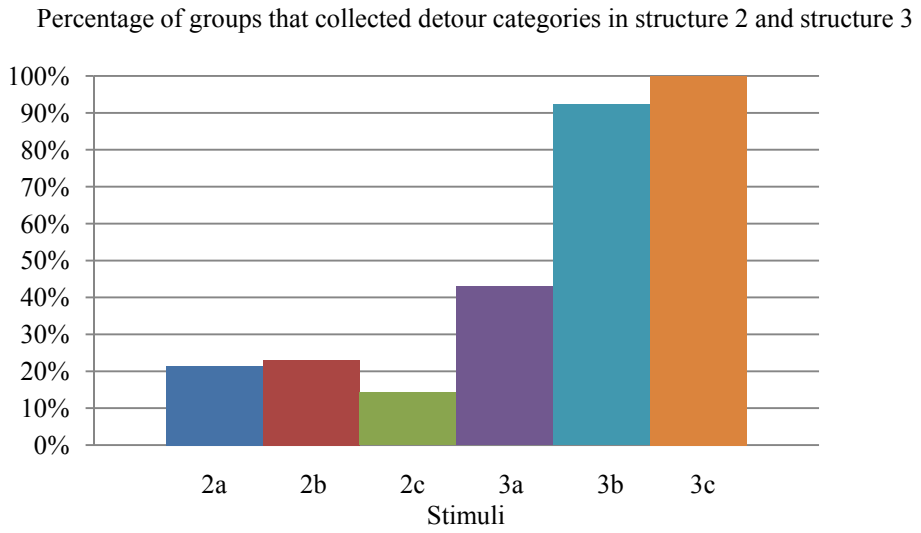


Figure 13: Percentage of groups that collected detour categories in Structure 2 and 3

5.1.4 Time spend holding detour categories

When four cards of a detour category were collected, the length of time it was held was recorded. The time spent holding detour categories for a group is the sum of the time when any detour category was held during the process of solving one problem.

Table 8 shows the result of time spent holding detour category. Figure 13 shows the same result in graphic presentation. Note that there is no data for the basic structure problems as there is no detour in basic structure problems.

Table 8: Time spent holding detour categories

Task Structure	Incentives	N	Minimum	Maximum	Mean	Std. Deviation
2- detour	a-individual	3	181	252	219	35.8
	b-mixed	3	84	647	343	284.2
	c-group	2	144	393	268.5	176.1
3-detour and restructuring	a-individual	6	93	700	269	221.2
	b-mixed	12	91	797	288.8	202.1
	c-group	14	42	860	342.2	228.4

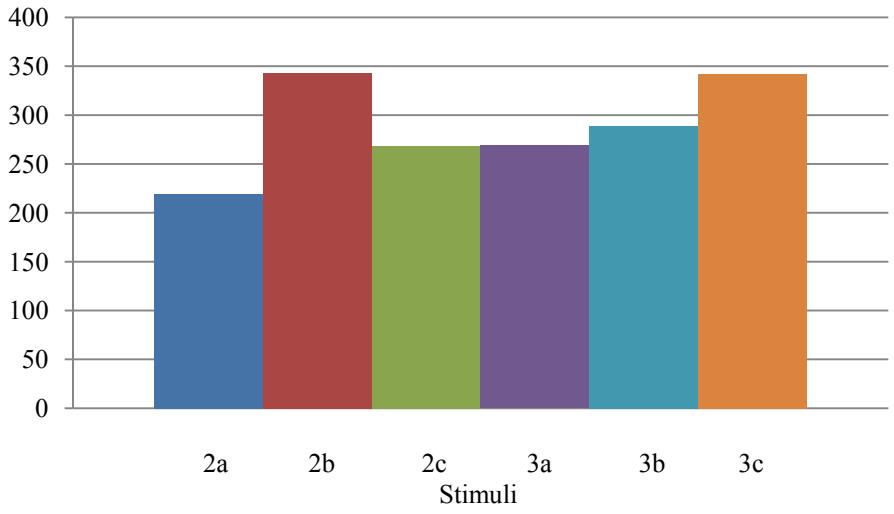


Figure 14: Time spent holding detour categories

5.1.5 Number of LIB trajectory direction reversals

As discussed earlier, high amount of LIB trajectory direction reversals indicates more incremental search and cognitive restructuring. The number of LIB trajectory direction reversals from its trajectories in the process of solving each task was calculated. Table 9 shows the results of the average number of card exchange reversals and standard deviation across the 9 experiment conditions and Figure 15 shows the graphic representation of the same results.

Table 9: Average number of LIB trajectory direction reversals and standard deviation across nine experimental conditions

Task Structure	Incentives	N	Minimum	Maximum	Mean	Std. Deviation
1-basic card sort	a-individual	14	0	2	.14	.535
	b-mixed	13	0	10	1.23	2.774
	c-group	14	0	8	1.00	2.320
2- detour	a-individual	14	0	4	.57	1.222
	b-mixed	13	0	6	1.38	2.022
	c-group	14	0	8	1.07	2.433
3-detour and restructuring	a-individual	14	0	4	1.50	1.698
	b-mixed	13	0	9	3.77	2.948
	c-group	14	1	22	6.57	6.394

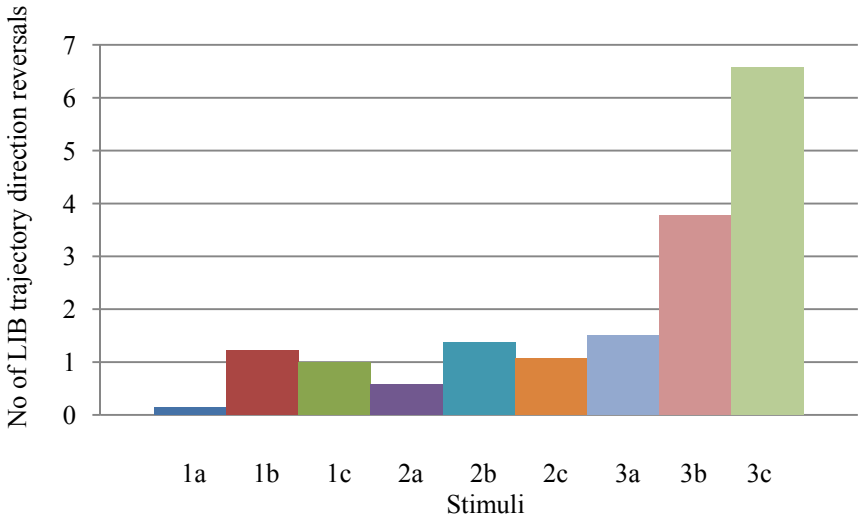


Figure 15: Average number of card exchange reversals across nine experimental conditions

For each fixed incentive, as the structure of the task increases and become more complicated, more LIB reversals were exhibited as expected, indicating that more restructuring and incremental search are required. Moreover, as the incentive switches from individual incentive to group incentive, the magnitude of the increasing in LIB reversals with increasing task complexity is amplified, indicating a potential interaction effect.

5.1.6 Typical LIB trajectory for three structures

This section provides examples of three typical LIB trajectories under three structures. Figure 16 shows a typical LIB trajectory in task structure 1(basic condition). LIB decreased monotonically over time. Figure 17 shows a typical LIB trajectory for task structure 2 (detour condition). Two small reversals are exhibited. Some flat distance after the reversal points show that participants were stuck for a while at this moment and searched in the problem space before they made another move. Figure 18 shows a typical LIB trajectory for task structure 3 (detour and restructuring condition). We can see big direction reversals which indicates restructuring and incremental search. The deep valley shape in Figure 18 implies that the group was stuck in a blind alley, then they discussed for about 4 minutes before they made the next move. As they came out of the blind alley, first their LIB

increases, because the group members had given up detour categories and tried new categories, then LIB decreased rapidly as they restructured their categorization and suddenly found the right solution.

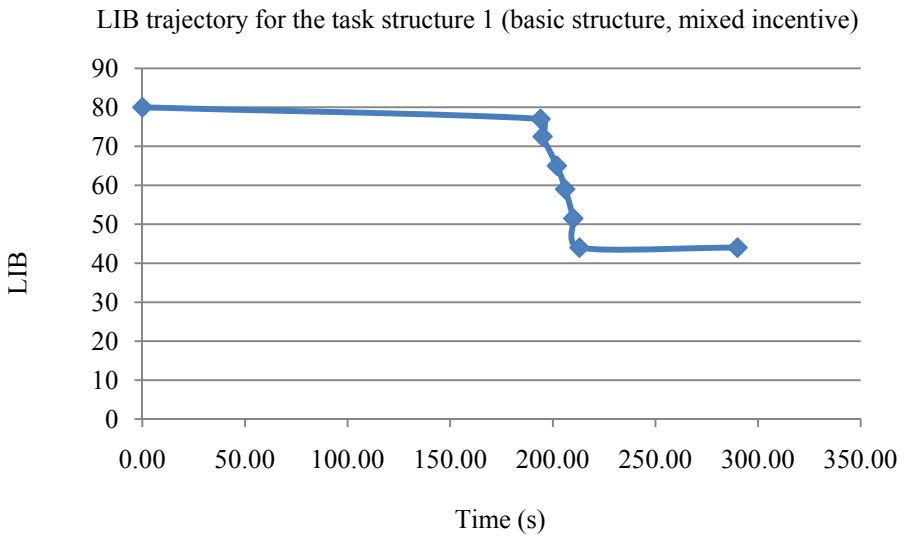


Figure 16: LIB trajectory for task structure 1 (mixed incentive)

