Collaborative Visual Search

A Dissertation Presented

By

Xin Chen

To

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

Doctor of Philosophy

In Experimental Psychology

Stony Brook University

December 2007

Stony Brook University

The Graduate School

Xin Chen

We, the dissertation committee for the above candidate for the
Doctor of Philosophy degree, hereby recommend
acceptance of this dissertation.

Gregory J. Zelinsky, Ph.D. – Dissertation Advisor
Associate Professor, Psychology Department

Susan E. Brennan, Ph.D. - Chairperson of Defense
Associate Professor, Psychology Department

Antonio L. Freitas, Ph.D.
Assistant Professor, Psychology Department

Dimitris Samaras, Ph.D.
Associate Professor, Computer Science Department

This dissertation is accepted by the Graduate School

Lawrence Martin
Dean of the Graduate School

Abstract of the Dissertation

Collaborative Visual Search

by

Xin Chen

Doctor of Philosophy

in

Experimental Psychology

Stony Brook University

2007

Previous work has demonstrated that people can collaborate during search by adopting a strategy of spatially dividing the search labor. But does this strategy hinge on the information available in the task? And can searchers coordinate their behavior in a fine-grained way, at the micro-level? I had groups of 2-4 people engage in 5 collaborative search tasks. In Experiment 1 subjects searched for an oval dot among circular dots. Experiment 2 added color, with red, blue, green, and black dots segregated into irregular regions. Experiment 3 used multiple potential targets (1, 4, or 8 photo-realistic objects) in 14-item displays. I quantified division-of-labor by correlating targets' properties (e.g., location, color, identity) with individuals' responses within the collaborating group. Consistent with previous work, subjects in Experiment 1 divided the search labor spatially, splitting the display in halves (2-person condition) or quadrants (4-person condition). Subjects in Experiment 2 divided the search labor by feature, with each of 4 members searching a different color. In Experiment 3 subjects divided the labor by targets rather space or feature, with each searcher taking responsibility for a different potential target from the preview set. Experiment 4 looked for tacit coordination during collaborative searches. Experiment 5 used a *shared-gaze* methodology to explore coordination at a micro-level; remotely located pairs viewed matching displays with each other's eyegaze cursor superimposed, so that each could monitor where the other was looking in real time. Stimuli were again colored dots, with only two colors and with one region 1/3 the size of the other. I found that when one person finished searching the smaller region, she often assisted her partner in searching the larger region. This assistance was targeted, such that she tended to look where her partner had not yet searched. I conclude that subjects use spatial division-of-labor collaborative strategies when the task does not allow for more meaningful divisions of labor. When the task allows, simple spatial division-of-labor strategies are replaced by collaborative strategies based on feature and target information. These analyses constitute the first demonstration of behavioral coordination during visual search at the micro-level.

Table of contents

List of figures .... **vi**
List of tables .... **viii**
Acknowledgement .... **ix**
Chapter 1: Introduction **1**
1.1. Overview 1
1.2. Group task typology 3
1.3. Group performance 8
1.4. Coordination 13
1.5. Verbal communication for language based tasks 18
1.6. Verbal communication for visual tasks 20
1.7. Visual communication for visual tasks 23
1.8. Group visual search 29
Chapter 2: Experiments **33**
2.1. Overview of Experiments 33
2.2. Experiment 1 35
Design and Participants 36
Stimuli and Apparatus 36
Procedure 37
Results and Discussion 38
2.3. Experiment 2 42
Design, Participants, and Procedure 42
Stimuli and Apparatus 43
Results and Discussion 43
2.4. Experiment 3 46
Design and Participants 46
Stimuli and Apparatus 47
Procedure 47
Results and Discussion 48
2.5. Experiment 4 54
Design and Participants 56
Stimuli and Apparatus 57
Procedure 57
Results and Discussion 58
2.6. Experiment 5 60
Design and Participants 61
Stimuli and Apparatus 62
Procedure 63
Results and Discussion 64
Chapter 3: Conclusion **68**
References **74**
Appendix: Transcripts of the communication groups about division of labor strategies **83**

List of figures

Figure 1: The shared gaze collaborative search system used in Brennan et al. (2007, in press) and Zelinsky et al. (2005). Searchers were seated in separate rooms but saw the same stimulus and performed the same search task. The eye movements from each searcher's eyetracker could be displayed as a gaze cursor on the other's monitor.

Figure 2: A sample of the search stimuli used in experiment 1.

Figure 3: The physical arrangement of the collaborative search system used in Experiments 1-3. One source of video signal was split into 4 monitors. And the 4 gamepads were connected to the same computer also. Participants were either seated back-to-back or separated by a barrier. This system enable up to 4 persons to perform search task collaboratively.

Figure 4: A: The locations the targets found by the members of the three 2-person groups who split the search scene by bottom-top. The white bars represent the members who were assigned to the bottom side of the scene, and the black bars represent the members who were assigned to the top side of the scene. B: The locations of the targets found by the members of the two 2-person groups who split the search scene by right-left. The white bars represent the members who were assigned to the left side of the scene, and the black bars represent the members who were assigned to the right side of the scene.

Figure 5: The black dot dash lines represent the RTs of experiment 1: the black dots search experiment. And the solid lines represent the RTs of experiment 2: the color dots search experiment. The error bars represent one standard error.

Figure 6: The locations of the targets found by the members of the three 4-person groups who split the search scene by quadrant. The different type bars represent the members who were assigned to different quadrant.

Figure 7: A sample of the search stimuli used in experiment 2.

Figure 8: The color of the targets found by the members of the six 4-person groups who divided the search display by dot color. The different type bars represent the members who were assigned to different color.

Figure 9: The search stimuli used in experiment 3. The target preview had 1, 4 or 8 objects. For the 4 target preview, the top 2 objects from left to right were labeled as target 1 and target 2; the bottom 2 objects from left to right were labeled as target 3 and target 4. For the 8 target preview, the top 4 objects from left to right were labeled as target 1, target 2, target 3 and target 4; the bottom 4 objects from left to right were labeled as target 5, target 6, target 7 and target 8. The duration of the target preview followed a "1 second per

object" rule. So a 1-object preview was displayed for 1 second, and an 8-object preview 8 seconds. The search display had constant 14 objects.

Figure 10: A: The relation between the target preview location and the target found by the five 2-person groups who divided the 4 targets. The white bars represent the members of the 2-person groups who were assigned to target 1 and target 2. The black bars represent the members of the 2-person groups who were assigned to target 3 and target 4. B: The relation between the target preview location and the target found by the five 4-person groups who divided the 4 targets. The different bars represent the members of the 4-person groups who were assigned to target 1, target 2, target 3 and target 4 respectively. The error bars represent one standard error.

Figure 11: A: The relation between the target preview location and the six 2-person groups who divided the 8 targets. The white bars represent the members of the 2-person groups who were assigned to target 1, 2, 3, 4. The black bars represent the members of the 2-person groups who were assigned to target 5, 6, 7, 8. B: The relation between the target preview location and the six 4-person groups who divided the 8 targets. The different bars represent the different two targets that were assigned to different member.

Figure 12: A: The reaction time of the target-present trials in experiment 3. B: The reaction time of the target-absent trials in experiment 3. C: The error rates of the target-present trials in experiment 3. D: The error rates of the target-absent trials in experiment 3. The error bars represent one standard error.

Figure 13: A: The mean sensitivity d' of the groups in experiment 3. The left is the 4-target trials data and the right is the 8-target trials data. The error bars represent one standard error. B: The mean response bias lnβ of the groups in experiment 3. The left is the 4-target trials data and the right is the 8-target trials data. The error bars represent one standard error.

Figure 14: A sample of the search stimuli used in experiment 5.

Figure 15: The scatterplots of the fixations of the dyads in experiment 5. The red triangles represent the fixations of the participants who were assigned to smaller dots group, and blue circles represent the fixations of the participants who were assigned to bigger dots group.

Figure 16: The average gaze distance of the dyadic searchers in experiment 5. The initial 2 groups of bars represent the gaze distance averaged over the whole experiment. The third group of bar represent the gaze distance averaged over from 10 seconds later of the start to the end of each trial of the experiment.

List of tables

Table 1: Average hits, false alarms, sensitivity (d'), and response bias (lnβ) for the individual searchers, non-communication groups and communication groups.

Table 2: The mean and SE of error rates and RTs in experiment 5. The RTs are in msecs.

Acknowledgement

There are lots of people I would like to thank for their support in the shaping of this dissertation. First and foremost, I would like to thank my advisor, Greg Zelinsky. He is the best advisor I can imagine. Without his support, inspiration and knowledge I would never have finished.

I would also like to thank the members of my committee, Susan Brennan, Antonio Freitas and Dimitris Samaras, for their time, suggestions and comments. I also want to express my gratitude to Hoi-Chung Leung and Arthur Samuel, for their kindness and help.

I want to thank my present and past labmates, Christopher Dickinson, Mark Neider, Hyejin Yang and Joe Schmidt, for sharing their knowledge, and all kinds of help to my various projects. I also want to thank our RAs, especially Shreena Bindra, Zainab Karimjee, Mariya Norenberg and Samantha Brotzen, for their help on data collecting and analysis.

Chapter 1: Introduction

1.1. Overview

This thesis investigated collaborative visual search. Visual search is a task requiring a person to make a judgment about a target, often whether it is present or absent when the target's location is not known in advance. Typically, many non-target items are also present, thereby requiring the searcher to scan the visual environment looking for the target. Collaborative visual search is a group task in which two or more than two searchers scan the visual environment for particular targets, with the emphases on group performance instead of individual performance.

Solitary visual search has been well studied in both the theoretical and applied domains (for a review, see Wolfe, 1998). According to one popular theoretical perspective (e.g., Treisman & Gelade, 1980; Wolfe, 1994), visual search has two stages. The first stage is a guidance process. During this stage observers will determine the most significant position in the display being viewed.

The saliency of the position is determined by two processes: one computes how this position is different from its surroundings along several feature dimensions. This process is called the bottom-up guidance (e.g., Koch & Ullman, 1985). Another process computes the similarity between this position and the target. This is called top-down guidance (e.g., Rao, Zelinsky, Hayhoe, & Ballard, 2002). At the end of the first stage or the beginning of the second stage, observers' attention will be guided to the most significant position of the scene. Observers will then compare the target definition with the object at the most significant position to make the search decision: whether it is the target or not. If it is, the target is detected and the visual search task ends; if it is not, observers will move to the next salient position of the scene till they find the target or conclude that there is no target in the scene.

I will assume the existence of the above described solitary search processes, and focus on the additional processes arising from the group behavior inherent to collaborative search. But what are these additional processes? Take a squad in a battle field for example. All squad members are required to be watching out for hostile targets. However the squad may spatially divide the search labor "you search the front, I will cover the back." The squad members may also conflict over whether an object is hostile or not, or they may search less or more carefully by knowing that they are working as a group. All these phenomena, dividing search labor, resolving group conflict, and group motivation, can not be simply generalized from the findings of solitary search, but can only be studied by using collaborative visual search tasks.

This thesis has three chapters. In chapter 1 I reviewed the previous studies relevant to collaborative visual search, to see what I can learn from the literatures about collaborative searches' task properties (section1.2), performance (section1.3), coordination (section1.4), effects of communication media (section 1.5-1.7). And finally, I also reviewed the previous modeling and empirical studies on collaborative search (section1.8).

Chapter 2 included five experiments designed to investigate group search behavior. Experiment 1 -3 focused on the relationships between the search task properties and the emerging of coordination strategy. Experiment 4 focused on tacit coordination and speed/accuracy trade-offs. And experiment 5 focused on the micro-level behavioral coordination.

Finally, in chapter 3 I gave the conclusions of this thesis.

1.2. Group task typology

It is impossible to meaningfully discuss either individual or group cognitive processes without first discussing the specific task. Group tasks can be classified along a number of dimensions.

1. Group tasks can be classified by the behavior requirements on the group members. Within this dimension, Carter, Haythorn and Howell (1950) classified group tasks into six categories: clerical, discussion, intellectual construction, mechanical assembly, motor coordination and reasoning. By applying a factor analysis to these categories, Hackman and Morris (1975) proposed that there were three types of tasks: production -meaning tasks that need the group to generate idea, discussion, and problem-solving tasks -meaning tasks that require coordinated planning.
2. Different group tasks have different requirements on the relationships among the group members. With this dimension, Laughlin (1980) classified tasks into two types: done by collaborative groups, or done by competitive/mixed-motive groups. Collaborative group means that all the group members have compatible shared goals. Competitive group means that the group members have conflicting goal; the success of one member often means the failure of other members, as when individuals compete for a reward or limited resource (Laughlin, 1980). A search

task has not to be a collaborative task. Indeed group members can be instructed to search either collaboratively or competitively. This thesis investigated collaborative visual search. Participants were instructed to collaborate, and collaboration was emphasized by providing feedback on group, rather than individual performance.

3. Group tasks can be categorized by how the individual contributions are combined into the group product. The group visual search task has a single group outcome or product, and this kind of group task is called a unitary task (Steiner, 1972). Unitary tasks can be further divided into disjunctive, conjunctive, and additive varieties. In a disjunctive task, if any one of the group members succeeds, the entire group succeeds. Group performance is therefore determined by its most capable member. In a conjunctive task, all members of the group must succeed in order for the group to succeed. Group performance is therefore determined by its least capable member. Lastly, if the group product is the sum of the contributions from every group member, this is referred to as an additive task. Group performance in an additive task is often determined by averaging the performance of all of the members. Rope pulling is an example of an additive group task.
4. Group tasks also differ in their requirements on reaching consensus. Every group member has a preference for the right answer. Social decision schemes theory (Davis, 1973, 1982) provided a framework for understanding the process of combining these preferences into one group outcome. Some tasks have a demonstrably correct answer Laughlin (1980). An extreme example of this is the Eureka type task: the correct answer is so compelling that once a group member

points it out the group immediately accepts it as the right answer. So for a Eureka type task, the decision scheme should be "truth wins": the correct answer will win out (Laughlin & Ellis, 1986). On the other hand, some group tasks have no correct answer thereby requiring the group to reach a consensus. For a consensus-reaching task, a possible scheme is "majority wins": the answer preferred by the majority of the group will win out.

5. There are other dimensions that can be used to categorize group tasks, two of which were discussed by Shaw (1981). One is tasks' intrinsic interest: to what degree a task is interesting, motivating or attractive to the group members. Tasks can vary from dull to interesting on this dimension. Another is tasks' population familiarity: how much the population has had experience with a task. Tasks can vary from highly familiar (encountered by every one in the population, such as visual search) to unfamiliar (encountered by no one in the population, such as space shuttle piloting).

In an attempt to integrate these group classification efforts, McGrath (1984) proposed a group task classification schema called Task Circumplex. Based on what the group is to do (i.e., the behavior requirements on the group; Hackman & Morris, 1975), McGrath (1984) proposed that there are four group processes: to generate; to choose; to negotiate; and to execute. Moreover, each of these four group processes can be further divided into two subtypes:

1. To generate, can be divided into generating a plan, similar to Hackman and Morris (1975)'s problem-solving task, and generating ideas, similar to Hackman and

Morris (1975)'s production task. Examples of tasks in this category include brainstorming, planning and agenda setting.

2. To choose, can be divided into intellective tasks and decision-making tasks. Intellective tasks require choosing a correct answer; decision-making tasks require reaching consensus on a preferred solution. In this sub-division we can see a similarity to Laughlin (1980)'s classification schema.
3. To negotiate, is divided into two subtypes: Resolving conflicting viewpoints and resolving conflicting interests.
4. To execute, requires physical movement or coordination. This execution process can be divided into contest-oriented and performance-oriented tasks. Contest tasks require the group to compete with an opponent. Performance tasks require the group to meet some standard of excellence.

So what can we learn from these group task typologies to guide our investigation on collaborative visual search?

1. Based on McGrath (1984)'s circumplex framework, group visual search requires both a “choosing" (quadrant 2) and “executing” (quadrant 4) function. A “choosing” function is required in that the subjects' task is to find a correct answer, target present or absent. An “execution” function is less apparent, but not less important. Search is a process that typically unfolds over time, thereby requiring the execution of multiple shifts in attention or gaze.
2. Group visual search does not require cooperation. Each member of the group has the knowledge and skills needed to complete the task, assuming that they each know what target they are searching for. Also, and unlike an assembly line task,

the visual search task cannot be divided into different procedures, such as guidance, comparing the target with distractor, decision making, etc., and assigned to different participants. All the participants have to perform the visual search as an integrated process. However, this does not mean that the search labor cannot be divided. Given that one searcher cannot detect the target in a single glance; the group can choose to divide the display, with each searcher inspecting a different region.

3. Assuming that the search target is present and well defined, visual search is a Eureka type task. The group decision process should be “truth wins”, with little or no conflict among the members. However, target absent visual search will usually require a consensus, as members must decide whether to end the search and declare that the target is absent.
4. Group visual search is a disjunctive task (Steiner, 1972). The fastest response from any group member will end the search. Group performance is therefore determined by the fastest searcher’s performance.

Although helpful, a note of caution should be used when applying McGrath (1984)’s circumplex framework to visual search tasks. The framework was developed for the small group research community to account for group performance and processes in tasks having substantial problem solving, decision making, and language use components. Given that visual search relies minimally on any of these components, it is not clear whether this typology is valid for a search task. Moreover, the schemas themselves were never systematically validated through empirical testing (for an exception, see Straus, 1999). Rather, the schemas were developed based on the post-hoc analysis of empirical data.

1.3. Group performance

Groups almost always outperform individuals. The over performances can come from (not an exhaustive list): 1. Adding effect, for additive tasks, such as rope-pulling, the group output is the aggregation of the individual outputs. 2. Pooling effect, the more members in a group, the more chance the group includes at least one member who is capable of doing the task (Steiner, 1972). 3. Individuals working in a group outperform individuals working alone.

Researchers have shown great interest in whether individuals working in a group outperform individuals working alone. And this question can be boiled down to whether group interaction facilitates or inhibits individual performance.

Unfortunately for some group tasks it is almost impossible to measure individuals' performance when they work in a group. Take group visual search as an example, the search will end when any member of the group finds the target. So the fastest searchers' responses will be counted as the group's response, but the slow searchers would never have a chance to contribute to the group performance. In such a case all we have is the group's actual performance, but no measures of individuals' performance.

Steiner (1972) formalized the relations between the actual group performance (AP), potential group performance (PP), performance gain (PG) and performance loss (PL) as follows:

$$AP = PP + PG - PL \tag{1}$$

Here the potential group performance represents the expected performance if the group interaction neither facilitates nor inhibits the group performance. The performance gain term corresponds to a benefit from group interactions, and the performance loss term corresponds to the cost resulting from group interactions. According to this framework, it is impossible to independently estimate performance gains and losses. But if $PG - PL$ is positive, there is a performance gain, otherwise there is a loss. Since $PG - PL = AP - PP$, if actual group performance were more than the potential group performance, we can say group interactions facilitate the group performance, otherwise group interactions inhibit the group performance.

Now the question is how to determine the potential performance (Hill, 1982; Steiner, 1972). The potential performance should take into account the adding effect and pooling effect, so its difference from actual performance can only be attributed to group interaction effect.

Creating nominal groups (Taylor & Faust, 1952; Lorge & Solomon, 1955) is a solution. Taking group search again as an example, a nominal group would be created by pooling the responses from N individual searchers, with N being the number of searchers in the interacting group, and using the fastest of each trial as the nominal group's response. The performance of the nominal group therefore captures the adding and pooling effects across group members, but eliminates the influence of group interaction. In this sense it represents the performance expected if the group interaction neither facilitated nor inhibited performance in the task.

Then how did group interaction affect the group performance? I reviewed the group performance of three kinds of group tasks. The first is brainstorming. Brainstorm is a method of idea generation first proposed by Osborn (1957). This method has four simple rules: 1. banning self-criticism and criticism from others; 2. encouraging quantity of ideas instead of quality; 3. speaking out freely all ideas in mind; 4. building on the ideas of others. Osborn (1957) believed that by following these rules "…the average person can think up twice as many ideas when working with a group than when working alone…" This hypothesis was based on the assumption that group interaction would mutually stimulate individuals and cause them to generate more ideas than they would alone.

Taylor, Berry and Block (1958) was the first to empirically test this hypothesis. They compared the performance of an interacting group with that of a nominal group, which was created by pooling the ideas generated by the same number of individuals working individually. They found that the nominal group outperformed the interacting group in both quality and quantity. Rather than facilitating idea generation, group interaction actually inhibited idea generation, a finding that has since been multiply replicated (Diehl and Stroebe, 1987), but with the exception of dyadic groups.

Also, Valacich, Dennis and Connolly (1994) pointed out the performance difference between nominal and interacting groups is moderated by group size. Larger interacting groups usually generate no more ideas than smaller ones (Bouchard, Drauden & Barsalou, 1974; Bouchard & Hare, 1970; Fern, 1982; Hackman & Vidmar, 1970; Lewis, Sadosky & Connolly, 1975). However, for nominal groups the number of ideas generated increases with the size of the group (Hogarth, 1978), meaning that the nominal groups' superiority over interacting groups also tends to increase with group size (Bouchard & Hare, 1970; Lewis et al., 1975).

Another extensively studied group task is group remembering. The typical group remembering paradigm involves subjects first being instructed to memorize some materials individually, such as a list of words or a story, then being asked to work face-to-face as a group to remember as much of the materials as possible. Technically speaking, it is therefore only the memory retrieval process that is under the influence of group interaction. Researchers originally expected to find a group benefit arising from cross-cuing (Meudell, Hitch & Boyle, 1995): by cuing each other, interacting group members might be able to remember more material than they would if they were remembering individually.

However, and contrary to this expectation, the results showed that interacting groups recalled less information than nominal groups (pooling together the information retrieved by the same number of individuals working individually, or as a group in the absence of communication). This unexpected result was termed "collective inhibition" by Weldon and Bellinger (1997) and has since been repeatedly replicated (Pavitt, 2003).

This finding is similar to that obtained from brainstorming tasks. As in brainstorming tasks, dyads again proved to be a special case. Meudell et al. (1995) investigated dyads and found no difference between nominal and collaborative groups; other group size conditions showed that the nominal groups robustly outperformed interacting groups.

Different from brainstorming and group remembering, rope pulling is an additive task in which every contribution from every group member is counted. In the early 1920s, a German psychologist by the name of Ringelmann (Kravitz & Martin,1986) investigated the group performance of rope-pulling and found that as the group size increased, the total force exerted by the group also increased, but the average force exerted by each group member decreased. This has since become known as the Ringelmann effect. The Ringelmann effect illustrates a case in which there is lower than expected group performance, despite the additive sum of individuals' outputs.

In general, with the exception of dyadic groups, nominal groups typically outperform real interacting groups in both quantity and quality, a finding that prompted McGrath (1984) to conclude: "...The difference is large, robust and general..." So most research suggested that group interaction hurts, rather than helps, the group performance: "a group will perform, at best, only as well as the sum of its parts and, at worst, the interaction of group members will create loss rather than gain" (Propp,2003).

What underlies the performance decrement associated with group interaction? Most researchers attributed this inhibition to the cost of the communication between the members. For example, Diehl and Stroebe (1987, 1991) attributed this inhibition in a brainstorming task to the turn taking property of natural language. Since only one group member was allowed to talk at one time, other members had to wait and listen while another member was speaking. During this waiting-and-listening period, members might either forget their ideas, or engage in rehearsal in an attempt to remember them. Either way, the waiting members are blocked from producing new ideas, making them more subject to self-censoring. They referred to this effect as "production blocking". According to this explanation the inhibition should disappear, or at least decrease if the group members can communicate in parallel. Valacich et al. (1994) confirmed this by allowing group members to communicate using a computer-conferencing system; members could type in their ideas in parallel and all ideas were showed on a screen to all members.

Valacich et al. (1994) found that with this communication system the nominal groups no longer outperformed the interactive groups, even with the slower pace of text-typing than speaking.

In the case of a group remembering task, Basden, Basden, Bryner, and Thomas (1997) attributed the group inhibition to communication interrupting member retrieval strategy. The retrieval-strategy disruption was thought to share the same process underlying part-list cuing inhibition. Part-list cuing inhibition is when people try to remember a list of words, telling them some of the words from the list will tend to impede their memory for the other words on the list (Nickerson, 1984). For example, when remembering a list of categorized words, people prefer to retrieve all the words from one category before moving to another category. When working in a group, group members would likely retrieve words from different categories at any moment, so the group would not benefit from organized retrieval, resulted in poorer memory performance compared to a nominal group.

For a group rope-pulling task, Steiner (1972) argued that the group members do not synchronize their efforts optimally; not every member directs their efforts in the same direction at the same time.

1.4. Coordination

I want to mention here that not all groups are teams. One of the elementary features of teams is the interdependence among the team members, the team work usually can only be done by teams but not individuals. Military combat teams, airplane cockpit teams or the baseball teams are all examples of teams. I prefer the term group search over team search because the searchers do not necessarily need each other, the individuals can search for the target on their own. But of course, the group will benefit from coordination.

Dividing search labor can not be done without coordination among the group members. Malone and Crowston (1994) defined coordination as “managing dependencies between activities”. In this thesis the group coordination referred to the temporal and spatial synchronization of group behavior to achieve optimal performance.

Strategy coordination Tschan and Cranach (1996) proposed that there are at least three hierarchical levels of coordination. At the highest level is knowledge based coordination. Taking group search as an example, group members share a common task goal and common knowledge about what is the general way of doing visual search. This is the fundamental of all kinds of coordination.

At the intermediate level are the rules, scripts, or strategies based coordination. A strategy can be regarded as a long term plan of action and is different from an immediate reaction to a situation. I can not say that group members will never change their strategy, but it is probably fair to say that members will often stick to a particular strategy.

Wittenbaum, Vaughan, and Stasser (1998) classified group strategy coordination into four categories on the basis of two dimensions: time and explicitness. Time refers to when the coordination emerges, which could be either pre-process or inprocess. Explicitness refers to whether the coordination is explicitly expressed among the group members or tacit. The four types of coordinations are therefore:

1. Preplans: this is explicit coordination and happens before the tasks begin. Job descriptions, rules, policies, schedules and standard procedures all fall in this category. Preplans often take the form of written memos or verbal instructions.

2. Inprocess plans: This is also a form of explicit coordination, but it happens while the task is being performed. When group members discuss strategies for performing a task while they are engaged in the task, they are demonstrating inprocess planning. Straus (1999) referred to this kind of coordination as “propose process”.
3. Tacit precoordination: This form of coordination happens before the group starts interacting. The main difference between preplan and tacit precoordination is that in the tacit precoordination group members do not talk about how they plan to coordinate their behaviors; they simply coordinate their behaviors based on unspoken assumptions about the other members.
4. Inprocess tacit coordination: This is tacit coordination that takes place during group interaction. Group members often tacitly adjust their behavior to fit with the observed behavior of others while they are performing a task. There is only a thin line between this kind of coordination and the micro-level behavior coordination, the topic of the following section.

<u>Tacit coordination</u> Evidences for tacit precoordination have come from studies of the Transactive Memory System (Wegner, 1986). Group members can share their memory through communication, such that group members can not only access their own memory, but also the memories of other members in the group. Group members therefore have an internal index to their own memory, as well as external indices to the other group members. Researchers manipulated the relationships between the collaborating members: the groups were composed of either intimate couples or strangers. All the groups were instructed to remember lists of words. The results showed that intimate couples recalled

more than strangers. It was explained that intimate couples knew each other's areas of expertise, and tacitly chose words that they were likely to remember and avoided words that their partner was likely to remember. When an explicit organizational scheme for remembering the words was imposed on groups, intimate couples recalled fewer words than strangers. This suggested that the explicit scheme impeded the tacit precoordination of intimate couples, but helped strangers who did not have such a system in the first place (Wegner, 1991; Hollingshead, 1998b, 1998a).

The work of Wegner (1991) showed that group members use unspoken assumptions to guide their selection of a collaborative strategy. Wittenbaum, Merry, and Stasser (1996) extended this idea to show that unspoken assumptions can be affected by task demands. Wittenbaum et al. (1996) used identical set of materials, but placed different demands on the groups. Some groups were instructed to collectively recall the materials and others were instructed to reach consensus. The hypothesis was that for collective recall, group members should try to remember information that they thought others would not likely remember. However, when reaching consensus, group members should try to remember information that they thought others would likely remember. This hypothesis was confirmed by the data.

Tacit coordinations are rarely optimal because the assumptions are unproven and could be wrong. The better way for group members to decide on a strategy is to discuss various possibilities and reach a consensus. For example, when groups were forced to plan before acting, group performance increased with increasing of amount of communications about coordination (Weingart, 1992; Weldon, Jehn, & Pradhan, 1991). But interestingly, even with the environment allowing, groups do not like to explicitly talk about their coordination strategy, "...there appears to be a pervasive norm in groups NOT to address such matters explicitly..." (Hackman & Morris, 1975). Even when a group knew that it was to their advantage to engage in planning before actual work, and it was easy to engage in such planning, group members still tended to begin immediately generating "real" products when they were presented with a task, rather than talking with each other about how to do the task. It is therefore very possible that group members were using a wrong assumption to adjust their behavior from the beginning to the end.

<u>Micro-level behavioral coordination</u> At the micro-level is the behavior or skill based moment-by-moment coordination, and it is finer-grained both temporally and spatially. It is usually studied in the context of group physical activities. Because group visual search has a substantial motor component (i.e., scanning eye movements), I believe that the group search task might benefit from micro-level coordination.

Occurrence of micro-level behavioral coordination requires real time information of the environment and group members' behavior. Salas, Bowers, and Cannon-Bowers (1995), Steven and Campion (1994) dissociated coordinate behaviors (i.e., the behaviors for coordination purpose) from task behaviors (i.e., the behaviors for the task) and pointed out that the primary requirements of successful behavioral coordination were (1) the group members' ability to maintain constant mutual awareness of each others' behavior, and (2) the availability of inter-group feedback. Also, by observing group collaborative behavior, Tang (1991) found that the group members' proximity was important for micro-level behavioral coordination: "... This proximity allows a peripheral awareness of the other participants and their actions. Many intricate and coordinated hand motions were observed, such as avoiding collisions with other hands or working closely together on a sketch. These coordination actions demonstrate an awareness of the other participants, enabled by being in close proximity with them...” This is why micro-level behavioral coordination is mainly observed in groups working face-to-face.

Verbal communication did not play an important role in micro-level behavioral coordination. Groups demonstrating micro-level behavioral coordination simply did so spontaneously or usually over practice, but not by discussing verbally.

1.5. Verbal communication for language based tasks

So what is known about the relationship between communication and group performance? Unfortunately there were no clear answers yet. A consistent finding in the literature is that the amount of verbal communication was positively correlated with group performance (Foushee, 1984; Kanki & Foushee, 1989; Williges, Johnston, & Briggs, 1966). However, this finding varies with task complexity; it is true for complex tasks but not for simple tasks (Lanzetta & Roby, 1960; Muller, 1992).

Media-richness theory is the most cited theory attempting to sort out the relationships among task characteristics, communication media (only verbal communication media here), and group performance. The first assumption of media-richness theory is that groups need information to reduce uncertainty and equivocality (Daft, Lengel, &Trevino, 1987). Galbraith (1977) defined uncertainty as “the difference between the amount of information required to perform the task and the amount of information already possessed by the organization.” Equivocality is defined as the ambiguity of the task, caused by conflicting interpretations about the situation. Therefore, when equivocality is high, an individual does not even know what questions to ask. When uncertainty is high the group at least knows the question but lacks the necessary information to address it.

The second assumption of media-richness theory is that communication media differ in the richness of the information processed. The criteria of richness include feedback capability, the communication channels utilized, language variety, and personal focus. The more a medium has these characteristics, the richer it is. Face-to-face communication is typically considered to be the richest medium, because it allows rapid mutual feedback, permits the simultaneous communication of multiple cues (e.g., body language, facial expression, tone of voice), uses high-variety natural language, and conveys emotion. The telephone, addressed written documents (e.g., notes, memos, letters), and unaddressed documents (e.g., bulletins, standard reports) follow face-to-face communication in media richness, in a descending order.

The third and most important assumption of media-richness theory is that rich media are appropriate for reaching agreement about unanalyzable, difficult, and complex issues, while lean media are appropriate for communicating about routine activities. In other words, the rich media are good at reducing equivocality, and lean media are good at reducing uncertainty. So if the media is too rich for the task, it will cause distractions. On the other hand, if the media is too lean for the task, it will be incapable of transmitting enough information (Daft et al., 1987; McGrath & Hollingshead, 1993; Suh, 1999).

Media richness theory has not been validated in the context of visual tasks. I believed that verbal communication is necessary to generate explicit strategies for group visual search. But after that, the verbal communication is less useful but more a source of distractions. After all, group members can not depend on verbal communication for micro-level behavioral coordination. I will explore this topic in the following section.

1.6. Verbal communication for visual tasks

For micro-level coordination of group visual search, the real time awareness of the group members' search progress is necessary. I will explain how it is difficult, if not impossible, to achieve micro-level coordination by verbal communication. I will use spatial referring as an example.

How does a person direct another person's spatial attention to a particular location? There are several possibilities. For example, if the location corresponds to a highly salient object (Treisman & Gelade, 1980; Koch & Ullman, 1985), one might only need to say "there" or "look at that" to communicate this location to the other person. Of course this requires that the salient object can be effortlessly distinguished from its surroundings, and that there is only one salient object in view (Logan, 1995). Alternatively, if the to-be-referred object has a name or some perceptual features that can uniquely distinguish it from all other objects in view, a person might simply use these labels when referring to the object; "look at the coffee cup" or "look at the red vertical bar". This referential strategy shifts responsibility to the other person to search the view and find the referred object, but search is often fast and efficient and may not impose a prohibitive cost.