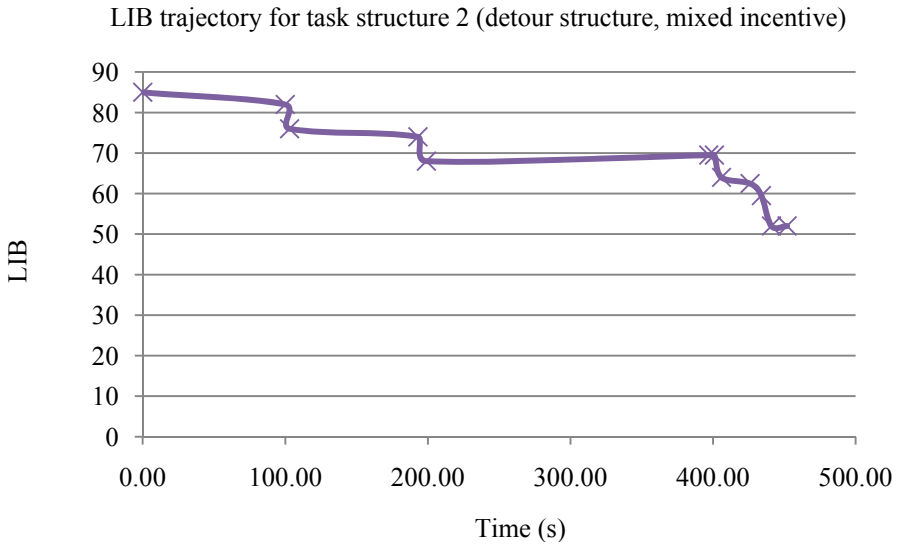


Figure 17: LIB trajectory for task structure 2 (mixed incentive)

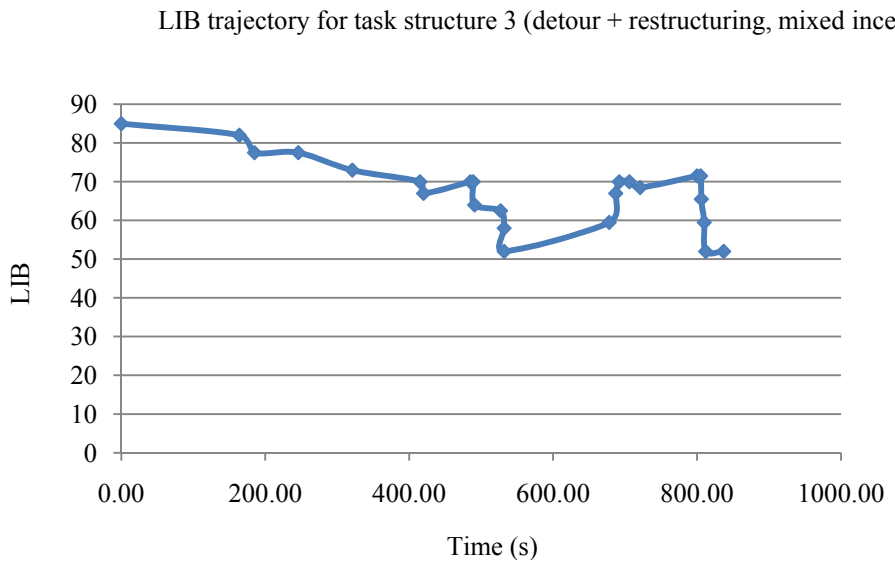


Figure 18: LIB trajectory for task structure 3 (mixed incentive)

Note that the initial LIB imbalance at time 0 is different for problem structure 1 (basic, LIB = 80) than for problem structure 2 (detour, LIB = 85), and 3 (detour and restructuring, LIB = 85). This is because detour category did not exist in problem structure 1, but existed in both problem structures 2 and 3. In problem structures 2 and 3, pictorial items from each detour category were positively associated to each other, however, dispersed among four participants. Thus this resulted in higher number of LIB imbalance of the solution state in problem structures 2 and 3 than problem structure 1.

5.2 Non parametric analysis

This section tests the effect of two factors—structure and incentive—on problem solving behaviour and performance. Since the variance for different stimuli on most variables are not equal, non-parametric analysis was performed.

5.2.1 Effect of structure under individual incentive

Results of three different structures (1, 2, and 3) were compared to test the effect of structure on group performance and behaviour under individual incentive (condition a). Table 10 shows the results from Kruskal-Wallis analysis on structure 1, 2 and 3 with fixed incentive (a). It shows that only time and number of card exchange reversals were significantly different across the three conditions.

Table 10: Kruskal-Wallis[a] analysis on three problem structures under individual incentive

	(1) Time to solve (s)	(2) No of Card Exchanges	(3)No of detours	(4) No of LIB reversals	LIB Rsquare by polynomial order				
					(5) 1st Order	(6) 2nd Order	(7) 3rd Order	(8) 4th Order	(9) 5th Order
Chi-Square	20.738	5.891	1.421	7.296	.310	4.069	1.826	2.319	2.535
df	2	2	1	2	2	2	2	2	2
Asymp. Sig.	<.001	.053	.233	.026	.856	.131	.401	.314	.281

[a] Results of Mann–Whitney test for pair-wise differences between experimental conditions: (2) vs. (3) variable 1 ($p < .0001$); (1) vs. (3) variables 2, 4 and 6 ($p < 0.05$), variable 1 ($p < 0.001$).

The Mann-Whitney test was used to test pair-wise differences between the 3 structure conditions. Mann-Whitney test results for structure (1 versus 2) show no significance on any variable. This shows that there was no significant difference between behaviours in solving task 1 and task 2 under individual incentive. Mann-Whitney test result for structure (2 versus 3) show significant difference only for "time to solve". This indicates that with the introduction of restructuring, the group spent more time to solve the problem; however, a lack of significance on other variables indicates little difference in problem solving behaviours between the two structures under individual incentive. Neither is there any significance indicates the number of detour a group fell into changes under individual incentive. Mann-Whitney test for structure (1, 3) shows significance on time, number of card exchanges, and reversals of card exchanges. This indicates that more time was spent to solve the task with detours and restructuring and group level search behaviour also increased; however, this increase was not obvious between either task 1 and task 2 or task 2 and task 3. It only becomes obvious when task 1 and task 3 were compared. This might show that the increases on these measurements were small between each level of structure that no significance can be detected unless the effects of two levels were counted together; however, no significance on the shape of LIB trajectories (variables 4-9 in Table 10) indicates any significant change on the shape of the path leading to the solution.

5.2.2 Effect of structure under mixed incentive

Results of three different structures (1, 2, and 3) were compared to test the effect of structure on group performance under mixed incentive (condition b). Table 11 shows the results from Kruskal-

Wallis analysis on structures 1, 2 and 3 under mixed incentive. Time, number of card exchanges, number of detours the group fell into, number of reversals, and LIB R square second order to fourth order were all significant. This indicates that the group performance and behaviours show significant difference on different aspects in solving three types of problems when the mixed incentive was given. This also shows a greater difference in behaviours across three structures than groups given the individual incentives.

Table 11: Kruskal-Wallis[a] analysis on three structures with mixed incentive

	(1) Time (s)	(2) No of Card Exchanges	(3)No of detours	(4) No of reversals	LIB Rsquare				
					(5) 1st Order	(6) 2nd Order	(7) 3rd Order	(8) 4th Order	(9) 5th Order
Chi-Square	19.333	13.120	8.526	7.822	.530	8.895	12.575	11.066	5.945
df	2	2	1	2	2	2	2	2	2
Asymp. Sig.	<.001	.001	.004	.020	.767	.012	.002	.004	.051

[a] Results of Mann–Whitney test for pair-wise differences between experimental conditions: (2) vs. (3) variable 8 ($p < .0001$), variable 1, 7 (p< .001), variable 2, 3, 6 (p< .01), variable 4, 9 (p< .05); (1) vs. (3) variable 1 ($p < 0.001$), variable 2 (p<.001), variable 6, 7 (p< .01), variable 4, 8, 9 (p<0.05).

Mann-Whitney test was used to test pair-wise differences between the 3 structure conditions. Mann-Whitney test results for structure (1 versus 2) shows no significant difference on any variable. Mann-Whitney test for structure (2 versus 3) show significance on time, number of card exchanges, detours the group fall into, number of reversals, and LIB R square second order to fifth order. This indicates that with the introduction of restructuring, the group spent more time to solve the problem. Significance of LIB R square second order to fourth order show that the paths leading to the solution were different between two conditions. The group tends to collect the detour category more with the existence of restructuring and thus the number of reversals also increases. This was not shown under individual incentive. Mann-Whitney test for structure (1, 3) show significance on time, number of card exchanges, detours the group fall into, number of reversals, and second order to fifth order. This is similar to the Mann-Whitney test result for structure (2, 3). This indicates that under the mixed incentive, restructuring are playing a more obvious role on these measurements on group performances than detours. Compare this to the individual incentive, the biggest difference is that under individual incentive, despite of more complicated structure, the path groups have leading to the solution is more or less smoother and more direct than these for groups under the mixed incentive.

5.2.3 Effect of structure under group incentive

Results of three different structures (1, 2, and 3) were compared to test the effect of structure on group performance and behaviours under group incentive (condition c). Table 12 shows the results from Kruskal-Wallis analysis on structures 1, 2 and 3 under group incentive. Time, number of card exchanges, detours the group fall into, number of reversals, and LIB R square third order to fifth order are all significant. This indicates that the group performance and behaviours show significant difference on different aspects in solving three types of problems when group incentive was given. This also shows a greater difference in behaviours across three structures than groups given the individual incentives. Compared to the groups given the mixed incentives, significances in measures are the same except for LIB R square fifth order.

Table 12: Kruskal-Wallis[a] analysis on three structures under group incentive

	(1) Time to solve (s)	(2) No of Card Exchanges	(3)No of detours	(4) No of reversals	LIB Rsquare				
					(5) 1st Order	(6) 2nd Order	(7) 3rd Order	(8) 4th Order	(9) 5th Order
Chi-Square	22.214	24.683	19.705	20.048	1.667	2.458	17.577	17.876	13.778
df	2	2	1	2	2	2	2	2	2
Asymp. Sig.	<.001	<.001	<.001	<.001	.435	.293	<.001	<.001	.001

[a] Results of Mann–Whitney test for pair-wise differences between experimental conditions: (2) vs. (3) variable 1, 2, 3, 4, 7, 8 ($p < .0001$), variable 9 ($p < .01$); (1) vs. (3) variable 1, 2, 4, 7, 8, 9 ($p < 0.001$).