A far more difficult problem is how to refer to an object that is unremarkable in its surroundings. For example, how does one refer to a window on the front of a building with lots of windows? Under these circumstances, there are no salient cues to serve as an implicit anchor or a name that unambiguously references the desired object. Given that humans are not computers and cannot reference objects in terms of Euclidean x, y coordinates, a method of using natural language to reference objects in space must be used. This typically involves specifying a target in terms of a deictic relationship relative to a landmark.

The first step in using a landmark referential communication scheme is to choose a landmark that is accepted by all members of the group. The most salient or distinguishable object in the view is typically a good candidate for a landmark. Talmy (1983) believed that the objects chosen for landmarks should be: permanently located, relatively large, geometrically complex, encountered earlier in the scene, and easily or immediately perceivable. A successful landmark should serve as a spatial anchor for the group's collective attention.

Having decided on a landmark, the next step is to ground a common "reference frame" with the landmark as the origin point. A reference frame is a coordinate system, which includes an origin point, orientation, direction, and scale. There are three kinds of reference frames in English, and each of them has its own semantic organization (for details, see Levinson, 1996). The reference frame will be aligned with the landmark as the origin point, and then the direction and orientation of the reference frame will be determined by viewers projecting them onto the landmark. For example, "the window is to the left of the door" would be a typical landmark-based reference because the window would be to the viewer's left if the viewer were in the scene and at the position of the door.

After both the landmark and the reference frame are established, the director must settle on a deictic relationship between the target and landmark. Logan and Sadler (1996) proposed that each deictic relation has associated with it a “spatial template”, which represents the regions surrounding the target. In order to obtain the desired deictic relation, the director must center the spatial template on the landmark and then adjust the template relative to the reference frame until the best match to the target is found.

Logan and Sadler (1996) suggested that the template mapping process occurs in parallel, and that different relations have different spatial templates. Also templates can be combined to represent compound relations, such as “up and to the left”. In English there are over 80 spatial relations lexicalized in a closed class: the spatial prepositions. Of course there are other linguistic representations of spatial location, but the spatial prepositions are devoted specifically to representing spatial locations.

Let us suppose the end product of the above described processes is the referential communication “the window is to the left of the door”. There are two arguments in this sentence, the first argument represents the “located object” (the window), and the second represents the landmark (the door). The position of the target is specified by its deictic relation with the landmark: to the left of.

According to Logan (1995), a person’s attention could then be directed from the landmark to the target object based on the specified deictic relation. However, this is not a very precise or efficient method of referring to an object in space; the target region may be large and coarsely defined, and the process will be slowed if there is more than one potential target in the referenced region of space. For example, the reference “A above B” will lack precision to the extent that, when there are other objects between A and B, a person’s attention will first move to the non-target objects closest to the landmark, with the “spotlight” of attention eventually reaching the target after all of the intermediate objects have been evaluated and rejected (Logan & Compton, 1996; Carlson & Logan, 2001).

In summary, the visual routine (Ullman, 1984) to reference an object in space would require: landmark indexing, reference frame alignment and adjustment, spatial template alignment, and computing a goodness of fit.

1.7. Visual communication for visual tasks

It is awkward to transfer visual information through verbal media. So what forms of visual information are useful for coordination of group visual search?

Mutual gaze, also known as eye contact, refers to the state in which people can tell whether they are looking at each other or not (Gale & Monk, 2000). Mutual gaze has been shown to play an important role in regulating turn-taking in conversations: both speaker and addressee can use eye gaze to signal that they understand each other, which is important for moving the conversation forward (Brennan, 2005; Clark & Brennan, 1991).

Mutual eye gaze and the perception of facial expression were also originally thought to be important for remote collaboration, as they might create an environment of telepresence. However, by comparing video-mediated with audio-mediated communication, researchers found that mutual gaze and facial expression alone, and in the absence of an immediate environment, failed to help referential conversation (Chapanis, Ochsman, Parrish, & Weeks, 1972). This basic finding was mirrored in the real world by the well-known failure of Picturephone (Egido, 1989).

The importance of the immediate environment, and the relative unimportance of mutual gaze, was further confirmed by studies on physical collaborative activities, which showed that the group’s shared work space was far more important than the view of the participants’ face or mutual gaze (Gergle, Kraut, & Fussell, 2004). For instance, Argyle and Graham (1977) asked a pair of subjects to plan a European trip, and found that when there was a map of Europe in between them, the amount of mutual gaze dropped from 77% to 6.4%, and 82% of the time was spent looking at the map. This suggests that the participants were communicating by looking at or pointing to the same object, instead of looking at each other.

Why is the immediate environment so much more important than mutual gaze and facial expression for group performance? Mutual gaze serves mainly to communicate understanding and expressions between participants; the immediate environment provides real time information about the ongoing progress in a task. In this sense, the immediate environment can serve as a hub that every group member can synchronize their behavior to. In a very nice demonstration of this, Nardi et al. (1993) observed the collaborative behaviors in a brain surgery operating room (OR), focusing on the use of live video in the OR. At some part of the surgery the neurosurgeon would look through a stereoscopic microscope to view the brain. A video camera mounted with the optics of the microscope captured what the neurosurgeon saw and broadcasted it to the other people in the OR, and to a remote audience over a network. Nardi et al. (1993) found that the OR coordination centered on this live video. The scrub nurse said: “the live video is the only indication we have of what’s going on in the head”, and used it to track the operation and anticipate which instruments and supplies the surgeon may need. The anesthesiologist reported not watching the video all the time, but still used it to keep aware of what was going in the OR so as to anticipate potential emergencies.

With regard to collaborative search, the importance of the immediate environment is less clear, as the visual search task does not explicitly change the environment, making it difficult to monitor ongoing progress. Group members may therefore have to rely on speech to communicate task progress, which may prove difficult given that this updating may require frequent references to spatial locations. This topic explored in the next section.

Communication media that use eye gaze Given the importance of establishing deictic relations between objects when referring to a location in space, it is not surprising that researchers would explore methods of efficiently communicating deictic information. For example, Argyle & Graham (1977) allowed people to use their finger to point to a map to indicate their focus of attention in a referential communication task involving locations on a map. Brennan (2005) also used a map reading task, but had her subjects use a computer mouse cursor to point. By providing the speaker with real time visual information about where the addressee thought the target location was, she found that the partners were able to "ground" on the target more quickly.

In addition to fingers and computer mice, people also can point with their eyes. Clark (1996) described how eye gaze can be used in spatial reference with a simple example: "I want you [gazes at A] and you [gazes at B] to come with me." In this case the addressee has to use the speaker's head-and-eye position to estimate who the director is "looking at". Gale and Monk (2000) demonstrated that the estimation was generally accurate, and that it remained accurate even when the head-and-eye position and work space were video-mediated, but this accuracy decreased with increasing distance between the speaker and the addressee.

The use of non-verbal pointing, however, may come with costs as well as benefits. Although people can perform visual search and talk at the same time, both pointing and search depend on vision, and might therefore conflict with each other. To the extent that a person would need to disengage from the search task to reference a partner's finger or gaze position, the ongoing search process may be disrupted.

It may be possible to minimize the potential interference between gaze pointing and a visual task by using an eye tracker to capture a person's gaze, then drawing it as a moving cursor on a screen-based representation of the immediate environment, just as Brennan (2005) did with a mouse cursor. This solution seamlessly integrates the gaze cursor with the work space. Now people can access the referred object in real-time, any time. Moreover, this solution can potentially make visible the often invisible ongoing search process. Group searchers can therefore know the regions of the display where their partners have looked, and how their search is progressing. In the following paragraphs I will review some existing efforts to integrate eye gaze in computer-mediated collaborative systems.

Clearboard is a video-mediated collaboration system that combines gaze awareness with the work environment (Ishii & Kobayashi, 1992; Ishii, Kobayashi, & Grudin, 1993). If participants work around a table, they cannot see each other's eye gaze and work space at the same time, meaning that they would have to switch their attention between the work space and their partner's eyes to obtain and use deictic information from gaze. This is not the case for the Clearboard scenario: where two participants work on opposite side of a half-transparent glass board. So they can look through glass workspace to obtain information about where their partner is looking at. This transparent drawing board, however, limits the materials that can be presented on it, and the Clearboard concept only works well for very small groups of collaborators.

The major technical difficulty in maintaining mutual gaze in video-mediated communication involves the alignment between the position of the camera and the position of the video on screen. Users have to look directly at the camera to make eye contact with their partners, but at the same time they must look to the screen to find out whether their partners were looking at them or not. “Video Tunnel” (Buxton & Moran, 1990; Acker & Levitt, 1987) addressed this problem by using half-silvered mirrors to align the camera and monitor in the same virtual position. A related device used by Monk and Gale (2002) combined the ideas of both Clearboard 2 and “Video Tunnel”. As in Clearboard 2, the participants can see through the transparent work display in order to access their partner’s eye position information. At the same time the device uses the “Video Tunnel” technique to achieve mutual gaze awareness.

Velichkovsky (1995) adopted a qualitatively different approach, illustrated in Figure 4. Two participants viewed the same display presented on two different monitors. The eye gazes of one of the participants were obtained using an eye tracker and projected on the partner’s monitor in real time. This person could therefore see where her partner was looking at any moment, thereby increasing awareness of the ongoing task process and enabling eye gaze to be used for spatial reference. Velichkovsky (1995) suggested that participants might even intentionally increase their fixation duration so as to purposefully communicate to their partner that they were referring to a particular point in space, and speculated that common ground regarding this use of fixation durations might be achieved very easily in a spatial referencing task. Despite its advantages, a weakness of the Velichkovsky system is that eye gaze information is communicated in only one direction. It is therefore restricted in its use as a collaborative system.

Vertegaal (1999)'s gaze groupware system (GGS) provided both video-mediated gaze awareness and a shared working space. In GGS, each participant was represented by a 2D image (persona) around a table in a virtual room, at a position that would be held by that remote participant. Each persona rotates around its x and y axes in 3D space, according to where the corresponding participant looks. The participants' eye movements were tracked with a desk-mounted eyetracking system. When person A looks at person B, B sees A's persona turn to face her. When A looks at person C, B sees A's persona turn towards C. This metaphorically conveys to whom each participant is listening or speaking to.

GGS also represented the shared table. When a participant looks at the shared table, a lightspot is projected onto the surface of the table, in line with her persona's gaze orientation. The color of this lightspot is identical to the color of her persona. This "miner's helmet" metaphor enables a participant to see exactly where the others are looking within the shared workspace. GGS even represented the shared document. When a participant looks within a file, all participants looking inside that file can see a lightspot with her color projected over the contents. This lightspot shows exactly what this person is reading.

Recently, Brennan, Chen, Dickinson, Neider, and Zelinsky (2007, in press), Zelinsky, Dickinson, Chen, Neider, and Brennan (2005) developed a collaborative visual search system that fully implemented bi-directional eye gaze communication. This system, illustrated in Figure 1, had two participants located in separate rooms but viewing the identical displays on synchronized computer monitors. Each participant also wore a head-mounted eye tracker. The display computers and eye trackers were all interconnected by an Ethernet hub. Central to this system is the exchange of gaze cursors between participants. The gaze position from each participant was transferred to their partner's display computer and drawn as a little yellow cursor on the partner's display screen in real-time. The participants could also communicate by a speech channel that was implemented using a bi-directional microphone-speaker system, routed through an audio mixer. Using this system, Neider, Chen, Dickinson, Brennan, and Zelinsky (2005) showed that this shared-gaze technique could be used with great success in a spatial referencing task. One person could refer to an object in space simply by fixating on its location, and their partner would know immediately that this extra long fixation meant that the target was found and located at the gaze cursor. Neider et al. (2005) also compared shared gaze collaboration to speech and found that shared gaze was the more efficient communication medium for spatial referring.

1.8. Group visual search

In this section I will review some modeling and empirical efforts to understand group visual search.

Modeling group visual search There have been some efforts on extending the solitary visual search model to cover multiple person conditions.

First I will brief introduce the solitary visual search model. Assuming that g is the probability of target detected by one fixation, the probability of target detected by n fixations, p_n, should be:

$$p_n = 1 - (1 - g)^n \tag{2}$$

Note that this model assumes an independent sampling of each fixation, meaning that previous fixations do not influence future fixations (Koopman, 1956). The process expressed in Equation 2 is therefore a form of memoryless search model (Horowitz & Wolfe, 1998, but see Dickinson & Zelinsky, 2005). Note that this simple model of number of fixations can also be extended to a model of target acquisition time:

$$p(t) = 1 - e^{-\gamma t} \tag{3}$$

Where $p(t)$ is the probability of target detection during time t, and γ is the instantaneous probability density of detection.

Following this "classical approach" of modeling target acquisition, Bailey (1970) proposed that the probability of acquiring a target, P_R, could be broken down into 3 components: P_1: the probability that a given fixation will locate the target region of a scene, P_2: the probability that the target will be detected once it is fixated, and P_3: the probability that the target will be identified once it is detected. In the equation $P_R = P_1 \times P_2 \times P_3$, only P_1 is dependent on search time, both P_2 and P_3 are search time independent.

The probability that a fixation will be made to the target area during time t, $P_1(t)$ can be defined as the first arrival time of a Poisson process:

$$P_1(t) = 1 - e^{-t/\tau FOV} \tag{4}$$

Where τFOV represents the mean acquisition time given that a target is fixated. Bailey (1970) also derived terms for P_2 and P_3, but these can be replaced with P_∞: leaving with unlimited search time, the proportion of the normal observer ensemble who will finally find the target:

$$P_2 \times P_3 \equiv P_\infty = \frac{(N/N_{50})^E}{1+(N/N_{50})^E} \tag{5}$$

And the probability of acquiring the target as a function of time as:

$$P(t) = P_\infty[1-e^{-p_0 t/t_f}] \tag{6}$$

Rotman (1989) extended the above-described model of solitary search to a group search situation, under the assumption that all observers were searching independently for a target in the same scene and from the same perspective, and that the group's detection time was determined by the fastest response of any group member. Applying Rotman's formulation to the two searcher case, the probability of at least one detecting the target at given time t should be:

$$P_{1+2}(t) = 2P_\infty[1-e^{P_0 t/t_f}] - \{P_\infty[1-e^{P_0 t/t_f}]\}^2 \tag{7}$$

Extending this two-person example to an N group search condition is trivial. However, although a formal expression of search time in terms of group size and number of fixations is impressive, one must keep in mind that the model assumes an independent sampling of fixations by each observer, and that the multiple observers in a group are searching independently. Given that both of these assumptions are highly questionable, the model's applicability to predict actual group search behavior is likely to be very limited.

Empirical investigations of group search Despite the large number of studies devoted to understanding solitary search processes, the additional group processes underlying collaborative search remain largely unexplored (Chatziastros & Bultho, 2006; Malcolmson, Reynolds, & Smilek, 2007). Malcolmson et al. (2007) investigated the collaboration influences on dyadic visual search performance. By using signal detection analyses, they compared the performances of the collaborative dyadic searchers with the nominal pairs, and found that collaborative pairs had a greater sensitivity and conservative response bias than the nominal pairs. To date, the most detailed investigation of collaborative visual search was conducted by Zelinsky et al. (2005), Brennan et al.(2007, in press), who used a shared gaze paradigm (see Figure 1) to study how groups of searchers divide the labor of search. Three main findings emerged from their study. First, collaborative search far surpassed solitary search performance, and was even significantly better compared to the nominal group. Even a highly efficient cognitive process like visual search can therefore benefit from collaboration. Second, collaborative benefits stemmed largely from a "spatial division of labor" search strategy. An analysis of the eye movements revealed that searchers typically divided the labor of the search task, with one person searching the top of the display and the other the bottom, or one person searching the left side of the display and the other searching the right. Moreover, searchers were apparently using gaze cursors communicatively. By restricting gaze to one side of the display, that searcher was sending her partner a message: "I will search this side of the display, you search the other side." And that message was clearly received. Third, collaboration mediated by gaze cursors was more efficient than collaboration mediated by speech. Indeed, adding a speech channel to shared gaze introduced a cost compared to having gaze cursors alone.

Chapter 2: Experiments

2.1. Overview of Experiments

This thesis investigated the intermediate strategy coordination and micro-level behavioral coordination in collaborative visual search. With regard to strategy coordination, this thesis explores the influences of group size, task characteristics, and communication medium on strategy coordination. With regard to micro-level behavioral coordination, I tried to demonstrate the existence of a particular behavioral coordination.

With regard to group size, Brennan et al. (2007, in press) already demonstrated the strategy collaboration of dyadic searchers. But given the special role that dyads seem to play in the group literature (Valacich et al.,1994), one is reluctant to generalize the findings from Brennan et al. (2007, in press) to larger groups. There are good reasons to believe that collaborative benefits may not scale linearly with group size. In Experiments 1-3, I investigated the potential effects of group size on collaborative searches.