Mann-Whitney test for structure (1 versus 2) show no significance on any variable. This shows that there was no significant difference between behaviours in solving task 1 and task 2 under the group incentive. This is consistent across three incentives, which might imply that detours are not playing an obvious role in the problem structure, and neither did any of the three incentives change that. Mann-Whitney test for structure (2, 3) show significance on time, number of card exchanges, detours the group fall into, number of reversals, and LIB R square second order to fifth order. This indicates that with the introduction of restructuring, the group spent more time to solve the problem. Significance of LIB R square second order to fifth order show that the paths leading to the solution were different between two conditions. The group tends to collect the detour category more with the existence of restructuring and thus the number of reversals also increases. This was not shown under

individual incentive. Compared to the mixed incentive condition, there is significance on LIB R square fifth order, which indicates that the path leading to the solution of task 3 becomes more complicated in shape and exhibits more reversals with the introduction of group incentive. In general, the path leading to the solution of task 3 becomes more and more complicated in shape and exhibits more reversals as the incentive changes from pure individual to mixed to group incentive. Mann-Whitney test for structure (1 versus 3) show significance on time, number of card exchanges, number of reversals, and LIB R square second order to fifth order. This is similar to the Mann-Whitney test result for structure (2 versus 3). This indicates that under the group incentive, restructuring plays a more significant role on these measurements on group performances and behaviours than detours.

5.2.4 Effect of incentives under problem structure 1—basic condition

Results of three different incentives (a, b, and c) were compared to test the effect of structure on group performance under the basic structure (structure 1). Table 13 shows the results from Kruskal-Wallis analysis on incentives (a, b, and c) on problem structure 1. None of the variables is significant. None of the Mann-Whitney test for incentive condition (a versus b), (a versus c) and (b versus c) show any significance on any variable. We can conclude that incentives do not play a significant role on performances on solving the simplest task where there is no blind alley and every group could solve the problem directly without struggle.

Table 13: Kruskal-Wallis analysis on three incentives under basic structure

	(1) Time to solve (s)	(2) No of Card Exchanges	(3) No of reversals	LIB Rsquare				
				(4) 1st Order	(5) 2nd Order	(6) 3rd Order	(8) 4th Order	(8) 5th Order
Chi-Square	1.569	.671	2.440	1.205	.480	1.306	1.815	.503
df	2	2	2	2	2	2	2	2
Asymp. Sig.	.456	.715	.295	.548	.787	.521	.404	.778

5.2.5 Effect of incentives under problem structure 2—detour

Results of three different incentives (a, b, c) were compared to test the effect of structure on group performance and behaviours under problem structure 2—detour. Table 14 shows the results from Kruskal-Wallis analysis on incentives (a, b, c) on detour problem structure. None of the variables are

significant. None of the Mann-Whitney pair-wise test for Incentive condition (a versus b), (a versus c) and (b versus c) show significance on any variable. We can conclude that incentives do not play a significant role on performances on solving task 2.

Table 14: Kruskal-Wallis analysis on three incentives under detour structure

	(1) Time to solve (s)	(2) No of Card Exchanges	(3)No of detours	(4) No of reversals	LIB Rsquare				
					(5) 1st Order	(6) 2nd Order	(7) 3rd Order	(8) 4th Order	(9) 5th Order
Chi-Square	1.533	2.495	.529	.386	2.031	4.379	.032	4.234	3.176
df	2	2	2	2	2	2	2	2	2
Asymp. Sig.	.465	.287	.768	.824	.362	.112	.984	.120	.204

5.2.6 Effect of incentives under problem structure 3—detour and restructuring

Results of three different incentives (a, b, c) under problem structure 3 were compared to test the effect of structure on group performance and behaviours under the problem structure of detour and restructuring. Table 15 shows the results from Kruskal-Wallis analysis on incentives (a, b, c) on problem structure 3. Number of card exchanges, number of detours the group fell into, number of reversals, and LIB R square 3rd order to 5th order are all significant. This indicates that the groups under different incentives behaved differently in solving the problem with structure 3. This makes structure 3 very different from the other two structures as this is the only task that incentives affect performances.

Table 15: Kruskal-Wallis[a] analysis on three incentives under detour and restructuring structure

	(1) Time to solve (s)	(2) No of Card Exchanges	(3)No of detours	(4) No of reversals	(5) Time spent in detour categories	LIB Rsquare				
						(6) 1st Order	(7) 2nd Order	(8) 3rd Order	(9) 4th Order	(10) 5th Order
Chi-Square	1.377	12.563	10.304	8.483	0.754	0.456	0.518	12.466	11.515	6.238
Df	2	2	2	2	2	2	2	2	2	2
Asymp. Sig.	0.502	0.002	0.006	0.014	0.686	0.796	0.772	0.002	0.003	0.044

[a] Results of Mann–Whitney test for pair-wise differences between experimental conditions: (1) vs. (2) variable 7, 8 ($p < .01$), 2, 4, (p, .05); (1) vs. (3) variable 2 ($p < .0001$), variable 8 ($p < .001$), variable 3, 4, 9 ($p < .01$), variable 10 ($p < .05$).

Mann-Whitney test for Incentive (a, b) show significance on number of card exchanges, number of reversals, and LIB R square second order to fourth order. This indicates that with the introduction of partial group incentive, the group exchanged more cards to solve the problem. Significance of LIB R square second order to fourth order show that the paths leading to the solution were different between two conditions. This indicates that the path leading to the solution of task 3 becomes more and more complicated in shape and exhibits more reversals as the incentive changes from pure individual to mixed. The group tends to struggle more with the existence of partial group incentive and thus the number of reversals also increases. This was not shown under the other two structures. Mann-Whitney test for incentive (a, c) show significance on number of card exchanges, detours the group fall into, number of reversals, and LIB R square third order to fifth order. This is similar to the Mann-Whitney test result for incentive (a, b) but only stronger. However, Mann-Whitney test for incentive (b, c) show no significance on any variable. This indicates that under structure 3, the existence of the concept of "group incentive" plays more significant role than the weight of "group incentive".

So far, incentives do not result in significance difference on the group performance and behaviours on task 1 or task 2. However, Table 15 shows that incentive plays a significant role on performances on task 3. The performances showed difference between individual incentive cases and both non-individual incentive cases; however, the performances showed no difference between partially group incentive and group incentive cases. In conclusion, the effect of incentives on group problem solving behaviours can be presented as follows:

1. With the introduction of group incentive, whether it counts for 50% of the bonus or 100% of the bonus, groups made significantly more exchanges of cards and more reversals in the trajectories simulating the searching for right solution.
2. With the introduction of group incentive, whether it counts for 50% of the bonus or 100% of the bonus, it becomes significantly more likely for groups to collect the detour category.
3. With the introduction of group incentive, whether it counts for 50% of the bonus or 100% of the bonus, the group performed much more varieties in searching behaviours.

5.3 Interaction between incentives and structures:

Two-way analysis of variance was used to test the interaction between incentives and structures. Although some variables do not have equal variance among all sample groups, results are still adopted since there is no alternative way to do a two-way non parametric analysis using available statistical software (SPSS 17.0). Table 16 provides a picture of the F value and significance level for different variables corresponding to the interaction between incentives and structures. It shows that the interaction between incentives and structures is significant on these variables: number of card exchanges, number of detours each group fall into and the number of reversals on card exchanges. The direction and detail of influence have been discussed in previous sections.

Table 16: Interaction between incentive and structure from Two-way ANOVA

Variable	F	Sig
Time to solve	.412	.800
No of card exchanges	4.382	.002
No of detours	6.365	.003
No of reversal	2.753	.031
R square 1st order	.477	.753
R square 2nd order	.119	.976
R square 3rd order	2.031	.095
R square 4th order	.974	.425
R square 5th Order	.301	.877

5.4 Test of hypotheses

Hypothesis 1: As the complexity of problem structure increases, the time to solve the problem and the complexity of search behaviour both increase.

Test of Hypothesis 1: Kruskal-Wallis analysis on time and complexity of search behaviour measures across three task structures has tested this hypothesis in previous sections.

Test result of time, Variable 1 in Table 10, Table 11 and Table 12 show that differences on time to solve the problem are significant ($p<.001$) across three task structures under every incentive condition. General results support Hypothesis 1. However, Mann-Whitney test show that this difference is mainly from task structure 3, and no difference between the time to solve task 1 and task 2 under any incentive. This will be discussed in the next section.

Test result of complexity of search behaviour measure, Variable 2 (No. of card exchange), Variable 4 (Reversals), Variable 6-9 (LIB R squared values) in Table 10, Table 11 and Table 12 also support this hypothesis in general.

> Hypothesis 2a: Under the basic task structure, individual versus group incentives will affect neither the time to solve the problem nor the complexity of search behaviour.

This hypothesis is tested by Kruskal-Wallis analysis across three incentives under basic structure. The results in Table 13 show that none of the measures used in this experiment is significant across three incentive conditions. Thus this hypothesis is supported by the results.

> Hypothesis 2b: Under the task structures of detour and detour plus restructuring, as incentives switch from group to individual, the time to solve the problem increases.

Under the structure of detour, this hypothesis is tested by Kruskal-Wallis analysis on measures across three incentives. Results in Table 14: Kruskal-Wallis analysis on three incentives under detour structure indicates that individual versus group incentive does not make any difference on either time or complexity of search behaviours.

Under the structure of detour plus restructuring, this hypothesis is tested by Kruskal-Wallis analysis on time and measures on complexity of search behaviour across three incentive conditions. Results in Table 15 show significant difference on complexity of search behaviours (variable 2, 3, 4, 8-10) as

incentive changes. However, the result on variable 1 (Time) is not significantly different across three incentive conditions. This will be discussed in the next section.

We also expected that individuals under individual incentive would be attracted to the sub-optimal solution to the group, in this case, detour categories. However, interestingly, the results show the opposite: groups under group incentive actually collected detour categories more than groups under individual incentive under the structure of detour plus restructuring. In conclusion, Hypothesis 2b is partially supported by the results and this will be discussed in the next section too.

> Hypothesis 2c: As incentives switch from group to individual, the effect of structural complexity on time to solve the problem and complexity of search behaviour will increase.

Two-way ANOVA analysis is used to test the interaction effect between structure and incentive. Results in Table 16 show that there is no interaction on time, however, the effect of task structure complexity show increasing effects on some variables of complexity of search behaviours (No. of card exchanges, No. of detour., and No. of reversals) as incentives switch from individual to group, opposite to the prediction. This will be discussed in the next section too.

Chapter 6
Discussion

6.1 Findings and discussion

This thesis investigated the effects of individual and group incentives on group performances and behaviours in three different types of problem solving scenarios. Previous research has focused primarily on individual problem solving or group behaviour. There has been limited study on the influence of the problem itself and incentives on group problem solving due to the difficulty of developing an efficient task that allows both cognitive restructuring and incremental search to be studied with flexible manipulation of problem structure and incentives. This study investigated the effects of individual and group incentives on three typical structures of problem by using a well developed system of problem structure and varying the external force on the group.