With regard to task characteristics, Brennan et al. (2007, in press) demonstrated that dyadic searchers readily adopted a spatial division-of-labor (DOL) strategy. But this specific DOL strategy might be determined by the specific search task used in Brennan et al. (2007, in press)'s experiment. It is possible that strategy selection is flexible, and that very different strategies may be used depending on the task. Experiment 1-3 explored this space of coordination strategies by devising tasks expected to promote division of search labor by space (experiment1), features (experiment2), and target templates (experiment3).

Verbal communication is crucial to the formation of explicit coordination strategies during visual search (Brennan et al., 2007, in press). But verbal communication may also introduce distractions and slow group search. These effects of communication medium are addressed in the context of Experiments 1-3. But does no communication means no coordination during search, as assumed by Rotman (1989)? Although explicit coordination would not be possible without communication, groups could still coordinate their behavior tacitly. In Experiment 4 I explored the potential for tacit coordination in the absence of communication using signal detection theory.

In addition to verbal communication there is also shared gaze communication: group searchers can see each others' gazes drawn on the search display in real time (Figure 1). Brennan et al. (2007, in press) demonstrated the utility of shared gaze cursors as a medium for forming coordination strategies, and they also found that at least for a spatial task, shared gaze was more efficient than verbal communication. My hypothesis is that the advantage of shared gazes resulted from the use of more dynamic micro-level coordination strategies, made possible by the awareness of each others gaze. Besides the intermediate level coordination: the relatively coarse "you look left, I'll look right" divisions of labor, partners sharing gaze might adjust their coordination strategies moment-by-moment, much like a couple dancing across a ballroom.

Crucial to understanding this micro-level behavioral coordination will be our ability to quantify how search behavior respectively changes as a result of one person seeing the other's gaze cursor. Informing this analysis is a study on group foraging behaviors of tonkean macaques. Drapier, Ducoing, and Thierry (1999) observed macaques foraging either singly or in a group for a hidden banana. Although the monkeys in a group appeared to divide the search space, they were also clearly monitoring each others behavior, as evidenced by the fact that, when one monkey lingered too long in one spot, the others flocked to his location thinking that the food was discovered. Colston and Schiano (1995) observed a related behavior with respect to human gaze, finding that observers could tell whether a person found a problem difficult based on how long their gaze lingered on a problem before moving on to the next.

I expect to find similar evidence for behavioral coordination using eye gaze in the context of a search task, although this effort is likely to be more exploratory compared to the other questions addressed in this thesis. Given the possibility of many different forms of micro-level behavioral coordination, I narrowed the focus to look specifically for evidence of targeted assistance. When dyadic searchers can see each other's eye gaze, they will be able to monitor the other's search progress and therefore might be able to assist the other. With shared gaze, this assistance probably will not be random, but rather *targeted* to those parts of the partners search area that have not yet been searched. Exploring the evidence for targeted assistance will be the topic of experiment 5.

2.2. Experiment 1

The main purpose of experiment 1 is to see whether the spatial division of labor collaborative strategy reported by Zelinsky et al. (2005), could also be extend to 4-person groups.

Design and Participants

The experiment had a 2 (verbal communication vs. no-communication) x 2 (group size 2, 4) between-subjects design. There were 6 groups for each condition. There were also 6 persons for the single visual search condition. So there were a total of 1 x 6 + (2 x 6 + 4 x 6) x 2 = 78 subjects. There were 20 practice trials and 60 experimental trials for each person or group. All the participants were from the Stony Brook University subject pool and have normal or corrected to normal visual acuity and color vision.

Stimuli and Apparatus

The subjects were instructed to search for a vertical or horizontally oriented oval dot embedded in an array of 494 circular dots (see Figure 2). The circular dots were 12 pixels in diameter, about 0.49° visual angle. The vertically oriented oval dot was 0.49° (12 pixels) in height and 0.41° (10 pixels) in width; the horizontally oriented ellipse is 0.49° in width and 0.41° in height.

Each search stimulus was created by covering the screen with a 38 (rows) x 26 (columns) invisible grid, then filling in 50% of the grid cells with dots. This yields 38 x 26 x 50% = 494 dots for each stimulus. One of these dots was randomly selected and replaced with an oval target dot. Half of these targets were vertically oriented, the other half were horizontally oriented. All stimuli were presented in black against a white background, and displayed using a 17 inch ViewSonic flat screen CRT color monitor at a screen resolution of 1280 x 960 pixels and a refresh rate of 70 Hz. The viewing distance was approximately 34 cm, creating a 49° (H) x 38° (V) view of the each stimulus array.

The position of the group members during testing is shown in Figure 3. All searchers were physically present in the same room. At a group size of 2, searchers sat side-by-side but were physically separated by a barrier, thereby preventing them from seeing each other (and exchanging visual cues). At a group size of 4, two additional searchers were positioned with their backs facing the backs of the other two, thereby again preventing the exchange of visual cues (an experimenter was in the room monitoring that the subjects do not turn around to face their partners).

A Share Plus® Super VGA Multiplier, driven by a single Pentium-based computer, was used to simultaneously display the identical search image to all four of the subject monitors. Manual responses were collected by using 4 Microsoft SideWinder gamepads, which were connected to 4 USB ports on the computer. Custom C++ software incorporating DirectX functionality and running under Microsoft Windows XP was used to control the experiment.

Procedure

Each trial started with the presentation of a central fixation point, which would be replaced by the search scene after 1 second. Observers were instructed to identify, as quickly and as accurately as possible, the target. They were asked to press the right trigger of the gamepad if the target was a horizontally oriented ellipse, or the left trigger if the target was a vertically oriented ellipse. A target was presented on each trial. The first person who pressed the button would terminate the search for all, and this response would be taken as the group response. After pressing the button, the subjects were shown a feedback displayed at the center of the screen, indicating who made the response, and whether it was correct or incorrect, along with a running total of the correct and incorrect responses made by the whole group. The feedback message disappeared after 2 seconds and the next trial would commence. A 200 msecs 2000 Hz tone would signal the incorrect responses. The experiment was auto-paced, but subjects could ask for a break between the trails at any time. The entire experiment was in one session and lasted approximately 30 - 40 minutes.

Results and Discussion

The overall error rate was only 3.9%. All incorrect trials and trials with a search time more than 100 seconds were excluded from further analysis.

Before analyzing the search data I first transcribed the digitally recorded speech in the communication condition, and then coded explicit references to collaborative division of labor strategies. Transcription was conducted by the author, and then confirmed by a native English speaking research assistant.

Communicating groups were classified into two categories based on whether or not their transcripts contained explicit references to a division of search labor. Five out of the six transcripts contained the division of labor consensus (see Appendix for the fragment transcripts that contain the discussion of division of labor). Among these five groups, three reached the consensus on dividing the display into top and bottom half, and two divided the display into right and left half. Only one out of the six 2-person communication groups showed no signs of reaching any consensus on division of labor. Also, all the five division of labor groups reached division of labor consensus during the practice trials, even before the start of the formal experiment.

For those data in which explicit division of labor assignments could be made (e.g., partner A agreed to search the top half of the display, and partner B agreed to search the bottom half), I then determined whether these division of labor strategies were actually used during search. I did this by correlating a subject's agreed upon region with target detected by that subject. These analyses are shown in Figure 4. Turning first to the data from the three groups who made top/bottom divisions (Figure 4A), we can see that partners who agreed to search the top half of a display were far more likely to detect targets appearing in the top half compared to partners who agreed to search the bottom half of a display, and vice versa. A similar pattern was found for the two groups who made a left/right division of search labor (Figure 4B). These data suggest that our reverse correlation technique is capable of illustrating spatial divisions of labor in a search task. Our data are also consistent with the evidence for spatial division of labor reported in Brennan et al. (2007, in press) and Zelinsky et al. (2005) using the shared gaze paradigm, suggesting that spatial divisions of labor search strategies are not linked to the use of an eye tracking methodology.

We know that dyads adopted a spatial division of labor collaborative strategy, but does this strategic collaboration extend to 4-person groups? To answer this question, I transcribed and coded the speech for the 4-person communication groups using the same method described for the dyads. Four of the 6 4-person communication groups used a spatial division of labor search strategy. Of these, three divided the search display into four quadrants, with each participant taking responsibility for a quadrant. The other group divided the display top and bottom, with 2 people searching each half. The remaining two groups did not reach an explicit consensus on a spatial division of labor strategy. See Appendix (Exp1: 4-person groups) for transcripts of these divisions of labor.

As in the case of the dyads, behavioral analyses of the 4-person communicating groups also showed strong evidence for a spatial division of search labor. Partners who agreed to search the upper-left quadrant (1) were more likely to detect targets when they appeared in that quadrant. The same pattern held for the upper-right (2) and lower-left (3) quadrants, although targets appearing in the lower-right were not preferentially detected by any one partner. However, this relatively clear evidence for a quadrant-based division of labor was found in only 50% of the communicating 4-person groups, a far lower percentage than what I found for dyads. This effect of group size might be due to difficulty in dividing a display into smaller and smaller regions (i.e., it is easier to divide a display in two than in fourths), or it may be harder to keep attention confined to a small region compared to a larger region (i.e., attention may be more prone to wander outside the confines of a small region). This matter will need to be addressed in future work.

The average RTs from Experiment 1 are shown in Figure 5 as the black dot dash lines. I conducted a 2 (communication, non-communication) x 2 (2 or 4-person groups) between-subject ANOVA (I will discuss the 1-person data in experiment 2) and found shorter RTs for the non-communication groups ($M = 9053$ msecs) than communication groups ($M = 12855$ msecs), $F(1,20) = 4.79$, $p < 0.05$, and shorter RTs for the 4-person groups ($M = 9176$ msecs) than 2-person groups ($M = 12732$ msecs), $F(1,20) = 4.19$, $p = 0.05$. The interaction between communication condition and group size was not significant, $F(1, 20) = 0.82$, $p = 0.38$. Given that individuals in communicating groups would in principle need to search only $1/n$ (where n = group size) of the display, why is it that non-communicating groups searched faster? I attribute this puzzling finding to the existence of speech costs. As argued in Brennan et al. (2007, in press), speech may incur costs as well as benefits associated with the strategic division of search labor. Using our non-communicating group as a baseline, the costs of speech clearly outweighed the benefits in our task.

To better determine the degree of costs and benefits in our communication and non-communication search conditions, I created 15 2-person and 15 4-person nominal groups. Nominal groups of size N (where N equals 2 or 4) were created by selecting N individuals from the pool of M ($M = 6$ in this experiment) solitary searchers. The number of nominal groups of size N that can be created from M individuals is determined by:

$$\frac{M!}{N!(M-N)!} \tag{8}$$

Because individual searchers performed the same sequence of search trials, I can combine the responses of participants on a trial-by-trial basis. The nominal group's response for a given trial was defined by the fastest individual response for that trial regardless of accuracy. The performance of a nominal group therefore represents the pooled best efforts from the same number of individuals working alone. As also illustrated in Figure 5, a 2 (nominal, communication) x 2 (2 or 4-person groups) between-subject ANOVA revealed significant main effects of search condition (nominal or communication) and group size, but no significant interaction. RTs in the nominal group condition were shorter than in the communication condition, F (1, 38) = 4.04, p = 0.05, and the RTs of 4-person groups were shorter than 2-person groups, F (1, 38) = 29.72, p < 0.01. Again, I attribute these faster RTs in the nominal condition as indication of communication costs overriding collaborative benefits. Analysis of the non-communication condition revealed a very different pattern of results. A 2 (nominal, non-communication) x 2 (2 or 4-person groups) between-subject ANOVA revealed that the interaction of search condition (nominal or non-communication) x group size was significant: F (1, 38) = 4.2675, p = 0.046. RTs in the 2-person non-communication condition (M =10046 msecs) were shorter than in the 2-person nominal condition, t-test revealed that it was marginally significant, t = 1.78, p = 0.06. There were no significant differences between the RTs of the 4-person non-communication groups (M = 8061 msecs) and the 4-person nominal groups (M = 7933 msecs), t = 0.18, p = 0.86. I will address the question of why only the non-communication dyads outperformed the corresponding nominal groups in experiment 4.

2.3. Experiment 2

All of the search elementsinExperiment1, as in the Brennan et al. (2007, in press) study, were identical, meaning that the spatial division of labor strategy observed in these experiments may have resulted from the unavailability of information needed to mediate more sophisticated collaborative search strategies. In this experiment I explored the possibility of dividing the search labor by features, rather than by space.

Design, Participants, and Procedure

Same as Experiment 1, none of the subjects from Experiment 1 participated in this experiment.

Stimuli and Apparatus

The same dot stimuli and search task as described in Experiment 1 were used, with the only difference being that the dots and targets were presented in color. The 494 elements per search display were divided evenly (as much as possible) into 4 color groups: red (RGB: 255, 0, 0), blue (RGB: 0, 0, 255), green (RGB: 0, 255, 0) and black (RGB: 0, 0, 0) (see Figure 7). Note that this similarity between experiments 1 and 2 extended to the spatial configurations of elements in each search display, meaning the target location and the positions of the dots on each trial were also identical between the two experiments. The elements composing each color group are spatially segregated into four irregularly shaped regions, thereby thwarting the use of a simple spatial division of labor collaborative search strategy.

Results and Discussion

The overall error rate was only 2.4%. All incorrect trials and trials with a search time longer than 100 seconds were excluded from further analysis.

The same conventions used to transcribe and code division of labor in the communication groups were also used here, with the exception that divisions of labor were now further categorized in terms of space or color. For the 2-person communication groups, five out of six reached a consensus on a division of labor strategy; one group showed no signs of attempting such a consensus (from the speech record). Among the five groups that reached a division of labor strategy, three divided the search display spatially as in experiment 1, and two divided the display by color, with each member taking responsibility for two colors (see Appendix, Exp 2: 2-person groups). In contrast, all 6 of the 4-person communication groups settled on a color division of labor strategy; each person took responsibility for one of the distractor colors (see Appendix, Exp2: 4-person groups).

Figure 8 groups the 4-person data by member of the collaborating group and target color. Clearly, target detection depended on the color of the target and the person assuming responsibility for searching items of the target color. This suggests that group members not only proposed a color division of labor strategy, they actually used it. These data also suggest that subjects preferred to divide the search labor by color (feature) rather than space, at least for the 4-person groups. I do not know why this color division of labor preference was less clear in the 2-person groups. One possibility is that spatial divisions of labor are very natural for dyads, and difficult to override, but that this division is less natural for larger sized groups. Another possibility is that the ambiguity involved in each partner picking two colors (rather than just one) introduced a cost that tipped consensus in favor of a spatial division of labor strategy.

The average RTs are shown in Figure 5 as the solid lines. A 2 (communication, non-communication) x 2 (2 or 4-person groups) between-subject ANOVA only revealed a significant main effect of group size, the RTs of the 4-person groups (M = 8884 msecs) were shorter than the 2-person groups (M =14818 msecs), $F(1, 20) = 37.12$, $p < 0.01$. However, and contrary to Experiment 1, RTs in the non-communication were not faster than those in the communication condition, $F(1, 20) = 0.83$, $p = 0.37$. I speculate that this may be due to a slightly larger collaborative search benefit with colored dots offsetting the communication costs. The colored regions would likely create better delineated regions for labor division compared to the coarser "you look left, I'll look right" divisions used in Experiment 1.

Figure 5 also shows an interesting trend towards longer RTs in the 1-person colored dots task (M = 25276 msec) compared to the1-person data from experiment 1 (M = 28014). Although this difference did not reach significance, t = 0.62, p = 0.55. The trend suggests that searching by color groups may not be as efficient as a grid search, perhaps because the groups are irregular and required the subjects to periodically back-track during their search. Further studies are needed to investigate the different roles of visual grouping in single and group searches.

As in experiment 1 I also compared the communication and non-communication conditions to 15 2-person and 15 4-person nominal groups constructed from the 1-person search data. A 2 (nominal, communication) x 2 (2 or 4-person groups) between-subject ANOVA again revealed only a significant main effect of group size, F (1, 38) = 112.16, p < 0.01. A similar analysis conducted on the non-communication groups and the nominal groups, a 2 (nominal, non-communication) x 2 (2 or 4-person groups) between-subject ANOVA also revealed only a significant main effect of group size, F (1, 38) = 107.25, p < 0.01. I interpret these data as suggesting that the costs of coordinating search using speech may largely offset any collaborative benefits, resulting in no net advantage relative to a nominal group.

2.4. Experiment 3

Every search task requires keeping in mind, presumably in visual working memory (VWM), the target long enough to successfully complete the search. But what if one is searching simultaneously for multiple targets? In this eventuality, the number of targets might exceed our limited VWM capacity (Alvarez & Cavanagh, 2004; Irwin & Zelinsky, 2002; Luck & Vogel, 1997) thereby resulting in a difficult and potentially error-prone search. Under these conditions, might a group of people searching collaboratively elect to divide the search labor by memorizing targets collaboratively?

Design and Participants

1, 2 or 4 subjects were asked to search for 1, 4, or 8 targets at the same time. Search displays were either target-present or target absent (divided 50/50), and in a target-present trial only one target would be presented, regardless of the number of possible targets (1, 4, or 8). This experiment therefore had a 2 (verbal communication vs. no-communication) x 3 (group size 1, 2, 4) x 3 (target set 1, 4, 8) x 2 (target present or absent) mixed design. Communication condition and group size were between-subjects variables; target presence was a within-subjects variable that was randomly interleaved over trials, and target set size (1, 4, 8) was a blocked and counterbalanced within-subjects variable. There were 6 practice and 60 formal trials for each target size, so there were totally 18 practice and 180 formal trials for each group. We had 6 groups for each communication (2) and group size (3) condition, resulting in a total of 1 x 6+2 x (2 x 6 + 4 x 6) = 78 participants. All participants were recruited from the Stony Brook University subject pool, with the constraint that they have normal or corrected to normal visual acuity and color vision, by self report.

Stimuli and Apparatus

The stimuli are photorealistic images of objects selected from the Hemera R Photo-Objects collection. The height and width of the objects are variable, but the images were resized so that they can fit within a 90 x 90 pixel bounding box (3.6 x 3.6). None of the objects were repeated throughout the experiment, and the apparatus used to present the images was the same as described in experiment 1.

The search display subtended a visual angle of 42° x 30°. The number of objects appearing in each search display was held constant at 14 for all trials. The 14 objects comprising each search display were positioned randomly with the following constraints: no objects were presented within 8° visual angle of the initial central fixation cross, and objects were separated by a minimum center-to-center distance of 8° visual angle. See Figure 9.

Procedure

Each trial started with a central fixation point presented for 1 second, followed by a target preview display. The target preview had 1, 4 or 8 objects. In the 1 target condition, the single target appeared at the center of the screen; in the 4 target condition one target appeared in each of the 4 display quadrants; in the 8 target condition the targets were arranged into 2 4-object rows.

The duration of the target preview depends on the number of objects appearing in the preview display, as determined by a “1 second per object” rule. A 1-object preview therefore was displayed for 1 second, and an 8-object preview was displayed for 8 seconds. Following the preview, the search display was presented and participants were instructed to indicate, as quickly and as accurately as possible, whether any one of the targets from the preview appeared in the search array. A target present judgment was indicated by a left trigger press; a target absent judgment was indicated by a right trigger press, again using a standard gamepad as described in Experiment 1. Only one target appeared in the search array on target present trials, and this object was selected randomly from the multiple preview targets in the 4 and 8 target preview conditions.

Results and Discussion

Conversation in the communication group condition was transcribed and coded as described in experiment 1, except now coding included explicit references to target division of labor (see Appendix).

Turning first to the 4-target data, five of the six 2-person groups in the communication condition divided the search task by target, each taking responsibility for searching two. Only one 2-person group failed to reach consensus on any division of labor search strategy. Similarly, five of the six 4-person groups reached a target division of labor strategy; again one group apparently failed to reach consensus on any search strategy. As in experiments 1 and 2, these explicit verbal agreements were supported by analysis of the search behavior. Figure 10A clearly shows a relationship between target position in the preview and the probability of a specific partner detected that target in the search array. Figure 10B shows the same relationship for 4-person groups. Partners therefore not only agreed to divide the search labor by target, they succeeding in executing this search strategy.

Evidence for the formation and use of a target division of labor search strategy was even more pronounced in the 8-target communication condition. All six of the 2-person groups reached consensus on dividing the search labor in half, with each searching for 4 targets. All of the 4-person groups reached a similar consensus, each taking responsibility for 2 targets. These division of labor strategies were again clearly expressed in the behavioral data; partners preferentially detected non-overlapping subsets of the search targets (Figure 11).

I conclude that groups in this task found it more useful to divide the search labor by preview target than by either feature or spatial region in the search display. However, there were exceptions, cases in which groups failed to reach consensus on a division of labor strategy. Interestingly, both of these exceptions occurred in the 4-target data. I speculate that, because 4 objects is still within an individuals visual working memory capacity, there was less need in the 4-target condition to divide the search labor. However, with 8 targets this limit would clearly be exceeded, thereby making collaboration essential.

Figure 12 shows the RTs (top two panels) and error rates (bottom two panels) for experiment 3. Given that high error rates complicate interpretations of RT, I will focus our analysis on errors.

A 2 (communication, non-communication) x 2 (2, 4-persongroup) x 2 (target present, absent) x 2 (4 or 8 targets) mixed ANOVA was conducted on the error rates. The communication condition and group size were between-subject factors; target presence and target set size were within-subject factors. I found significant main effects of communication condition ($F(1, 20) = 49.12$, $p < 0.01$), target presence ($F(1, 20) = 13.64$, $p < 0.01$), and target size ($F(1, 20) = 12.42$, $p < 0.01$). More interestingly, there were also 3 significant or marginally significant two-way interactions.

The two-way interaction between communication condition and target set size was significant, $F(1, 20) = 8.62$, $p < 0.01$. For the communication groups, when the target size increased from 4 to 8, the error rates did not increase significantly, $t = 0.38$, $p = 0.71$. But for the non-communication groups, when the target size increased from 4 to 8, the error rates increased dramatically, $t = 2.24$, $p = 0.03$, almost to the 50% level expected by chance. Communicating groups were not affected by increasing the number of targets because they divided the targets among them. Eight targets were beyond an individual searcher's memory capacity, but not the group's memory capacity.

The interaction between communication condition and group size was significant, $F(1, 20) = 4.73$, $p = 0.042$. For the communication groups, there was a marginally significant trend towards lower error rates with the increasing group size, $t = 1.89$, $p = 0.07$. For the non-communication groups, the error rates tended to increase with group size, although this was again not quite significant, $t = 1.77$, $p = 0.08$. For communicating groups, increasing group size means each member will have less search labor, resulting in improved search performance. Because efficient target division of labor strategies could not be used in the non-communication groups, search efficiency would not be expected to increase substantially with group size. Finally, the interaction of group size x target presence was significant, $F(1, 20) = 5.12$, $p = 0.035$. For the target present trials, the miss rates were not significantly affected by increasing group size from 2 to 4, $t = 0.99$, $p = 0.33$, despite a small downward trend. However, for the target absent trials, increasing group size from 2 to 4 resulted in more false alarms, $t = 1.90$, $p = 0.06$.

As I mentioned before, target present search is a Eureka type task: the searcher who finds the target does not need to obtain consensus from the other members. However, target absent search does require group consensus, at least in a limited form. The searcher who responds "target absent" must believe that all of the other members also had adequate time to search yet failed to find the target. So for target present trials, the larger the group the more chances at least one searcher might find the target, which is why increasing group size decreased the error rates (or at least did not increase the error rates). However, the reason for errors increasing with group size in the target absent condition is harder to explain. Based on a pooling effect alone, one might expect errors to also decrease with larger groups, but the opposite pattern emerged in the data. Why did false alarms increase with group size?

To attempt an answer to this question I computed the perceptual sensitivity (d') and response bias ($\ln\beta$) based on an analysis of errors in experiment 3 (Figure 13). First I conducted a 2 (communication, non-communication) x 2 (2, 4-persongroup) x 2 (4 or 8 targets) mixed ANOVA on the perceptual sensitivity, d'. The main effect of communication was significant, $F(1, 20) = 48.50$, $p < 0.01$, and the two-way interaction of communication x group size was marginally significant, $F(1, 20) = 3.38$, $p = 0.08$. For the communication groups, the sensitivity increased when group size increased from 2 ($d = 2.78$) to 4 ($d = 3.16$), but this increase was not statistically significant, $t = 1.28$, $p = 0.22$. For the non-communication groups, the sensitivity decreased when the group size increased from 2 (d =1.36) to 4 (d = 0.72), and this difference was marginally significant, t = 1.84, p = 0.079. For the communicating groups, I attribute the stable (or at least non-decreasing) sensitivity to a target division of labor between the group members. However, the drop in sensitivity with group size in the non-communication condition suggests that members in larger groups, when prevented from communicating, may elect to compete, and thereby spend less time on the perceptual task.

The main effect of target set size was significant, F (1, 20) = 6.67, p = 0.018, as was the two-way interaction of communication x target size, F (1, 20) = 7.10, p = 0.015. For the communication groups, sensitivity was not affected by increasing the targets from 4 to 8. However, for the non-communication groups, sensitivity decreased as the number of targets increased from 4 to 8, t = 2.3174, p = 0.03. Because the non-communicating groups were not able to efficiently divide the target labor, their sensitivity dropped as the number of targets increased.

I also conducted a 2 (communication, non-communication) x 2 (2, 4-person group) x 2 (4 or 8 targets) mixed ANOVA on the response biases. The main effect of communication was significant, F (1, 20) = 14.67, p = 0.001, as was the two-way interaction between communication condition and target set size, F (1, 20) = 7.69, p = 0.012. The communication groups became more conservative in their responses when the number of targets increased from 2 to 4, t = 1.79, p = 0.08, which is reasonable given the increased difficulty of the task. However, the non-communication groups became more liberal when the number of the targets increased, although this trend was not significant, t = 1.34, p = 0.19. Together, these sensitivity and bias analyses suggest a clear difference between communication and non-communication groups in Experiment 3. As expected, the increase in task difficulty resulting from an increase in the number of targets resulted in poorer sensitivity in the non-communications groups. This increase in task difficulty was largely offset in the communicating groups by the target division of labor strategy, resulting in no significant change in sensitivity. However, despite the greater task difficulty and reduced sensitivity, the non-communicating groups decided to adopt more liberal response criteria; perhaps reflecting some frustration with the inability to do the task, or the emergence of some competitive behavior under non-communication conditions.

For completeness, Figure 12A & B show the RTs from Experiment 3. First I conducted a 2 (communication, non-communication) x 2 (2, 4-persongroup) x 2 (target present, absent) x 2 (4 or 8 targets) mixed ANOVA on the RTs, the communication condition and group size were between-subject factors; target presence and target size were within-subject factors.

The 3-way interaction of communication x target presence x target size was significant, $F(1, 20) = 16.83$, $p < 0.01$. I therefore further split the data into communication and non-communication subsets, and conducted a 2 (2, 4-persongroup) x 2 (target present, absent) x 2 (4 or 8 targets) mixed ANOVA on each group separately.

Only for the communication groups was the two-way interaction between target presence and target set size significant, $F(1, 20) = 30.52$, $p < 0.01$. Although search times increased for both target present trials and target absent trials as the number of targets increased from 4 to 8, this increase was far larger in the target absent trials. I interpret this interaction as suggesting that the communication groups needed more time to reach consensus in the target absent trials.

There were also other significant effects in the RT data, such as an effect of communication condition and group set size. However, recall that these differences were also characterized by extreme differences in error rates. Given this potential for speed-accuracy trade-offs, interpretations of response time become impossible.

2.5. Experiment 4

In experiments 1 & 2 I found that the 2-person non-communication groups searched an average of about 3 seconds faster than the corresponding nominal groups. But that benefit disappeared for the 4-person non-communication groups. This benefit of non-communication search over nominal search is curious, and was left unexplained. In experiment 4 I attempt to understand the cause of this mysterious benefit.

As I already mentioned, the nominal groups were constructed by pooling the fastest responses of individual searchers on a trial-by-trial basis. Nominal group performance therefore represents the best performance with the assumption that the searchers were acting independently and were not influenced by group membership. The fact that the 2-person non-communication groups outperformed the corresponding nominal groups suggests that the members of the non-communication groups had to somehow be influencing each other's behavior, even in the absence of communication. How might this be possible?

There are two possible explanations. One possibility appeals to tacit communication. Even without communication, a group of searchers could have adjusted their search behavior based on some blind assumptions of their partner's search behavior. For example, most searchers prefer to start their search from the upper-left corner of the search display (Zelinsky, 1996). Realizing this, a perceptive search partner might therefore start their search from the bottom-right so as to reduce search redundancy. Such tacit coordination might explain the performance benefit observed in the non-communication groups.

How can we determine if tacit coordination was used in experiments 1 and 2? If a group of searchers divided the search labor, regardless of whether this was done explicitly or tacitly, each searcher would be able to confine their search to a smaller region of the display resulting in a greater perceptual sensitivity compared to individual searchers or a group of searchers who are not dividing search labor. The communication groups, because they explicitly divided the search display by quadrant or by color, certainly should have a greater perceptual sensitivity than individual searchers, which is confirmed by the large group advantage in this study, and in previous research (Malcolmson et al., 2007). Given that division of labor can be inferred from increased perceptual sensitivity, if the 2-person non-communication groups had a greater perceptual sensitivity than individual searchers, I can infer that they must had divided the search labor. Because they could not communicate, this division would have to have happened tacitly.

A second possible explanation is that the performance benefit in the 2-person non-communication groups could come from the occurrence of competition among the group members, similar (although probably less pronounced) to the competition observed in Experiment 3. Given that collaborative search is a kind of disjunctive unitary task (Steiner, 1972), one in which only the fastest responder contributes to the group performance on each trial, it is possible that group members might set aside the instructions to collaborate and start to compete under conditions of non-communication.

We know that people are ordinarily conservative in making a search judgment, and they tend to become even more conservative when searching as part of a communicating group (Malcolmson et al., 2007). This means that searchers, especially communication group searchers, require a lot of evidence before making a judgment (e.g., indicating that a target is present). However, when they are competing, searchers may adopt a more liberal decision criterion to speed up their responses, thereby requiring less evidence of target presence or absence when indicating their response. Such a hasty search might also result in reduced estimates of sensitivity, as people may spend relatively little time inspecting some regions of the display, or skip some regions entirely. Evidence for competition within a group might therefore be expressed as a more liberal response bias and less perceptual sensitivity.

In summary, to the extent that the non-communication groups had a greater sensitivity than individual searchers, we can infer the existence of tacit coordination. However to the extent that the non-communication groups had a liberal response bias and less sensitivity, we can infer that the performance benefit might come from the occurrence of competition.

Design and Participants

The experiment was a 3 (individual, 2-person non-communication group, and 2-person communication group) x 2 (target present, absent) mixed design. Target presence was a within-subject factor, and the search condition was a between-subject factor. There were 13 persons or groups for each search condition, so there were a total of 1 x 13 + 2 x 13 + 2 x 13 = 65 subjects. There were 10 practice trials and 120 formal trials for each person or group. The formal experiment was separated into two sessions and each session had 60 trials. All the participants were from the Stony Brook University subject pool and had normal or corrected to normal visual acuity and color vision, by self report. None participated in the previous experiments.

Stimuli and Apparatus

The search stimuli were same as the ones used in Experiment 2. The one exception was that in this experiment the task was detection rather than discrimination: the subjects had to report the presence or absence of the oval target. The target (a vertical or horizontal oval dot, in equal chance) was present on 50% of the trials. This experiment also differed from Experiment 2 in that it was time limited. The search display was presented for 20 seconds in the 1-person condition and 10 seconds in the 2-person conditions. The 10 and 20 seconds presentation time for 1-person and 2-person conditions were derived from experiment 2 and they are the times that generated 50% error rates. Both the shift to a detection task and the use of time-terminated displays was done to encourage more errors in this task (compared to experiment2), a prerequisite for an error analysis.

Procedure

Each trial started with a central fixation point, which was replaced by the search scene after 1 second. Observers were instructed to press the right trigger of the gamepad if the target was present, and the left trigger if the target was absent. The search display would terminate upon either button press, or after a constant display time (20 seconds for

1-person and 10 seconds for 2-person conditions), If the search display ended by time-out, a blank display would be presented with the text “please press a button to continue” Subjects were therefore forced to make a target-present or target-absent response in order to continue to the next trial. Feedback would be displayed at the center of the screen after the button press, indicating who pressed the button, whether it was correct or incorrect, and also the total number of the correct and incorrect responses made by the whole group. Incorrect responses were indicated by a 200 msecs 2000 Hz tone. The feedback message disappeared after 2 seconds and the next trial commenced.

Results and Discussion

Table 1 shows the mean hit rate, false alarm rate, sensitivity, and response bias for the individual searchers, non-communication groups, and communication groups, respectively. A one-way ANOVA revealed a significant difference in sensitivity between these conditions, $F(2, 36) = 6.54$, $p < 0.01$. A t-test with bonferroni correction confirmed that the communication groups had a greater sensitivity than both the individual searchers, $p = 0.02$, and the non-communication groups, $p < 0.006$. Moreover, the sensitivity did not significantly differ between the individual searchers and the non-communication groups; indeed, the trend was towards lower, not greater, sensitivity in the non-communication group compared to individual searchers.

A one-way ANOVA of response bias also revealed that there were significant differences among individual searchers, the communication groups, and the non-communication groups, $F(2, 36) = 5.04$, $p = 0.011$. A t-test with bonferroni correction confirmed that non-communicating groups had a more liberal response bias than communicating groups, $p = 0.01$, although no other differences were significant.