6.1.1 Incentives

Results of the study show that incentives influenced the group performance and behaviour by affecting the path groups took to approach the problem. Although groups under different incentives solved the problem in similar time ranges, the trajectory of LIB of each process show differences under different incentives in terms of the pattern of the path between the start of the problem space and the solution. The path leading to the solution simulated by the LIB trajectory shows that different strategies were developed associated with different incentives under certain type of problems. Groups under the group incentive were more willing to take random approaches and exploration while groups under the individual incentive were less willing to make card exchange prior to careful discussion which leads to discovery of the path to the answer.

When the path to the solution is clear to most groups, this difference in strategy is not obvious in terms of performances because either random exploration or planning leads to the solution directly in similar time range. An example is the basic structure (structure 1) in this study. Structure 2 has two detour sets of cards, thus instead of four sets of categories, six categories exist in structure 2; however, only 4 of them can form the right answer. If one picks a category to collect randomly, then this leaves a one third chance for the two detour categories to be collected. The detour categories were also designed to be more or equally attractive, which was expected to increase the chance for them to be collected. However, results show that the percentages of groups who collected the detour

categories is lower than one third regardless of incentives. The percentages are 21%, 23% and 14% under incentives a, b and c respectively. The results show that participants did not select category to collect randomly; instead, participants tend to give priorities to categories that can accommodate with other group members' categories. This implies that some groups acknowledged the existence of some detour category but decide not to collect them because they realize that this solution would cause failure of others. Thus the explanation for the low percentage of detours groups fell into under structure 2 is that the complexity of this structure is lower than participants' capacity to foresee steps before making a decision of which category to.

When the path to the solution is not clear to most groups, this difference in strategy is clear in terms of performance because random exploration result in a much higher chance to collect the detour category than careful planning that leads to the solution directly. The numbers of groups that collected detour category under structure 3 (detour and restructuring) showed significant difference between incentive a) and the other two incentives. The percentages were 43%, 92% and 100% under incentives a, b and c respectively. This indicates that the restructuring items with detour together are so complicated that a group cannot identify the right solution as easily as in structure 1 and 2.

In terms of interaction between incentives and structure, structure 3 (detour and restructuring) amplified the difference of strategies groups used to search for the solution within a problem space under different incentives.

Problem solving can be viewed as the process of moving from an initial state to the solution (goal) state. The results indicated that under structure 3 (detour and restructuring) groups adopt different strategies to do this depending on incentives. Groups under individual incentive tend to discuss the solution thoroughly and formulate a good model of the problem structure before making moves to exchange cards. Thus if they identify a blind alley in the problem space, they will construct a path to avoid the blind alley. Note that the chance they fall into the detour in the first place is lowered because of this process. This is because when groups begin to solve a problem in structure 3, there are only two sets of cards that are obviously complete—the two detour categories. The four correct solution categories are initially answer is hidden because of the restructuring items which at first sight seem to be irrelevant. For the individual incentive, this situation under individual incentive has properties similar to a zero-sum game, where any two people can obtain the gain at the cost of the

other two members' loss. Consequently, no one is willing to sacrifice by exchanging cards to help someone else collect a detour category. This hesitation to exchange cards forces them to seek other possibilities and thus helps them restructure the correct solution categories. On the other hand, groups under group incentive do not have such concerns because their bonus mark is gained at the same time through group performance. Thus rather than formulating a complete representation of the problem, they tend to go ahead and explore randomly throughout the problem space, and the high chance of getting stuck in the blind alley is not surprising.

Each of these strategies may offer benefits for different types of problems. If there is one and only one fixed answer to a problem, the individual incentive might have advantages in solving the problem because groups under individual incentive pay more attention on the solution state and develop strategy that would lead to the solution directly. Similarly in cases where wrong move would cause terrible consequence or huge cost, individual incentive may be able to minimize the likelihood of any mistake. On the other hand, if the cost of mistakes is low relative to potential rewards, or if the problem has more than one fixed answers or requires creative thinking or exploration, group incentive may be appropriate because groups under group incentive are more willing to explore and take random incremental search.

Note that in task 3, the results showed differences between the individual incentive cases and both non-individual incentive cases; however, the results showed little difference between the mixed incentive condition and the group incentive condition. With the introduction of group incentive, regardless of whether it counts for 50% (mixed incentive condition) of the bonus or 100% (group incentive condition) of the bonus, more problem search activities were observed, such as significant more exchanges of cards more reversals in the LIB trajectories simulating the searching for right solution, and greater likelihood of collecting the detour category. This indicates that it is the concept of "group" that is important to the change of behaviour. With the existence of "group" concept, it seems that individuals in the group are more willing to explore and take random approaches or even risky approaches.

6.1.2 Problem structures and interaction with incentives

Results also show that the three problem structures were associated with differences in performances and problem solving behaviour. With any fixed incentive, the group performance and behaviours

were significantly different between structure 2 and structure 3, as well as between structure 1 and structure 3. In general, the factor of restructuring complicated the problem and resulted in longer time to solve the problem, more card exchanges, more LIB trajectory reversals and higher order of polynomial regression to simulate the path. Incentives interacted with problem structure and these differences were shown to be amplified as the incentive changed from individual to mixed/group incentive.

However, the factor of detours didn't complicate the problem very highly. The results show no significance on any variable when comparing problem structure 1 and 2. This could result from two possibilities: 1. Detour structure itself might not be significantly more difficult than basic structure. As we discussed earlier, the participants tend to give priorities to categories that can accommodate with other group members' categories. It shows the complexity of pure detour is low enough compared to the human intelligence and capacity to avoid the detour and make decisions oriented to an ideal solution. 2. Groups took longer time in task 1 than expected due to two reasons. One is that observation shows that many groups discussed the problem solving strategy during task 1, and then used the same strategy in task 2 and 3. Thus, the time recorded to solve task 1 includes the time they spent to discuss strategy, which is absent in the time recorded to solve task 2 and 3. Another possible reason is that, the meaning of the solution changed from the two training sessions to the three experimental tasks. In the two training sessions, "four cards of the same kind" meant four identical in cards; while "four cards of the same kind" in task 1-3 meant four similar but different pictures. This change itself was another factor that may have added difficulty and time required to complete task 1 relative to task 2. It would be ideal if future researchers could eliminate this effect. For example, each group could only solve one task instead of three.

Based on above discussion, we draw the conclusion that task structures influence group performance and behaviours and incentives play an important role in the strategy the group uses to solve the problem. Individual incentive encourages the group to focus on the formulating a good conceptual model of the problem and identifying a path that take them directly to the solution while group incentives encourage random exploration.

6.1.3 Comparisons to existing literature

Fisher (1994) suggests that individual incentives be used to motivate key workers in traditional assembly lines, where tasks are straightforward in a group. This case can be viewed as a situation where the path to the solution state is straightforward, and our findings support this view as our finding suggests that individual incentives encourage direct discovery of the solution state.

Mitchell and Silver (1990) explained that a mixed incentive system works best when the accomplishment of individual goals are aligned with the accomplishment of the group goal. They also explained, however, that when group and individual goals conflict, dysfunctions can result. Our findings do not support this argument as it is ambiguous to determine if an individual goal is aligned with the accomplishment of the group goal. For example, when an individual was collecting a detour category under task 3, the individual or even the whole group might view this action as aligned with the accomplishment of group goal if they think by doing so they are half way to the end; however, in reality, this action creates difficulty and leads the group to a blind alley which does not lead them to the group goal.

Libby and Thorne (2009) found that when group members can provide useful information about how to perform the task better, group incentives resulted in higher group performance and behaviours than under individual or mixed incentives. One possible explanation is that group incentives encourage knowledge sharing, positive interaction, as well as lower competition within the group. In our experiment, one would not finish the task without information from any other group member because the only way to reach the solution is to exchanges of cards. Group incentive resulted in higher group activities such as more card exchanges; however, group incentive does not show significant difference from mixed incentive condition.

This lab experiment was unique in the sense that it provides an insightful view of a situation where group members need to share information to solve a problem without any existing answer. First, it simulates a real working environment better than previous tasks such as the tower building experiment where a few group members together use blocks to build a tower. Besides, results of this study show that no behaviour difference was observed on simple tasks where the path to the answer is clear and straightforward. Thus the observation in the tower building experiment was limited because the path to the goal is transparent to every member. Second, the design of the task forces participants

to communicate verbally and physically (by exchanging cards). This simulation of the information flow in a modern working environment was rarely obtained in other experiments. Third, little research has investigated problem solving situations where group and individual goals conflict, or when individuals in a group are not aware of a clear path towards their goal. The conflict of individual and group goals contributed to the long discussion among group members prior to physical card exchange, and thus shaped a unique path that lead to the solution directly. This is an additional interesting discovery of this study.

6.2 Limitations

This study has tried to simulate real-world problem solving in a lab environment. Although the design of experimental tasks has been carefully controlled, there are still some limitations:

1. Some groups were not sure about the solution they got even when they had the right solution, and thus spent part of their time discussing even after they had solved the problem. The time recorded in this study was the time from the beginning of a task to the time a group verbally told the experimenter that they were done, so the time recorded in this study was the time a group used to solve the problem and also cognitively believed that they had got the right answer. It is hard to determine whether the time a group used to convince themselves that they had the right answer should be counted or not. However, in real problems in organizations, this is usually the case. This reality of ambiguity in the definition of "problem solving" is one of the limitations of the study. It will be ideal if all groups could have the same degree of discussion behaviour; however, this is highly dependent on personalities of group members, and thus difficult to control.
2. Limited students and time. Ideally each task could be performed by one group who had not participated in any other task to avoid any learning effects across tasks. However, due to a limited number of students and time and difficulties with organizing groups of four students, we chose to have each group of students perform three tasks structure in the order of simplest to hardest: basic→detour→detour plus restructuring. Although we have attempted to avoid learning effects by using different stimuli and pictorial items for each task, and by arranging it in such order, there may still be some learning effect. Students in one group usually don't know each other, and they tend to discuss strategy of the whole group in the first task (basic condition), while they follow the same pattern in the following two tasks. This caused one

inequality in behaviours in three tasks; that is, the first task being performed by the group involves the behaviour of discussing strategy while other tasks usually do not. This could be avoided if each group only performs one task; however, due to the limited number of students that can be accessed, we were only able arrange experimental tasks in this way. On the other hand, such learning effect should reduce observed performance and behavioural differences between conditions, thus our results here should provide conservative measures of these differences.