The greater sensitivity in the communication groups likely resulted from their explicit division of search labor by color, resulting in each member scrutinizing only half of the display. The fact that sensitivity was significantly less in the non-communication groups suggests that they were not dividing the search labor, which argues against a tacit coordination of search in this experiment.

Dovetailing with the sensitivity analysis is the analysis of bias. The individual searchers showed the predictable conservative bias, and consistent with previous work, this conservative bias strengthened in the communicating groups. However, non-communication groups showed a definitive trend in the opposite direction, a significant shift towards more liberal responses. This liberal response bias, combined with the overall lower sensitivity, provides compelling evidence for the emergence of competition in the non-communication groups. I therefore attribute the fast no-communication search times in Experiments 1 and 2 to an influence of competition.

One final point deserves mention regarding the above interpretation. Search times in the non-communication condition were not only faster than those in the nominal and communicating groups, but sometimes substantially so (i.e., several seconds). How can such a large difference result from a bias attached to a single search decision? First, as our data in this experiment show, it was not just a bias contributing to this difference, but also a change in sensitivity. Second, I speculate the bias was not influencing just a single search decision on a trial. Rather, a liberal bias may affect the amount of evidence that is required to distinguish a target from a distractor at each fixation. A liberal bias may therefore shorten the duration of each fixation (also consistent with competition), which when added up might amount to the several second benefits observed in Experiments 1 and2.

2.6. Experiment 5

In experiments 1-4 I investigated at the level of strategy coordination. I conceptualize this coordination as plan-based, with the strategies being largely pre-determined (or determined very early in a given context), and not a momentary reflection of the environment.

Coordination can also exist at a more micro level, which does capture the moment by moment interactions between group members in response to their environment. Take two-person dancing as example. Surely two experienced dancers know the basic dance steps and are therefore moving according to a sort of script or plan. However, they also need to momentarily adjust their moves according to their partner's moves and the environment, and this is an example of micro-level behavioral coordination.

Micro-level coordination requires the group members' awareness of each other's behavior and the task progress in real time. By using the collaborative visual search system developed by Zelinsky et al. (2005) (Figure 1), the dyadic searchers can see each other's gaze cursors drawn on the search display. This system therefore creates the opportunity for the members of a group to be aware of each other's search behavior and task progress in near real time, thus meeting the basic requirements for studying micro-level behavioral coordination.

The specific focus of this experiment was targeted assistance. We already know that dyadic searchers choose to divide the search display, regardless of whether this division is by space or feature. However, it is not known whether a partner who quickly finishes the search of her region will choose to help the other partner who is still actively searching. Moreover, it is not known whether this assistance would be random or more "targeted" and intelligent. Random assistance would suggest that one partner is not taking into consideration where the other partner has already searched, thereby indicating a relative lack of micro-level coordination. Alternatively, targeted assistance, defined here by a shift in gaze to a yet unsearched region of a partner's search area, would indicate a relatively high level of moment-by-moment micro-coordination between partners. The purpose of this experiment was to evaluate the evidence for such targeted assistance in the context of shared-gaze search.

Design and Participants

Given that the shared gaze system can only accommodate 2-person groups, this experiment was limited to dyads of searchers. Searchers were allowed to communicate with each other using either of two methods in this experiment. In the shared gaze (SG) condition they could see each other's gaze cursors superimposed over the search display in real time, but they could not speak to each other. In the shared voice condition (SV) they could speak to each other via open microphones, but one partner could not see where the other was looking. There were also target present and target absent trials. In this experiment, 1/3 of the trials were target present and 2/3 was target absent. I adopted this ratio because target absent trials have a greater potential of illustrating targeted assistance, due to the greater time spent searching for the target. Experiment 5 therefore had a 2 (SG vs. SV) x 2 (target present or absent) experimental design. Communication condition was a between-subject factor and target presence was a within-subject factor.

There were 5 dyads comprising each communication condition, for a total of 20 participants. All the participants were recruited from the Stony Brook University subject pool, and were screened to have normal or corrected to normal visual acuity and color vision, by self report.

Stimuli and Apparatus

The search stimuli were arrays of 494 red and blue dots, similar in size and type to those used in Experiment 2. However, and different from Experiment 2, in each display there was a small region of one color juxtaposed with a larger region of the other color (see Figure 14 for a representative target present display). Across the experiment, there were an equal number of trials in which the small region was red and blue, and this color of the small region was randomly changed over trials. The small region was also likely to appear in any one of the four display quadrants, also randomly interleaved over trials. There were therefore 8 possible display configurations, based on the 2 color x 4 quadrant of the small region. As in experiments 1 and 2, the non-target dots were circles and the targets were ovals, 50% oriented horizontally and 50% oriented vertically.

The apparatus used in this experiment was the shared gaze system developed by Zelinsky et al. (2005) and Brennan et al. (2007, in press) (Figure 1). Two partners (A and B), each wearing ahead-mounted EyelinkII eyetracker (SR Research), were seated in front of identical 19 inch SVGA computer monitors in separate rooms. Synchronized

Pentium-based computers outputted displays to each monitor, meaning that partners were seeing the same stimulus and were performing the same search task. Partners registered their responses using a gamepad (Microsoft Sidewinder 1.0), and the first partner to make a response terminated the search trial for both. The display and eyetracker computers were connected via an Ethernet hub, enabling the bi-directional exchange of gaze signals between the collaborators. Specifically, the eye position from each partner's eyetracker was sent to the other's screen and displayed as a gaze cursor (a 1.7° visual angle yellow ring). The total time required between obtaining a gaze sample from A and drawing a gaze cursor on B's monitor was estimated to be less than 20 msecs, based on a 500 Hz sampling frequency and a 100 Hz monitor refresh rate. Voice information could also be exchanged between the partners using a bi-directional microphone-speaker system, routed through an audio mixer.

Procedure

Dyadic searchers were asked to look at a central fixation dot and to press a button as part of the eyetracker's drift correction procedure. After the first partner passed drift correction, the fixation point would be turned into a cross. After the second partner passed drift correction the fixation target would be changed back to a dot, was then replaced by the search display. The searchers were instructed to indicate the presence or absence of the target as quickly and accurately as possible by pressing the left or right gamepad triggers, respectively. The first searcher to register a response terminated the trial for both. Accuracy feedback was displayed following the response. The dyadic searchers were instructed to collaborate instead of compete, and the strategy of dividing the labor by color was suggested. This was done in order to increase the chances that one partner would take responsibility for the small region, and the other partner would take responsibility for the large region (and alternate over trials), a condition central to our experimental logic. The whole experiment had 12 practice trials and 60 experimental trials, and it took approximately 1.5 hours, including the set up time for the eyetrackers.

Results and Discussion

The expectation was that, when the partner assigned to the small region finished her search, she would begin to assist the partner still searching the larger display region. The primary question of interest was whether this assistance would be random or targeted. Figure 15 shows the scatterplots of fixation points from all dyads in the 5 seconds period starting 1 second after the search display onset, divided by quadrant. The red points indicate the fixations of searchers who assumed responsibility for the small region; the blue points indicate the fixations from searchers who assumed responsibility for searching the large region. The scatterplots confirm that the dyads actually divided the search display by color, which was consistent with the post-experiment debriefing in which all subjects reported following this color division of labor strategy.

Table 2 provides the error rates and RTs for the dyadic searchers in experiment 5. I did a 2 (SG, SV) x 2 (target present, absent) mixed ANOVA on the error rates and RTs with communication media (SG or SV) a between-subject factor and target presence a within-subject factor. For error rates, only the main effect of target presence was significant, $F(1, 8) = 23.57$, $p < 0.01$. Both SG and SV dyads had high but comparable miss rates. They all demonstrated the usual conservative bias (preferring to respond "target absent"), which may have been exaggerated in this experiment due to 2/3 of the trials actually being target absent. RTs were only analyzed from correct trials, and this analysis again revealed only a significant main effect of target presence, F (1, 8) = 21.31 p <0.01. Predictably, both the SG and SV dyads searched faster when the target was present than absent.

How does one go about quantifying targeted assistance in this task? I define targeted assistance as fixations on a partner's region that tend not to overlap with where that partner has already searched, but this seemingly straightforward definition presupposes an answer to a seemingly simpler question: when is one person trying to assist another? However, answering this simple question is anything but simple when dealing with noisy free viewing behavior, as one partner might occasionally direct gaze into a partner's region without this being an attempt to offer assistance (e.g., a check on the other person's progress). I therefore adopted the following criteria for defining the start of assistance, which rely on (1) the fact that there is a small and large region in each search display, and (2) the assumption that when the person searching the small region finishes, she will attempt to help the person searching the larger region. First, the helper (i.e., the partner responsible for the small region) had to search their own region for at least 5 seconds. Second, the helper would then have to shift gaze to the large region and search this region for at least 1 second without re-entering the small region. Assuming that these two criteria are satisfied, this 1 second "forward window" would be subtracted from the trial clock, and this time would be defined as the moment that assistance started. Using these criteria, on 80% of the trials I could determine a start time for assistance.

Having determined the start time of assistance, I could now compute overlap, which is the defining characteristic of targeted assistance. First I determined the locations of the first 10 fixations made by the helper in the large region. Second, I computed the distances between these 10 fixations and all of the partner's fixations in the large regions since the start of the trial. If any of these distances was less than 1° visual angle (50 pixels), this helper fixation would be counted as overlapping with the partner's fixation. By dividing the number of the overlapping fixations by 10 and multiplying by 100 I obtained the percentage of the helper's fixations that overlapped with his or her partner's fixations. I also calculated a baseline overlap estimate by creating the same number of nominal dyadic pairs by randomly pairing partners from the real dyadic groups. Fixation overlap was computed using the method described above, although the onset of assistance was defined as a fixed 9 seconds after search display onset (which was the average start time for the real dyads). The partners of these nominal groups were therefore searching the same display regions, but would not be offering each other targeted assistance.

The mean overlap rate of the SG dyads was 11.12%, with a SE of 2.81%. The mean overlap rate of the SV dyads was 21.52%, and with a SE of 1.76%. The baseline overlap rate was 18.38%. T-tests confirmed that the overlap in the SG condition was significantly less than overlap in the SV condition, $t = 3.14$, $p = 0.017$, and that the SG overlap rate was also less than the overlap expected by chance, $t = 2.58$, $p = 0.06$. In contrast, the SV dyads' overlap rate was not different from the baseline, $t = 1.78$, $p = 0.14$, and in fact was slightly larger rather than smaller. These analyses strongly suggest that searchers who shared eye gaze can, and did, offer each other targeted assistance. Because one partner knew where the other had already searched, she was able to optimize her assistance by targeting it to the yet unsearched region of the display. Our data also suggest that targeted assistance was not practical or possible for partners limited only to verbal communication. I speculate that the time and effort needed to create and verbalize unambiguous requests for assistance is prohibitive in the context of a search task.

I also addressed the question of whether seeing each other's gaze cursors might affect the moment by moment search behavior more broadly. To do this, I measured the gaze distance between the dyadic searchers throughout each search trial. I analyzed these data using a 2 (SG, SV) x 2 (target present, absent) mixed ANOVA and found a significant interaction between target presence and communication medium, F (1, 8) = 7.51, p = 0.025 (see Figure 16). A post hoc t-test on the target absent data revealed that the average gaze distance between searchers in the SG condition (M = 591 pixels) was shorter than the gaze distance between searchers in the SV condition (M = 661pixels), p = 0.016. No difference in average gaze position was found between searchers in the target present data, t = 0.44, p = 0.67. Why did the gaze difference between searchers differ between SG and SV conditions only in the target absent data? Recall that at the start of a search each partner would focus on their region and there would be no gaze coordination or "dancing". Searchers would likely only start gaze dancing after one partner finished his or her region, but by this time the target would often be detected in target present trials. I speculate that the reason why gaze dancing was limited to the target absent data was because the longer target absent trials simply afforded more opportunity for this behavioral coordination to emerge. To test this hypothesis, I only analyzed the target absent gaze distances starting 10 seconds into each trial (when possible). As is clear from Figure 16, the effect of communication media increased; partners sharing gaze tended to keep a shorter distances between their fixations. As for why shared gaze should affect gaze distance, one possibility is that the efficient use of this medium requires a close monitoring of the partner's gaze cursor, which is best accomplished if the gaze cursor is relatively close to where the other partner is looking. Another possibility is that the knowledge obtained from shared gaze allows for more precisely demarcated divisions of labor, thereby enabling people to confidently search closer together while still avoiding redundancy. Future work will be needed to tease apart these possibilities.

Chapter 3: Conclusion

This thesis investigated coordination in group visual search. It focused on two levels of coordination: intermediate level strategic coordination, and micro-level behavioral coordination.

Turning first to strategy coordination, I addressed the influences of task characteristics, group size, and communication media on the formation and efficient use of collaborative search strategies.

Regarding the effects of task characteristics, Brennan et al. (2007, in press) (see also Zelinsky et al., 2005) already demonstrated that dyads naturally elect to divide the search display by space, and I replicated this finding in Experiment 1 in the context of a difficult oval discrimination task. However, I hypothesized that the subjects in these studies might have used this spatial division of labor search strategy, not necessarily because it was the preferred way of dividing the search labor, but rather because there were no visual grouping cues available to divide the labor in any other way. I tested this hypothesis in Experiment 2, which gave subjects the option of dividing labor by space (left/right or up/down) or by feature (blue, black, red, or green regions). When given this option, I found that most subjects chose to divide the search labor by color (particularly for 4-person groups); a fact that was confirmed by correlating individual subject responses with targets located in different colored regions. I speculate that low level visual grouping processes easily separate these displays into well defined color regions, making a color division of labor a more natural choice than a division of labor by space, which may be more ambiguous and effortful.

The search scene is not the only divisible component of a visual search task, the target templates might also be divided among the group members, especially if the number of targets was far beyond an individual searcher's short-term memory capacity. I explored this possibility in Experiment 3, and again found clear evidence for a division of search labor, in this case by target. In fact, this target division of labor produced our clearest evidence for a collaborative benefit, as without it subjects were largely unable to perform the task (as indicated by an increase in error rates). Combined with Experiments 1-2, I have therefore demonstrated, for the first time, collaborative divisions of search labor by feature and target, in addition to space. Note also that a target division of labor is in a sense qualitatively different from the other two types of labor divisions. Whereas spatial and feature divisions of labor are not mandatory in the sense that any individual member of the group is still free to search the entire display independently, a target division of labor requires group coordination, as any one member of the group does not know all of the targets. This is particularly true for target absent trials, where input from all of the group members would be essential in knowing when to stop a search. This demand for coordination was expressed in our data by an inordinately large effect of target presence/absence.

Regarding the effect of group size on collaborative search, it is clear from Experiments 1-3 that 4-person groups searched faster than 2-person groups. I have therefore determined in this study that strategic divisions of labor are not limited to dyads in a collaborative search task. Less clear from our data is whether collaborative search was more efficient than the same number of subjects searching independently. In general it seems that the benefits gained from a collaborative division of labor were offset by the costs of coordinating this group behavior, resulting in either no net difference or a small search cost. It is also the case that effects of group size interacted with task characteristics, communication condition, and even target presence or absence, making it difficult to discern clear and generalizable patterns in these exploratory data. Future work will be needed to better understand how group size affects collaborative search behavior.

The effects of communication in this study are tied to the previously summarized evidence for collaborative strategies - without communication it would be impossible to form and efficiently implement a strategic coordination of behavior. There are two points regarding this coordination, however, that deserves emphasis. First, and unexpectedly, I found that search in the non-communication conditions from Experiments 1-2 was generally more efficient than search in the communication conditions. In Experiment 4 I applied a signal detection analysis to this mysterious data pattern and found that non-communication search was accompanied by a decrease in perceptual sensitivity (d') and an increase in liberal response bias (lnβ) relative to the communication condition. I therefore speculate that group members in the non-communication conditions had a tendency to compete with each other, and that it was this expression of competition that resulted in the faster search times. A corollary to this argument is that, by giving group members the opportunity to communicate, such competitive influences can be suppressed. Second, and as already mentioned, the costs of verbal communication typically outweighed the benefits, resulting in less efficient search relative to nominal group performance. One should therefore consider carefully a plan for reducing these communication costs, and the potential for competition, before applying a collaborative search method to a task. One possibility might be to limit the potential for communication to certain times during a task, perhaps at the very start and at the very end. This would enable group members to devise collaborative strategies, and to discuss the success or failure of these strategies, while minimizing task-irrelevant chatter that might be introducing unnecessary communication costs. Future work will be needed to explore this possibility.

Rather than operating at the level of strategies or agreed upon search plans, micro-level behavioral coordination refers to the moment by moment and largely unconscious response of one partner to the behavior of another. Given that micro-level coordination requires a real-time access to a partner’s behavior, it is most often studied in the context of face-to-face communication (Tang, 1991). However, face-to-face communication is probably not the most efficient medium for collaborative search, as it requires the error prone extrapolation of where someone is looking based on their direction of gaze. To remedy this problem, we developed a shared gaze methodology, which uses electronically mediated gaze cursors to show one partner where the other is looking. Using this technique, we found that dyads limited to only shared gaze (i.e., no speech) searched more efficiently than dyads that could speak but not share gaze (Brennan et al., 2007, in press; Zelinsky et al., 2005). We attributed this shared gaze benefit to micro-level coordination made possible by the shared gaze medium; one person should see moment by moment where her partner was looking, and adjust her behavior accordingly.

Although plausible, the evidence for micro-level behavioral coordination in the Brennan et al. (2007, in press), Zelinsky et al. (2005) study was indirect. Shared gaze benefits might have emerged due to early strategic coordination and the absence of speech costs, rather than micro-level behavioral coordination throughout the search task. In Experiment 5 I report the first direct evidence for micro-level behavioral coordination. Under shared gaze conditions, I found that searchers who finished inspecting a small region, shifted their gaze to their partner's larger search region in an effort to lend assistance. However, I found that this assistance was intelligent or targeted, meaning that they looked to a part of this region that their partner had not yet searched. Searchers limited to speech also attempted to help each other, but these efforts were less targeted, as indicated by a higher overlap in fixations.

In addition to this evidence for micro-level behavioral coordination in the form of targeted assistance, I also observed an even more subtle form of coordination in the form of "gaze dancing". Partners in the shared gaze condition tended to keep a smaller distance between their points of regard compared to searchers in the shared voice condition, a pattern that I attribute to the moment-by-moment monitoring of each other's gaze cursors. However, despite this evidence for micro-level coordination, I also note that it can be largely overridden by mid-level strategic coordination. I say this because at the start of each Experiment 5 trial, partners tended to restrict themselves to their own search areas (per their strategic agreement), and did not engage in any obvious micro-level coordination.

References

Acker, S., and Levitt, S. (1987). Designing videoconference facilities for improved eye contact. *Journal of Broadcasting & Electronic Media*, *31(2)*, 181–191.

Alvarez, G. A., and Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science*, *15*, 106–111.

Argyle, M., and Graham, J. (1977). The central europe experiment - looking at persons and looking at things. *Journal of Environmental Psychology and Nonverbal Behaviour*, *1*, 6–16.

Bailey, H. H. (1970, February). *Target detection through visual recognition: A quantitative model* (Tech. Rep. No. RM-6158-PR). Santa Monica: Rand Memorandum.

Basden, B. H., Basden, D. R., Bryner, S., and Thomas, R. L. (1997). A comparison of group and individual remembering: Does collaboration disrupt retrieval strategies? *Journal of Experimental Psychology: Learning, Memory and Cognition*, *23(5)*, 1176–1189.

Bouchard, T. J., Drauden, G., and Barsalou, J. (1974). Comparison of individual, subgroup, and total group methods of problem-solving. *Journal of Applied Psychology*, *59(2)*, 226–227.

Bouchard, T. J., and Hare, M. (1970). Size, performance, and potential in brainstorming groups. *Journal of Applied Psychology*, *54(1)*, 51–55.

Brennan, S. E. (2005). How conversation is shaped by visual and spoken evidence. J. Trueswell and M. Tanenhaus (Eds.), *Approaches to studying world-situated language use: Bridging the language-as-product and language-action traditions* (pp. 95–129). Cambridge, MA: MIT Press.

Brennan, S. E., Chen, X., Dickinson, C., Neider, M., and Zelinsky, G. (2007, in press). Coordinating cognition: The costs and benefits of shared gaze during collaborative search. *Cognition*.

Buxton, W. A. S., and Moran, T. (1990). Europarc's integrated interactive intermedia facility (iiif): Early experience. S. Gibbs and A. A. Verrijn-Stuart (Eds.), *Multi-user interfaces and applications* (pp. 11–34). Amsterdam: Elsevier.

Carlson, L. A., and Logan, G. D. (2001). Using spatial terms to select an object. *Memory & Cognition*, *29*, 883–892.

Carter, L. F., Haythorn, W. W., and Howell, M. (1950). A further investigation of the criteria of leadership. *Journal of Abnormal and Social Psychology*, *45*, 350–358.

Chapanis, A., Ochsman, R. B., Parrish, R. N., and Weeks, G. D. (1972). Studies in interactive communication: 1. effects of 4 communication modes on behavior of teams during cooperative problem-solving. *Human Factors*, *14(6)*, 487–509.

Chatziastros, A., and Bulthoff, H. (2006). Spatial partitioning during visual search of a dyad [abstract]. *Journal of Vision*, *6(6)*, 521.

Clark, H. H. (1996). *Using language*. Cambridge, England: Cambridge University Press.

Clark, H. H., and Brennan, S. E. (1991). Grounding in communication. L. B. Resnick, J. Levine, and S. D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 127–149). Washington, DC: APA.

Colston, H., and Schiano, D. (1995). Looking and lingering as conversational cues in vmc. *In proceedings of chi'95.* Denver, Colorado.

Daft, R. L., Lengel, R. H., and Trevino, L. K. (1987). Message equivocality, media selection, and manager performance: Implications for information systems. *MIS Quarterly*, *11(3)*, 355–366.

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

Davis, J. H. (1982). Social interaction as a combinatorial process in group decisions. H. Brandstatter, J. H. Davis, and G. Stocker-Kreichgauer (Eds.), *Group decision making* (pp. 27–58). London: Academic.

Dickinson, C., and Zelinsky, G. (2005). Marking rejected distractors: A gaze-contingent technique for measuring memory during search. *Psychonomic Bulletin and Review*, *12(6)*, 1120–1126.

Diehl, M., and Stroebe, W. (1987). Productivity loss in brainstorming groups : Toward the solution of a riddle. *Journal of personality and social psychology*, *53(3)*, 497–509.

Diehl, M., and Stroebe, W. (1991). Productivity loss in idea-generating groups - tracking down the blocking effect. *Journal of Personality and Social Psychology*, *61(3)*, 392–403.

Drapier, M., Ducoing, A. M., and Thierry, B. (1999). An experimental study of collective performance at a foraging task in tonkean macaques. *Behaviour*, *136*, 99–117.

Egido, C. (1989). Teleconferencing as a technology to support cooperative work: Its possibilities and limitations. J. Galegher and R. Kraut (Eds.), *Intellectual teamwork: Social and technological foundations of group work* (pp. 351–372). Hillsdale, NJ: Lawrence Earlbaum Associates.

Fern, E. F. (1982). The use of focus groups for idea generation: The effects of group size, acquaintanceship and moderator on response quantity and quality. *Journal of Marketing Research*, *19*, 1–13.

Foushee, M. C. (1984). Dyads and triads at 35,000 feet: Factors affecting group process and aircrew performance. *Amercian Psychologist*, *39*, 885–893.

Galbraith, J. (1977). *Organization Design,* Addison-Wesley, Reading, MA.

Gale, C., and Monk, A. F. (2000, April). Where am i looking? the accuracy of video-mediated gaze awareness. *Perception & Psychophysics*, *62(3)*, 586–595.

Gergle, D., Kraut, R. E., and Fussell, S. R. (2004). Language efficiency and visual technology: Minimizing collaborative effort with visual information. *Journal of Language and Social Psychology*, *23(4)*, 491–517.

Hackman, J. R., and Morris, C. G. (1975). Group tasks, group interaction process, and group performance effectiveness: A review and proposed integration. L. Berkowitz (Ed.), *Advances in experimental social psychology, 8* (pp. 45–99). New York, NY: Academic Press.

Hackman, J. R., and Vidmar, N. (1970). Effects of size and task type on group performance and member reactions. *Sociometry*, *33*, 37–55.

Hill, G. W. (1982, May). Group versus individual performance: Are n + 1 heads better than one? *Psychological Bulletin*, *91(3)*, 517–39.

Hogarth, R. M. (1978). A note on aggregating opinions. *Organizational Behavior and Human Performance*, *21*, 40–46.

Hollingshead, A. B. (1998a). Communication, learning, and retrieval in transactive memory systems. *Journal of Experimental Social Psychology*, *34(5)*, 423–442.

Hollingshead, A. B. (1998b). Retrieval processes in transactive memory systems. *Journal of Personality and Social Psychology*, *74(3)*, 659–671.

Horowitz, T. S., and Wolfe, J. M. (1998). Visual search has no memory. *Nature*, *357*, 575–577.

Irwin, D. E., and Zelinsky, G. J. (2002). Eye movements and scene perception: Memory for things observed. *Perception & Psychophysics*, *64(6)*, 882–895.

Ishii, H., and Kobayashi, M. (1992, May). Clearboard: A seamless medium for shared drawing and conversation with eye contact. *Proceedings of chi '92 human factors in computing systems* (pp. 525–532).

Ishii, H., Kobayashi, M., and Grudin, J. (1993, October). Integration of interpersonal space and shared workspace: Clearboard design and experiments. *ACM Transactions on information systems*, *11(4)*, 349–375.

Kanki, B. G., and Foushee, H. (1989). Communication as group process mediator of aircrew performance. *Aviation, space, and environmental Medicine*, *60*, 402–410.

Koch, C., and Ullman, S. (1985). Shift in selective visual attention: towards the underlying neural circuitry. *Human Neurobiology*, *4*, 219–227.

Koopman, B. (1956). The theory of search 2: Target detection. *Operations Research*, *4*, 503–531.

Kravitz, D. A., and Martin, B. (1986). Ringelmann rediscovered : The original article. *Journal of Personality and Social Psychology*, *50(5)*, 936–941.

Lanzetta, J. T., and Roby, T. B. (1960). The relationship between certain group process variables and group problem solving efficiency. *Journal of Social Psychology*, *52(1)*, 135–148.

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

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

Levinson, S. (1996). Frames of reference and molyneux's question: cross-linguistic evidence. P. Bloom, M. Peterson, L. Nadel, and M. Garrett (Eds.), *Language and space* (pp. 109–169). Cambridge, MA: MIT Press.

Lewis, A. C., Sadosky, T. L., and Connolly, T. (1975). The effectiveness of group brainstorming in engineering problem solving. *IEEE Transactions on Engineering Management*, *22*, 119–124.

Logan, G. D. (1995). Linguistic and conceptual control of visual spatial attention. *Cognitive Psychology*, *28*, 103–174.

Logan, G. D., and Compton, B. J. (1996). Distance and distraction effects in the apprehension of spatial relations. *Journal of Experimental Psychology: Human Perception and Performance*, *22*, 159–172.

Logan, G. D., and Sadler, D. D. (1996). A computational analysis of the apprehension of spatial relations. P. Bloom, M. A. Peterson, L. Nadel, and M. F. Garrett (Eds.), *Language and space* (pp. 493–529). Cambridge, MS: MIT Press.

Lorge, I., and Solomon, H. (1955). Two models of group behavior in the solution of eureka-type problems. *Psychometrika*, *20*, 139–148.

Luck, S. J., and Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*, 279–281.

Malcolmson, K. A., Reynolds, M. G., and Smilek, D. (2007). Collaboration during visual search. *Psychonomic Bulletin & Review*, *14(4)*, 704–709.

Malone, T. W., and Crowston, K. (1994). The interdisciplinary study of coordination. *ACM Computing Surveys*, *26(1)*, 87–120.

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

McGrath, J. E., and Hollingshead, A. B. (1993). Putting the group back in group support systems: some theoretical issues about dynamic processes in groups with technological enhancements. L. M. Jessup and J. S. Valacich (Eds.), *Group support systems: New perspectives* (pp. 78–96). New York: Macmillan.

Meudell, P. R., Hitch, G. J., and Boyle, M. (1995). Collaboration in recall: Do pairs of people cross cue each other to produce new memories? *Quarterly Journal of Experimental Psychology*, *48*, 141–152.

Monk, A. F., and Gale, C. (2002). A look is worth a thousand words: Full gaze awareness in video-mediated conversation. *Discourse Processes*, *33(3)*, 257–278.

Muller, G. F. (1992). Psychological and communicational influences on coordination effectiveness. *Basic and Applied Social Psychology*, *13(3)*, 337–350.

Nardi, B., Schwarz, H., Kuchinsky, A., Leichner, R., Whittaker, S., and Sclabassi, R. (1993). Turning away from talking heads: The use of video-as-data in neurosurgery. *Proceedings of interchi '93* (pp. 327–334). NY: ACM Press.

Neider, M., Chen, X., Dickinson, C., Brennan, S., and Zelinsky, G. (2005). *Sharing eyegaze is better than speaking in a time-critical consensus task.* Toronto, Ontario, Canada.

Nickerson, R. S. (1984). Retrieval inhibition from part-set cueing - a persisting enigma in memory research. *Memory & Cognition*, *12(6)*, 531–552.

Osborn, A. F. (1957). *Applied imagination: Principles and procedures of creative thinking* (2nd ed.). New York, NY: Scribner.

Pavitt, C. (2003). Colloquy: Do interacting groups perform better than aggregates of individuals? why we have to be reductionists about group memory. *Human Communication Research*, *29(4)*, 592–599.

Propp, K. M. (2003). In search of the assembly bonus effect: Continued exploration of communication's role in group memory. *Human Communication Research*, *29(4)*, 600–606.

Rao, R., Zelinsky, G., Hayhoe, M., and Ballard, D. (2002). Eye movements in iconic visual search. *Vision Research*, *42*, 1447–1463.

Rotman, S. R. (1989). Modeling human search and target acquisition performance: 2. simulating multiple observers in dynamic scenarios. *Optical Engineering*, *28(11)*, 1223–1226.

Salas, E., Bowers, C. A., and Cannon-Bowers, J. A. (1995). Military team research: 10 years of progress. *Military Psychology*, *7*, 55–75.

Shaw, M. (1981). *Group dynamics: The psychology of small groups*. New York: McGraw Hill.

Steiner, I. D. (1972). *Group process and productivity.* New York: Academic Press.

Steven, M., and Campion, M. A. (1994). The knowledge, skill and ability requirements for teamwork. *Journal of Management*, *20(2)*, 34–48.

Straus, S. G. (1999). Testing a typology of tasks - an empirical validation of mcgrath's (1984) group task circumplex. *Small Group Research*, *30(2)*, 166–187.

Suh, K. S. (1999). Impact of communication medium on task performance and satisfaction: An examination of medium richness theory. *Information and Management*, *35*, 295–312.

Talmy, L. (1983). How languages structure space. H. Pick and L. Acredolo (Eds.), *Spatial orientation: theory, research, and application.* New York: Plenum Press.

Tang, J. C. (1991). Findings from observational studies of collaborative work. *International Journal of Man-Machine Studies*, *34(2)*, 143–160.

Taylor, D., and Faust, W. (1952). Twenty questions: Efficiency in problem solving as a function of size of group. *Journal of Experimental Psychology*, *44*, 360–368.

Taylor, D. W., Berry, P. C., and Block, C. H. (1958). Does group participation when using brainstorming facilitate or inhibit creative thinking? *Administrative Science Quarterly*, *3(1)*, 23–47.

Treisman, A. M., and Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology*, *12*, 97–136.

Tschan, F., and Cranach, M. von. (1996). Group task structure, processes, and outcome. M. A. West (Ed.), *Handbook of work group psychology* (pp. 92–121). Chichester, England: John Wiley.

Ullman, S. (1984). Visual routines. *Cognition*, *18*, 97–159.

Valacich, J. S., Dennis, A. R., and Connolly, T. (1994). Idea generation in computer-based groups - a new ending to an old story. *Organizational Behavior and Human Decision Processes*, *57(3)*, 448–467.

Velichkovsky, B. M. (1995). Communicating attention: Gaze position transfer in cooperative problem solving. *Pragmatics and Cognition*, *3*, 199–222.

Vertegaal, R. (1999). The gaze groupware system: Mediating joint attention in multiparty communication and collaboration. *Proceedings of chi.* Pittsburgh.

Wegner, D. M. (1986). Transactive memory: A contemporary analysis of the group mind. B. Mullen and G. R. Goethals (Eds.), *Theories of group behavior* (pp. 185–205). New York: Springer-Verlag.

Wegner, D. M. (1991). Transactive memory in close relationships. *Journal of Personality and Social Psychology*, *61(6)*, 923–929.

Weingart, L. (1992). Impact of group goals, task component complexity, effort, and planning on group performance. *Journal of applied psychology*, *77*, 682–693.

Weldon, E., Jehn, K. A., and Pradhan, P. (1991). Processes that mediate the relationship between a group goal and improved group performance. *Journal of personality and social psychology*, *61*, 555–569.

Weldon, M. S., and Bellinger, K. D. (1997). Collective memory: Collaborative and individual processes in remembering. *Journal of Experimental Psychology - Learning Memory and Cognition*, *23(5)*, 1160–1175.

Williges, R. C., Johnston, W. A., and Briggs, G. E. (1966). Role of verbal communication in teamwork. *Journal of Applied Psychology*, *50*, 473–478.

Wittenbaum, G. M., Merry, C. J., and Stasser, G. (1996). Tacit coordination in anticipation of small group task completion. *Journal of Experimental Social Psychology*, *32*, 129–152.