6.3 Suggestions for future study

As discussed earlier, under structure 3, both mixed incentive and group incentive resulted in much more group level activities such as card exchanges, reversals in path simulating the searching for right solution, etc. However, there was no observed significance on variables when comparing the mixed incentive condition with the group incentive condition. The concept of "group" and how it works on the perception of individuals in a group is of great interest. We hypothesize that with the concept of "group", a group of individuals are more willing to take risk, more open in mind, and more flexible in terms of making decisions. The labelling of "group" might decrease the sense that the output belongs to an individual, and weaken the association of responsibility to the output and each individual in a group. However, at which ratio of individual verse group incentive does this switch in behaviour occur? Future studies could manipulate additional ratios and investigate the "switch point".

So far researchers have investigated the effect of the problem itself (structure) and external forces (incentives), but have not yet investigated the influence of individual personalities. Both mental operations oriented to the environment and the self contribute to the process of problem solving. As the participant of the experiment, each individual handles the information perceived, absorbs input from the environment (i.e. the incentive, and influence from each other) and responds to the information, thus contributing to the output (group behaviour and performance). Each individual's personality might play a role in two ways: one is that the personality influences how one reacts to an environmental input; the other is that the personality influence other group members and contribute to the dynamic of this process. This also leaves a large field to be explored in future research.

References

Abimbola, G. A. (2006). *Effects of task structure on group problem solving.* University of Waterloo, Waterloo.

Adejumo, G., Duimering, P. R., & Zhong, Z. (2008). A balance theory approach to group problem solving. *Social Networks, 30*(1), 83-99.

Bavelas, A. (1973). The five squares problem: an instructional aid in group cooperation. *Studies in Personnel Psychology, 5*(2), 29-38.

Cartwright, D., & Harary, F. (1956). Structural balance: a generalization of Heider's theory. *Psychological Review, 63*, 277-293.

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

Doreiana, P., & Mrvarb, A. (1996). A partitioning approach to structural balance *Social Networks, 18*(2), 149-168.

Duncker, K. (1945). *On problem solving*: American Psychological Association

Fisher, J. G. (1994). Technological interdependence, labor production functions and control systems. *Accounting, Organizations and Society, 19*(6), 493-505.

Goldberg, M. H., & Maccoby, E. E. (1965). Children's acquisition of skill in performing a group task under two conditions group formation. *Journal of Personality and Social Psychology, 2*(6), 898-902.

Guthrie, J. P., & Hollensbe, E. C. (2004). Group Incentives and Performance: A Study of Spontaneous Goal Setting, Goal Choice and Commitment. *Journal of Management, 30*(2), 263-284

Heider, F. (1946). Attitudes and cognitive organization. *The Journal of Psychology, 21*, 107-112.

Kershaw, T., & Ohlsson, S. (2001). Training for insight: The case of the nine-dot problem. In J. Moore & K. Stenning (Eds.), *Proceedings of the Twenty-third Annual Conference of the Cognitive Science Society* (pp. 489-493). Mahwah, NJ: Lawrence Erlbaum Associates.

Köhler, W. (1969). *The Task of Gestalt Psychology*. Princeton, NJ: Princeton Univ Pr

Laughlin, P. R., Bonner, B. L., & Miner, A. G. (2002). Groups perform better than the best individuals on Letters-to-Numbers problems. *Organizational Behavior and Human Decision Processes, 88*(2), 605-620

Lewin, K. (1936). *Principles of topological psychology*. New York: McGraw-Hill.

Libby, T., & Thorne, L. (2009). The Influence of Incentive Structure on Group Performance in Assembly Lines and Teams. *Behavioral Research in Accounting, 21*(2), 57-72.

Locke, E. A., & Latham, G. P. (1990). *A theory of goal-setting and task performance*. Englewood Cliffs, NJ: Prentice Hall.

Locke, E. A., & Latham, G. P. (2002). Building a Practically Useful Theory of Goal Setting and Task Motivation: A 35-Year Odyssey. *American Psychologist, 57*(9), 705-717.

Mitchell, T. R., & Silver, W. S. (1990). Individual and group goals when workers are interdependent: Effects on task strategies and performance. . *Journal of Applied Psychology, 75*, 185–193.

Newcomb, T. M. (1953). An approach to the study of communication acts. *Psychological Review, 60*(6), 393-404.

Newell, A., Simon, H.A. (1972). *Human Problem Solving*. Englewood Cliffs, NJ: Prentice Hall

Nijstad, B. A. (2009). *Group performance* (2nd ed.). New York: Psychology Press.

Parksa, C., & Kerrb, N. (1999). Twenty-Five Years of Social Decision Scheme Theory. *Organizational Behavior and Human Decision Processes, 80*(1), 1-2.

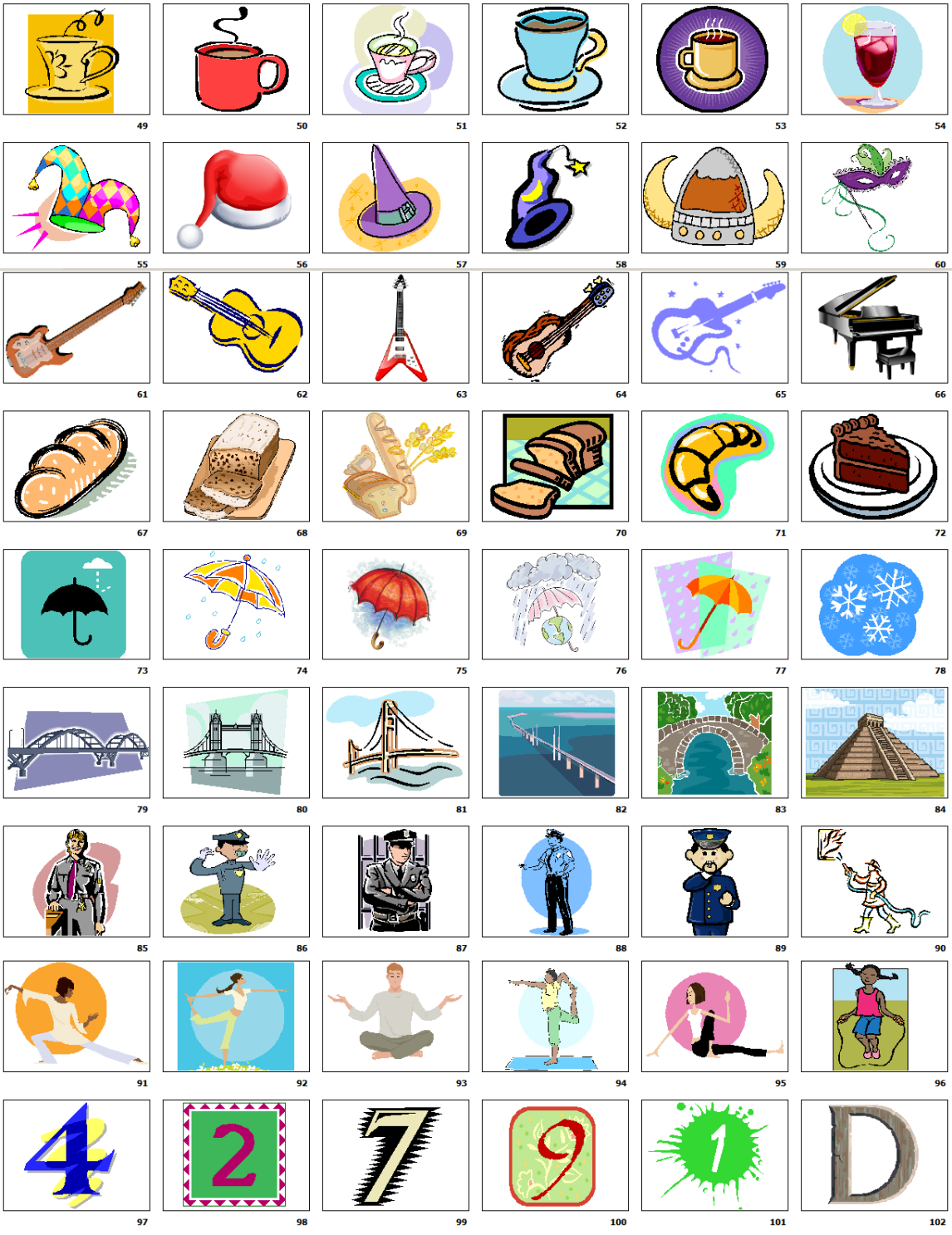
Rosch, E. (1978). Principles of Categorization. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization*. Hillsdale, NJ: Lawrence Erlbaum.

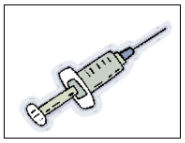
Wertheimer, M. (1982/45). *Productive thinking (Enlarged ed.)*. Chicago: University of Chicago Press.

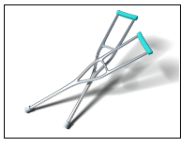
Young, S. M., J. Fisher, and T. M. Lindquist. (1993). The effects of intergroup competition and intragroup cooperation on slack and output in a manufacturing setting. *The Accounting Review, 68*(3), 466-481.

Appendix A: STIMULUS USED IN PRE-EXPERIMENT

All 138 pictures used in Pre-Experiment

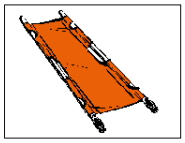

49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102

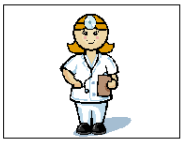

103

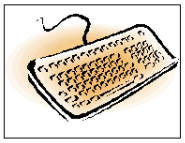

104

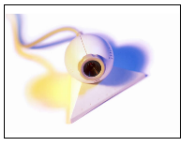

105


106


107


108


109


110


111

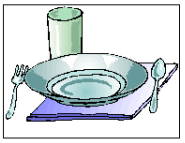

112


113


114


115

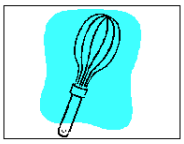

116


117

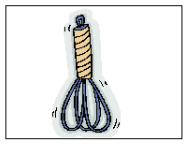

118


119

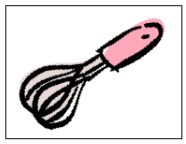

120


121

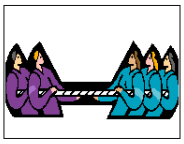

122

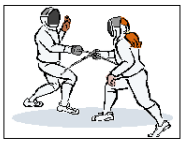

123


124

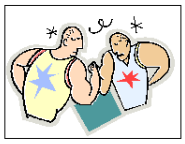

125

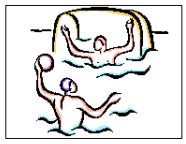

126


127


128


129


130


131


132


133


134


135


136


137


138

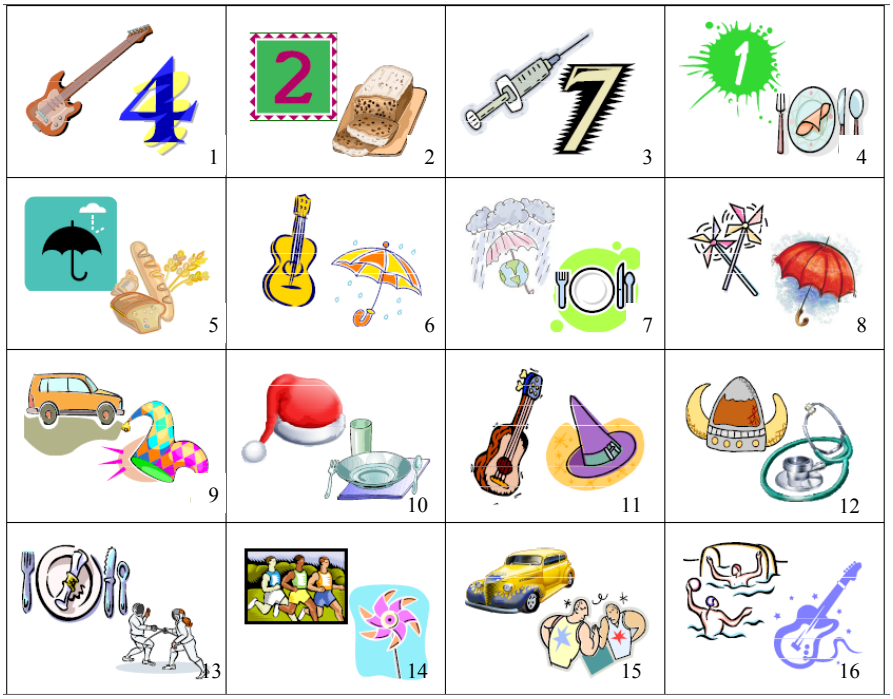
Appendix B: STIMULUS SETS USED IN EXPERIMENT

Training 1

Training 2

Stimulus 1 Structure 1- Basic Card Sort

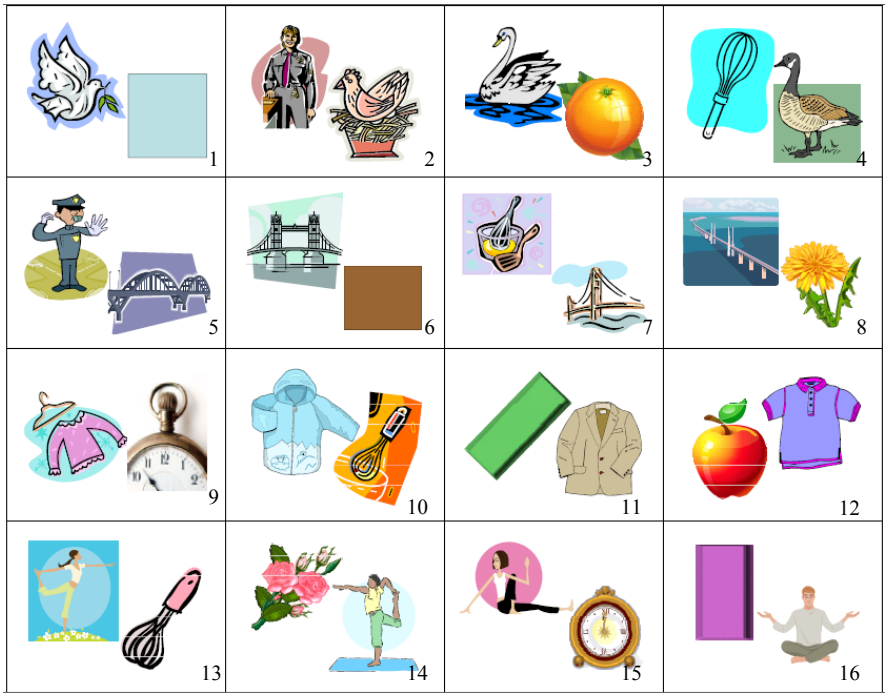
Stimulus 1 Structure 2- Detour



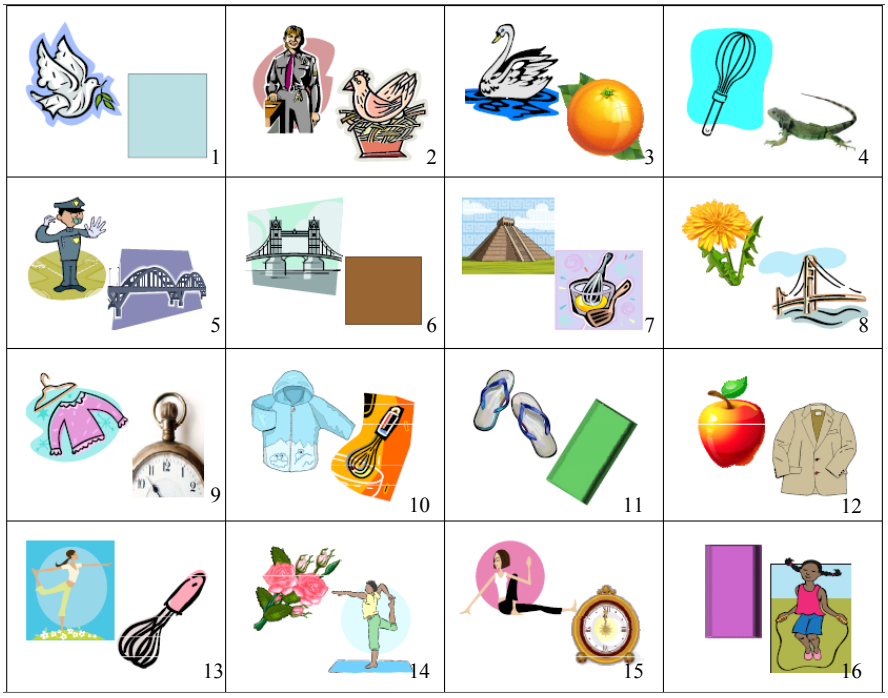
Stimulus 1 Structure 3- Detour and Restructuring

Stimulus 2 Structure 1- Basic Card Sort

Stimulus 2 Structure 2- Detour

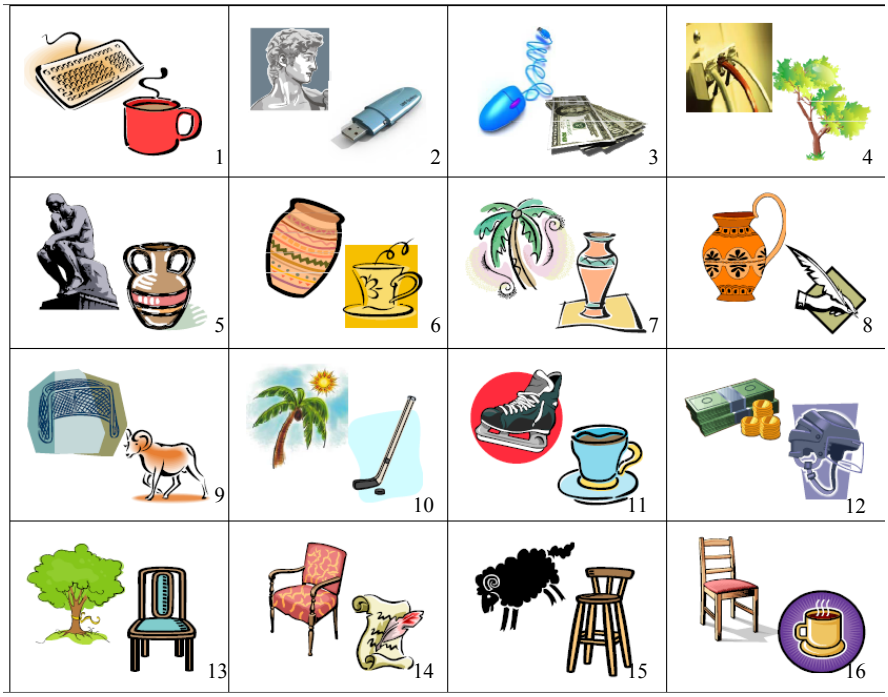


Stimulus 2 Structure 3- Detour and Restructuring

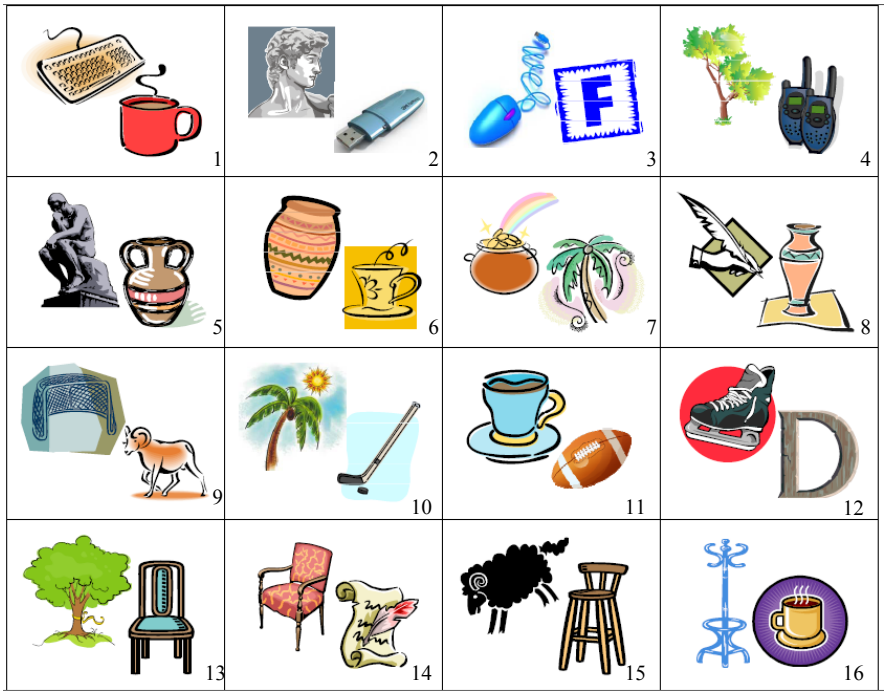


Stimulus 3 Structure 1- Basic Card Sort

Stimulus 3 Structure 2- Detour



Stimulus 3 Structure 3- Detour and Restructuring



Appendix C: INSTRUCTIONS TO PARTICIPANTS

Instructions for incentive a) - individual incentive:

1. This experiment is part of a study of how groups solve problems. Your task as a group is to work together to solve three different problems by sorting through a set of cards. After working on each of the three problems each of you will also complete a short questionnaire.
2. Before you begin, you will solve two training problems to help you understand the basic requirements of the task. Please feel free to ask any question during these training sessions. After the training sessions, your group will solve the three real problems.
3. For each problem, you will each be given four random cards to start with. The goal is to exchange cards until each group member gets a set of four of a kind.
4. You can speak to one another and describe what is on your cards. But you may not show each other your cards or look over the barrier at other peoples' cards.
5. You can trade cards as often as you like. But you may only trade one card at a time and no person can hold more than 5 cards or less than 3 cards at once.
6. You will earn up to 1 bonus mark in MSCI 311 for each problem, depending on your performance. ***In each task, if you get four cards of a kind for yourself, you will get 1 bonus mark. It does not matter if anyone else gets four cards of a kind.***
7. Please spread your cards on your section of the table so that they can be seen by the video camera overhead. Do not overlap your cards or hold them in your hand.
8. You have a maximum of 15 minutes to solve each problem, although you probably won't need that much time.
9. These tasks comply with the UW ethics and research guidelines. All data collected will be used solely for research purposes.

Thank you for your time and participation.

Good luck!

Instructions for incentive b) - mixed incentive:

1. This experiment is part of a study of how groups solve problems. Your task as a group is to work together to solve three different problems by sorting through a set of cards. After working on each of the three problems each of you will also complete a short questionnaire.
2. Before you begin, you will solve two training problems to help you understand the basic requirements of the task. Please feel free to ask any question during these training sessions. After the training sessions, your group will solve the three real problems.
3. For each problem, you will each be given four random cards to start with. The goal is to exchange cards until each group member gets a set of four of a kind.
4. You can speak to one another and describe what is on your cards. But you may not show each other your cards or look over the barrier at other peoples' cards.
5. You can trade cards as often as you like. But you may only trade one card at a time and no person can hold more than 5 cards or less than 3 cards at once.
6. You will earn up to 1 bonus mark in MSCI 311 for each problem, depending on your performance. ***In each task, if you get four cards of a kind, you will get 0.5 bonus marks. If every person in your group gets four cards of a kind at the same time, you will get another 0.5 bonus marks.***
7. Please spread your cards on your section of the table so that they can be seen by the video camera overhead. Do not overlap your cards or hold them in your hand.
8. You have a maximum of 15 minutes to solve each problem, although you probably won't need that much time.
9. These tasks comply with the UW ethics and research guidelines. All data collected will be used solely for research purposes.

Thank you for your time and participation.

Good luck!

Instructions for incentive c) - group incentive:

1. This experiment is part of a study of how groups solve problems. Your task as a group is to work together to solve three different problems by sorting through a set of cards. After working on each of the three problems each of you will also complete a short questionnaire.
2. Before you begin, you will solve two training problems to help you understand the basic requirements of the task. Please feel free to ask any question during these training sessions. After the training sessions, your group will solve the three real problems.
3. For each problem, you will each be given four random cards to start with. The goal is to exchange cards until each group member gets a set of four of a kind.
4. You can speak to one another and describe what is on your cards. But you may not show each other your cards or look over the barrier at other peoples' cards.
5. You can trade cards as often as you like. But you may only trade one card at a time and no person can hold more than 5 cards or less than 3 cards at once.
6. You will earn up to 1 bonus mark in MSCI 311 for each problem, depending on your performance. ***In each task, if every person in your group gets four cards of a kind, you will get 1 bonus mark. However, if anyone in your group cannot get four cards of a kind, you will get nothing.***
7. Please spread your cards on your section of the table so that they can be seen by the video camera overhead. Do not overlap your cards or hold them in your hand.
8. You have a maximum of 15 minutes to solve each problem, although you probably won't need that much time.
9. These tasks comply with the UW ethics and research guidelines. All data collected will be used solely for research purposes.

Thank you for your time and participation.

Good luck!

Appendix D: QUESTIONNAIRE FOR EXPERIMENT

UW Questionnaire – Group Problem Solving

This questionnaire is designed to help us better understand your role as it relates to the group in the process of solving the problems. Please be as explicit as possible to enable us understand your perception of the group dynamics. Please be assured of the complete confidentiality of your comments.

Each section of this questionnaire is labeled with a Task number and will be filled after the completion of the corresponding task.

Please circle either yes or No where applicable.

Name: _____

Student ID: _____

Department: _____

Range of Age: Please choose from below:

- Below 17
- 17-20
- 20-25
- 25-30
- 30-35
- Above 35

Which letter is your team letter? (See the piece of paper on your section of the table; choose from A, B, C, and D)

Please do not write here. For experimenter use only. Task Information:

Task 1

1) Did your group identify 4 groups of objects? Yes / No.

If yes, what were the 4 groups? Please list them:

If no, did you individually get 4 of a kind? Which group did you form?

2) Rate the difficulty of this task on a scale of 1-7:

Very easy			Neutral			Very difficult
1	2	3	4	5	6	7

3) Can you give an example of something happened that made the problem difficult?

4) Can you give an example of something happened that helped lead you to the solution?

Task 2

1) Did your group identify 4 groups of objects? Yes / No.

If yes, what were the 4 groups? Please list them:

If no, did you individually get 4 of a kind? Which group did you form?

2) Rate the difficulty of this task on a scale of 1-7:

Very easy			Neutral			Very difficult
1	2	3	4	5	6	7

3) Can you give an example of something happened that made the problem difficult?

4) Can you give an example of something happened that helped lead you to the solution?

Task 3

1) Did your group identify 4 groups of objects? Yes / No.

If yes, what were the 4 groups? Please list them:

If no, did you individually get 4 of a kind? Which group did you form?

2) Rate the difficulty of this task on a scale of 1-7:

Very easy			Neutral			Very difficult
1	2	3	4	5	6	7

3) Can you give an example of something happened that made the problem difficult?

4) Can you give an example of something happened that helped lead you to the solution?

The remaining questions focus on how well you performed and worked together as a group.

1) Please rate your individual performance and your team's overall performance on a scale of 1-7:

Very poor performance			Satisfactory performance			Very good performance
1	2	3	4	5	6	7

2) Please rate each group member in terms of their helpfulness and cooperativeness on a scale of 1-7:

Group Member	Never helpful/ cooperative	Often not helpful/ cooperative	Sometimes not helpful/ cooperative	Neutral	Sometimes helpful/ cooperative	Often helpful/ cooperative	Always helpful/ cooperative
A	1	2	3	4	5	6	7
B	1	2	3	4	5	6	7
C	1	2	3	4	5	6	7
D	1	2	3	4	5	6	7

3) Please give a specific example of something one of the group members did that was helpful (e.g. Person X did…)

4) Please give a specific example of something one of the group members did that was not helpful (e.g. Person X did…)

5) Please rate the relative importance of your own individual performance vs. your group performance in these problem solving tasks:

The only thing that mattered was my individual performance			Individual and group performance were equally important			The only thing that mattered was my group performance
1	2	3	4	5	6	7

6) Do you have any comments about the three tasks? What made it difficult?

Reminder: Please be sure not to speak to others in your class about any details of this study (because it may bias the results)!

Appendix E: TIMES OF OVERLAPPING BETWEEN CATEGORIES IN PRE-EXPERIMENT

Times (number of cards) of overlapping between any non-restructuring cards belong to two categories

	Category	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1	Animal																							
2	Baking																							
3	Building																							
4	Clothes																							
5	Competition																							
6	Costume				3																			
7	Alphanumeric																							
8	Drink																							
9	Emergency																							
10	Furniture																							
11	Instruments																							
12	Medical				1																			
13	Plant										1													
14	Shape																							
15	Sports				4	2	2																	
16	Stretching					8																		
17	Technology																							
18	Transportation					1																		
19	Utensils								1															
20	Vase								1															
21	Vegetable																							
22	Weather				1								1											
23	Whisk		1																	4				

Times (number of cards) of overlapping between any restructuring card and any category

	Restructuring item	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1	Animal[a]																							
2	Baking'								1													2		
3	Building'																				1			
4	Clothes'						2				2		1			2	1		2				1	
5	Competition'									1						2	1							
6	Costume'				1								1											
7	Alphanumeric'																							
8	Drink'										1			1										
9	Emergency'																							
10	Furniture'				4																			
11	Instruments'										1													
12	Medical'									2														
13	Plant'																							
14	Shape'																							
15	Sports'																							
16	Stretching'					4				1														
17	Technology'									2														
18	Transportation'												1											
19	Utensils'																							
20	Vase'			1																				
21	Vegetable'		3[b]																					
22	Weather'				2																			
23	Whisk'																							

a: Animal' means the restructuring item in the category of "animal". Thus the second column has restructuring items from different categories.

b: Numbers mean the times the restructuring item from the category that correspond to the row be mistakenly put into the category correspond to the column. For example, 3 here mean the restructuring item from "vegetable" category was mistakenly put into category of "Baking" for three times in total across all 9 participants.

Appendix F: NON-PARAMETRIC ANALYSIS USING MANN-WHITNEY TEST

Effects of incentives on group performance and behaviours under structure 3- detour and restructuring

1.Mann-Whitney Test: individual versus mixed

	Time to solve (s)	No of Card Exchanges	No of detour categories the group collected	No of LIB trajectory reversals	LIB Rsquare				
					1st Order	2nd Order	3rd Order	4th Order	5th Order
Mann-Whitney U	70	41.5	55	49	78	79	34	29	52
Wilcoxon W	175	146.5	160	154	183	170	125	120	143
Z	-1.019	-2.421	-1.826	-2.106	-0.631	-0.582	-2.766	-3.009	-1.667
Asymp. Sig. (2-tailed)	0.308	0.015	0.068	0.035	0.528	0.560	0.006	0.003	0.096
Exact Sig. [2*(1-tailed Sig.)]	.325[a]	.014[a]	.085[a]	.043[a]	.550[a]	.583[a]	.005[a]	.002[a]	.101[a]

2.Mann-Whitney Test: individual versus group

	Time to solve (s)	No of Card Exchanges	No of detour categories the group collected	No of LIB trajectory reversals	LIB Rsquare				
					1st Order	2nd Order	3rd Order	4th Order	5th Order
Mann-Whitney U	79	22.5	36	39.5	90	84	30	37	39
Wilcoxon W	184	127.5	141	144.5	195	189	135	142	144
Z	-0.873	-3.486	-3.087	-2.733	-0.368	-0.643	-3.124	-2.803	-2.523
Asymp. Sig. (2-tailed)	0.383	0.000	0.002	0.006	0.713	0.520	0.002	0.005	0.012
Exact Sig. [2*(1-tailed Sig.)]	.401[a]	.000[a]	.004[a]	.006[a]	.734[a]	.541[a]	.001[a]	.004[a]	.011[a]

3.Mann-Whitney Test: mixed versus group

	Time to solve (s)	No of Card Exchanges	No of detour categories the group collected	No of LIB trajectory reversals	LIB Rsquare				
					1st Order	2nd Order	3rd Order	4th Order	5th Order
Mann-Whitney U	80	76	57	73	83	89	69	85	83
Wilcoxon W	185	167	148	164	188	194	174	176	188
Z	-0.534	-0.729	-1.738	-0.880	-0.388	-0.097	-1.068	-0.291	-0.388
Asymp. Sig. (2-tailed)	0.593	0.466	0.082	0.379	0.698	0.923	0.286	0.771	0.698
Exact Sig. [2*(1-tailed Sig.)]	.616[a]	.488[a]	.105[a]	.402[a]	.720[a]	.943[a]	.302[a]	.793[a]	.720[a]

a. Not corrected for ties.

Effects of problem structure on group performance and behaviours under individual incentive

a)

1.Mann-Whitney Test: basic versus detour

Test Statistics[b]								
				LIB Rsquare				
	Time to solve (s)	No of Card Exchanges	No of LIB trajectory reversals	1st Order	2nd Order	3rd Order	4th Order	5th Order
Mann-Whitney U	88	93	84	92	83	91	94	77
Wilcoxon W	193	198	189	197	188	196	199	155
Z	-.460	-.288	-1.096	-.276	-.689	-.322	-.184	-.054
Asymp. Sig. (2-tailed)	.646	.773	.273	.783	.491	.748	.854	.957
Exact Sig. [2*(1-tailed Sig.)]	.667[a]	.839[a]	.511[a]	.804[a]	.511[a]	.769[a]	.874[a]	.979[a]

a. Not corrected for ties.

b. Grouping Variable: Stimulus (1a,2a)

2.Mann-Whitney Test: detour versus detour + restructuring

Test Statistics[b]									
					LIB Rsquare				
	Time to solve (s)	No of Card Exchanges	No of detour categories the group collected	No of LIB trajectory reversals	1st Order	2nd Order	3rd Order	4th Order	5th Order
Mann-Whitney U	16	62.5	77	68	87	70	73	68	48
Wilcoxon W	121	167.5	182	173	192	175	178	173	139
Z	-3.769	-1.775	-1.192	-1.617	-.505	-1.287	-1.149	-1.378	-1.632
Asymp. Sig. (2-tailed)	.000	.076	.233	.106	.613	.198	.251	.168	.103
Exact Sig. [2*(1-tailed Sig.)]	.000[a]	.104[a]	.352[a]	.178[a]	.635[a]	.210[a]	.265[a]	.178[a]	.110[a]

a. Not corrected for ties.

b. Grouping Variable: Stimulus (2a,3a)

3.Mann-Whitney Test: basic versus detour + restructuring

Test Statistics[b]								
				LIB Rsquare				
	Time to solve (s)	No of Card Exchanges	No of LIB trajectory reversals	1st Order	2nd Order	3rd Order	4th Order	5th Order
Mann-Whitney U	10	53.5	54	90	55	73	71	63
Wilcoxon W	115	158.5	159	195	160	178	176	154
Z	-4.044	-2.226	-2.543	-.368	-1.976	-1.149	-1.241	-1.103
Asymp. Sig. (2-tailed)	.000	.026	.011	.713	.048	.251	.215	.270
Exact Sig. [2*(1-tailed Sig.)]	.000[a]	.039[a]	.044[a]	.734[a]	.050[a]	.265[a]	.227[a]	.287[a]

a. Not corrected for ties.

b. Grouping Variable: Stimulus (1a,3a)

Effects of problem structure on group performance and behaviours under mixed incentive b)

1.Mann-Whitney Test: basic versus detour

Test Statistics[b]								
	Time to solve (s)	No of Card Exchanges	No of LIB trajectory reversals	LIB Rsquare				
				1st Order	2nd Order	3rd Order	4th Order	5th Order
Mann-Whitney U	68	68	73	74	81	82	84	60
Wilcoxon W	159	159	164	165	172	173	175	138
Z	-.846	-.906	-.675	-.538	-.179	-.128	-.026	-.369
Asymp. Sig. (2-tailed)	.397	.365	.500	.590	.858	.898	.980	.712
Exact Sig. [2*(1-tailed Sig.)]	.418[a]	.418[a]	.579[a]	.614[a]	.880[a]	.920[a]	1.000[a]	.740[a]

a. Not corrected for ties.

b. Grouping Variable: Stimulus (1b,2b)

2.Mann-Whitney Test: detour versus detour + restructuring

Test Statistics[b]									
	Time to solve (s)	No of Card Exchanges	No of detour categories the group collected	No of LIB trajectory reversals	LIB Rsquare				
					1st Order	2nd Order	3rd Order	4th Order	5th Order
Mann-Whitney U	24	31.5	30.5	44	71	34	21	19	41
Wilcoxon W	115	122.5	121.5	135	162	125	112	110	132
Z	-3.103	-2.742	-2.920	-2.145	-.692	-2.590	-3.258	-3.359	-2.013
Asymp. Sig. (2-tailed)	.002	.006	.004	.032	.489	.010	.001	.001	.044
Exact Sig. [2*(1-tailed Sig.)]	.001[a]	.005[a]	.004[a]	.039[a]	.511[a]	.009[a]	.001[a]	.000[a]	.046[a]

a. Not corrected for ties.

b. Grouping Variable: Stimulus (2b,3b)

3.Mann-Whitney Test: basic versus detour + restructuring

Test Statistics[b]								
	Time to solve (s)	No of Card Exchanges	No of LIB trajectory reversals	LIB Rsquare				
				1st Order	2nd Order	3rd Order	4th Order	5th Order
Mann-Whitney U	1	21.5	38.5	82	35	29	39	35
Wilcoxon W	92	112.5	129.5	173	126	120	130	126
Z	-4.282	-3.302	-2.494	-.128	-2.538	-2.847	-2.333	-2.115
Asymp. Sig. (2-tailed)	.000	.001	.013	.898	.011	.004	.020	.034
Exact Sig. [2*(1-tailed Sig.)]	.000[a]	.001[a]	.016[a]	.920[a]	.010[a]	.003[a]	.019[a]	.035[a]

a. Not corrected for ties.

b. Grouping Variable: Stimulus (1b,3b)

Effects of problem structure on group performance and behaviours under group incentive c)

1.Mann-Whitney Test: basic versus detour

Test Statistics[b]								
	Time to solve (s)	No of Card Exchanges	No of LIB trajectory reversals	LIB Rsquare				
				1st Order	2nd Order	3rd Order	4th Order	5th Order
Mann-Whitney U	90	97.5	97.5	68	91	93	81	90.5
Wilcoxon W	195	202.5	202.5	173	196	198	186	195.5
Z	-.368	-.027	-.032	-1.378	-.322	-.230	-.781	-.345
Asymp. Sig. (2-tailed)	.713	.979	.974	.168	.748	.818	.435	.730
Exact Sig. [2*(1-tailed Sig.)]	.734[a]	.982[a]	.982[a]	.178[a]	.769[a]	.839[a]	.454[a]	.734[a]

a. Not corrected for ties.

b. Grouping Variable: Stimulus (1c,2c)

2.Mann-Whitney Test: detour versus detour + restructuring

Test Statistics[b]									
	Time to solve (s)	No of Card Exchanges	No of detour categories the group collected	No of LIB trajectory reversals	LIB Rsquare				
					1st Order	2nd Order	3rd Order	4th Order	5th Order
Mann-Whitney U	15	12	7	21	82	66.5	19	21	32.5
Wilcoxon W	120	117	112	126	187	171.5	124	126	137.5
Z	-3.814	-4.001	-4.439	-3.663	-.735	-1.448	-3.630	-3.538	-3.010
Asymp. Sig. (2-tailed)	.000	.000	.000	.000	.462	.148	.000	.000	.003
Exact Sig. [2*(1-tailed Sig.)]	.000[a]	.000[a]	.000[a]	.000[a]	.482[a]	.150[a]	.000[a]	.000[a]	.002[a]

a. Not corrected for ties.

b. Grouping Variable: Stimulus (2c,3c)

3.Mann-Whitney Test: basic versus detour + restructuring

Test Statistics[b]								
	Time to solve (s)	No of Card Exchanges	No of LIB trajectory reversals	LIB Rsquare				
				1st Order	2nd Order	3rd Order	4th Order	5th Order
Mann-Whitney U	5	5.5	20	78	71	20	18	25
Wilcoxon W	110	110.5	125	183	176	125	123	130
Z	-4.274	-4.363	-3.709	-.919	-1.241	-3.584	-3.676	-3.354
Asymp. Sig. (2-tailed)	.000	.000	.000	.358	.215	.000	.000	.001
Exact Sig. [2*(1-tailed Sig.)]	.000[a]	.000[a]	.000[a]	.376[a]	.227[a]	.000[a]	.000[a]	.000[a]

a. Not corrected for ties.

b. Grouping Variable: Stimulus (1c, 3c)