Wittenbaum, G. M., Vaughan, S. I., and Stasser, G. (1998). Coordination in task-performing groups. R. Tindale and L. Heath (Eds.), *Theory and research on small groups. social psychological applications to social issues, vol. 4* (pp. 177–204). New York: Plenum Press.

Wolfe, J. M. (1994). Guided search 2.0: A revised model of visual search. *Psychonomic Bulletin & Review*, *1*, 202–238.

Wolfe, J. M. (1998). Visual search. H. Pashler (Ed.), *Attention* (pp. 13–73). Psychology Press.

Zelinsky, G., Dickinson, C., Chen, X., Neider, M., and Brennan, S. (2005). Collaborative search using shared eye gaze. *Journal of Vision*, *5(8)*, 700a.

Zelinsky, G. J. (1996). Using eye saccades to assess the selectivity of search movements. *Vision Research*, *36*, 2177–2187.

Appendix: Transcripts of the communication groups about division of labor strategies

<u>Exp 1: 2-person groups</u>
Group 1
A: You know we can do it... can do it like top half bottom half.
B: OK, I'll take top.
A: OK, I'll take bottom.

Group 2
A: You look at the left side; I'll look at the right side.
B: OK.
A: You heard it?
B: Yeah.

Group 3
A: You want to start on one side and I'll start on the other.
B: Sure, I'll start on the left.

Group 4
A: Do you want to look on the right and I'll look on the left?
B: Sure.
A: OK.

<u>Exp 1: 4-person groups</u>
Group 1
A: You know what? How about we separate this?
B: I said that. Everyone take a quadrant.
A: OK, I'll take top-right.
C: I'll take top-left.
B: I'll take bottom-right.
D: OK, I'll take bottom-left.

Group 2
A: I'll take the top-left corner.
B: OK, I'll take the top-right corner.
C: I'll take bottom-left.
D: I'll take bottom-right.

Group 3
A: We should divide it into four.
C: So, A and B you want to take the top half? And C and D do you want to take the bottom half?

A: Sure.

Group 4
A: Why not we just look in different corners.
B: OK, I'll take top-left.
A: Top-right.
C: Bottom-right.

Group 5
A: We should... like... split up by corners... Is that exactly the same scene? You want to split the corners?
B: OK.
C: I am alright.
(But they gave up and there were no further division of labor discussion.)

Exp 2: 2-person groups
Group 1
A: How about this, I look at the top and you look at the bottom?
B: OK.

Group 2
A: You want to split it in half?
B: OK, I'll take the right side.
A: OK, I'll take the left side.

Group 3
A: You know what we should do? You should cover two colors, I'll cover two colors. That's an easier way to do that.
B: I got it. Blue and green.
A: OK, I'll take red and black.

Group 4
A: How about I look at the bottom and you look at the top? So?
B: OK.

Group 5
A: We should help each other. You always look at the blue and the black and I'll always look at the green and red.
B: OK.
A: OK.

Exp 2: 4-person groups
Group 1
A: Do you guys want to divide up the color?
B: Yeah.
A: I'll do black.

B: I'll do blue.
C: Red.
D: what's the other color?
A, B: Green.
D: I'll do green.

Group 2
A: I think it would be easier if we divide the colors?
B: Yeah.
A: So, I'll be green.
B: Yeah, I'll be red.
C: Alright, I'll be blue.
D: I'll be black.

Group 3
A: You guys want to assign colors?
B: Yeah, let's do that really quickly. It doesn't matter.
A: I'll get black.
B: I'll get red.
C: I'll get blue.
D: OK, I guess I'll get green.

Group 4
A: There are four color. Why don't we all just take a different color?
B: Yeah, that's what I said. What do you guys want to do?
A: I'll take blue.
C: I'll take green.
B: I'll take black.
D: Red.

Group 5
A: Should we each pick a color?
B: Yeah.
A: which ones?
B: I'll take black.
C: I pick red.
D: Blue.
A: I'll take green.

Group 6
A: Wait, maybe we should all take a color?
B: Yeah. I got black.
C: I'll take green.
D: I'll take blue.
A: I'll take red.

Exp 3: 2-person groups

Group 1
For 4 targets:
A: I'll take the two on the left, you take two on right.
B: OK.
For 8 targets:
B: I'll take the top row, you get bottom.
A: OK.

Group 2
For 8 targets:
A: I think, maybe, we each remember four.
B: I'll take the bottom four.
A: OK, I'll get the top four.
For 4 targets:
B: I'll just do the bottom two.
A: OK.

Group 3
For 8-targets:
A: You want to try different rows?
B: I'll do the bottom and you do the top?
A: Alright. Try that.

Group 4
For 4-targets:
A: Why don't you look for what's on the top and I'll for what's on the bottom?
B: You mean top two and bottom two?
A: Yeah, you do the top two and I do the bottom two.
B: OK.

Group 5
For 8-targets:
A: How about I take the top row and you take the bottom row?
B: OK, sounds good.
A: OK.
For 4-targets:
A: You want to take the bottom again?
B: Yeah.

Group 6
For 8-targets:
A: How about I do the top line and you do the bottom line?
B: You want to do that?
A: Yeah.
B: Alright.

Exp 3: 4-person groups

Group 1
For 8-targets:
A: Let's split it up. Half of us will get the top, half of us will get the bottom.
B: We'll get the top, you guys get the bottom.
C: So you get the top?
A: We got the top, you guys get the bottom.

Group 2
For 8-targets:
A: You guys want to take the top half and we'll take the bottom half, or some way like that?
B, C, D: Yeah.
B: Do you want to divide it even more? Like, I can take the top two left.
A: Yeah. And the other guy over there can take the top two right.
C: Alright.
A: And I'll take the bottom left and you take the bottom right?
D: Alright.

Group 3
For 4-targets:
A: I figure out we should all look in certain corner. There are four corners of the screen and four of us.
B: I'll do the bottom-right.
A: I'll do the top-left.
C: I'll do the top-right.
For 8-targets:
A: We'll take the same corner. Same things as before except now you are looking at two of them.
B: Can we do rows? I like rows better.
C: No.
B: Then what do you mean corners?
A: left corner, like left top.
B: OK.
C: Same as before.

Group 4
For 4 targets:
A: I'll look for top-right.
B: OK, I'll look for the one in the bottom-right.
C: I'll take top-left.
D: OK.
For 8 targets:
C: There are eight now.
A: I'll take top right two.

B: Bottom right two.
D: Two bottom left two.

Group 5
For 8-targets:
C: Let's do it in order of letter.
B: OK.
C: So A, you memorize the first two, B, you memorize the second two.
B: OK.
D: So we memorize it by pairs, right?
A, B, C: yeah.
For 4-targets:
A: We should just memorize one.
B: Yeah, in order again, right?
C, D: Yeah.

Group 6
For 8-targets:
A; So, there are four of us, so why don't we each memorize two?
B: Yeah.
A; So I'll memorize the top-right.
C: I'll memorize top-left.
B: Bottom-left.
D: Bottom-right.
For 4-targets:
A: Well, let's just keep the same as last time.
B, C, D: alright.

Table 1: Average hits, false alarms, sensitivity (d'), and response bias (lnβ) for the individual searchers, non-communication groups and communication groups.

Conditions	Hits	False Alarms	Sensitivity(d')	Response Bias(lnβ)
Individual	0.561	0.163	1.462	1.152
Non-communication	0.561	0.204	1.098	0.531
Communication	0.595	0.085	1.850	1.424

Table 2: The mean and SE of error rates and RTs in experiment 5. The RTs are in msecs.

Communication	Target	Error mean(se)	RTs mean(se)
SG	present	32.7%(3.5%)	12848.71(2867.58)
SG	absent	0.5%(0.5%)	30256.68(7042.11)
SV	present	28.9%(11.8%)	16373.81(2022.50)
SV	absent	1.0%(0.6%)	28872.49(6426.86)

Figure 1: The shared gaze collaborative search system used in Brennan et al. (2007, in press) and Zelinsky et al. (2005). Searchers were seated in separate rooms but saw the same stimulus and performed the same search task. The eye movements from each searcher's eyetracker could be displayed as a gaze cursor on the other's monitor.

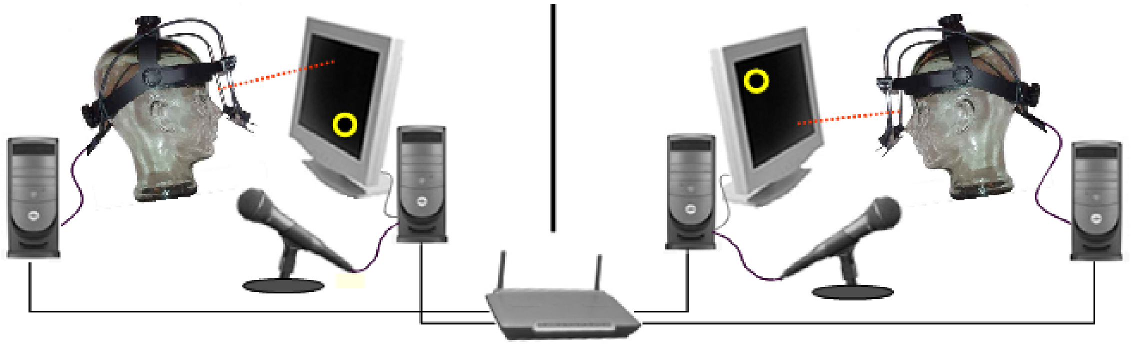

Figure 2: A sample of the search stimuli used in experiment 1.

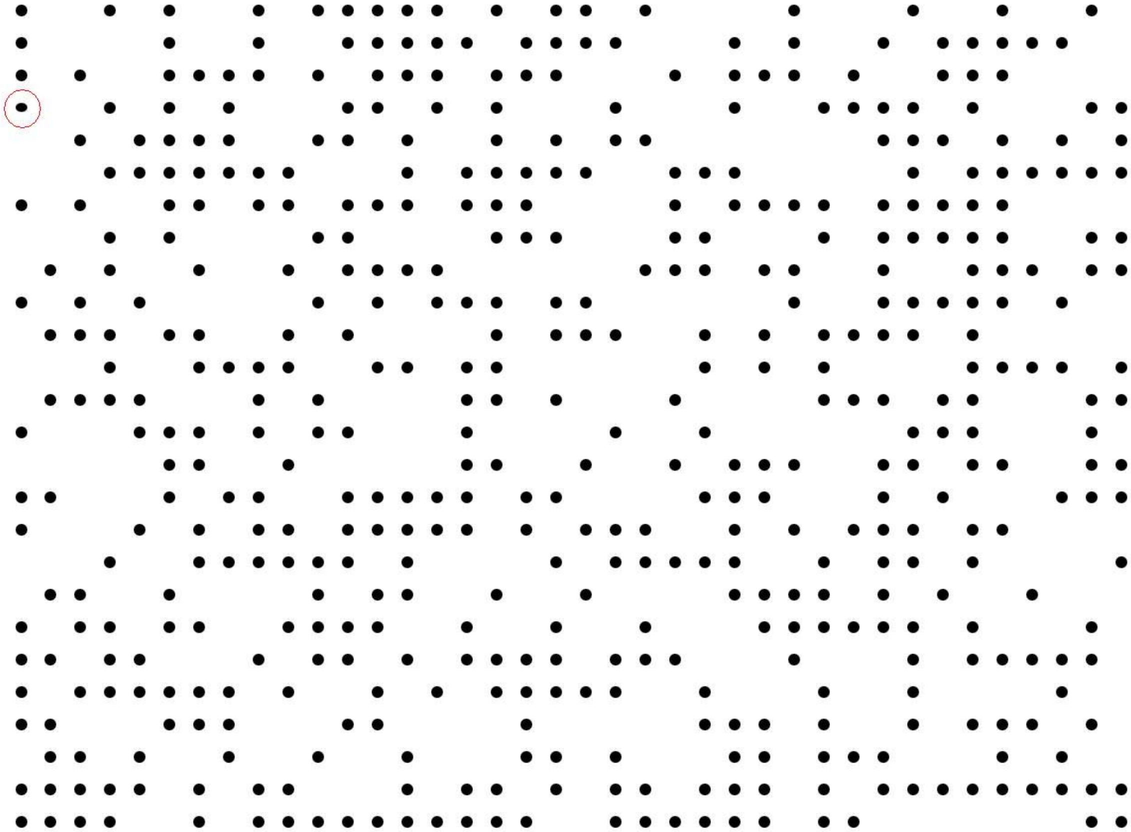

Figure 3: The physical arrangement of the collaborative search system used in Experiments 1-3. One source of video signal was split into 4 monitors. And the 4 gamepads were connected to the same computer also. Participants were either seated back-to-back or separated by a barrier. This system enable up to 4 persons to perform search task collaboratively.

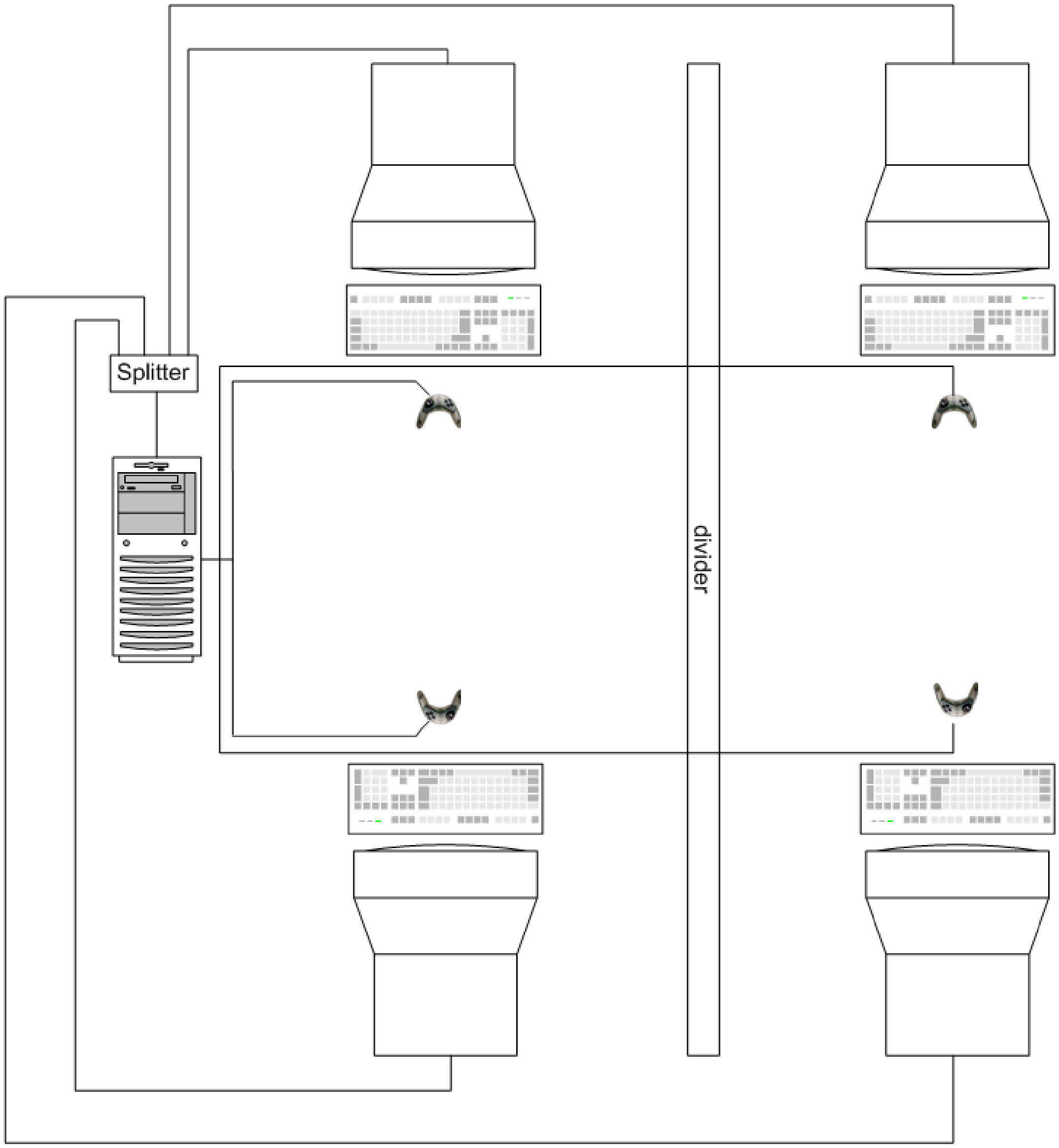

Figure 4: A: The locations the targets found by the members of the three 2-person groups who split the search scene by bottom-top. The white bars represent the members who were assigned to the bottom side of the scene, and the black bars represent the members who were assigned to the top side of the scene. B: The locations of the targets found by the members of the two 2-person groups who split the search scene by right-left. The white bars represent the members who were assigned to the left side of the scene, and the black bars represent the members who were assigned to the right side of the scene.

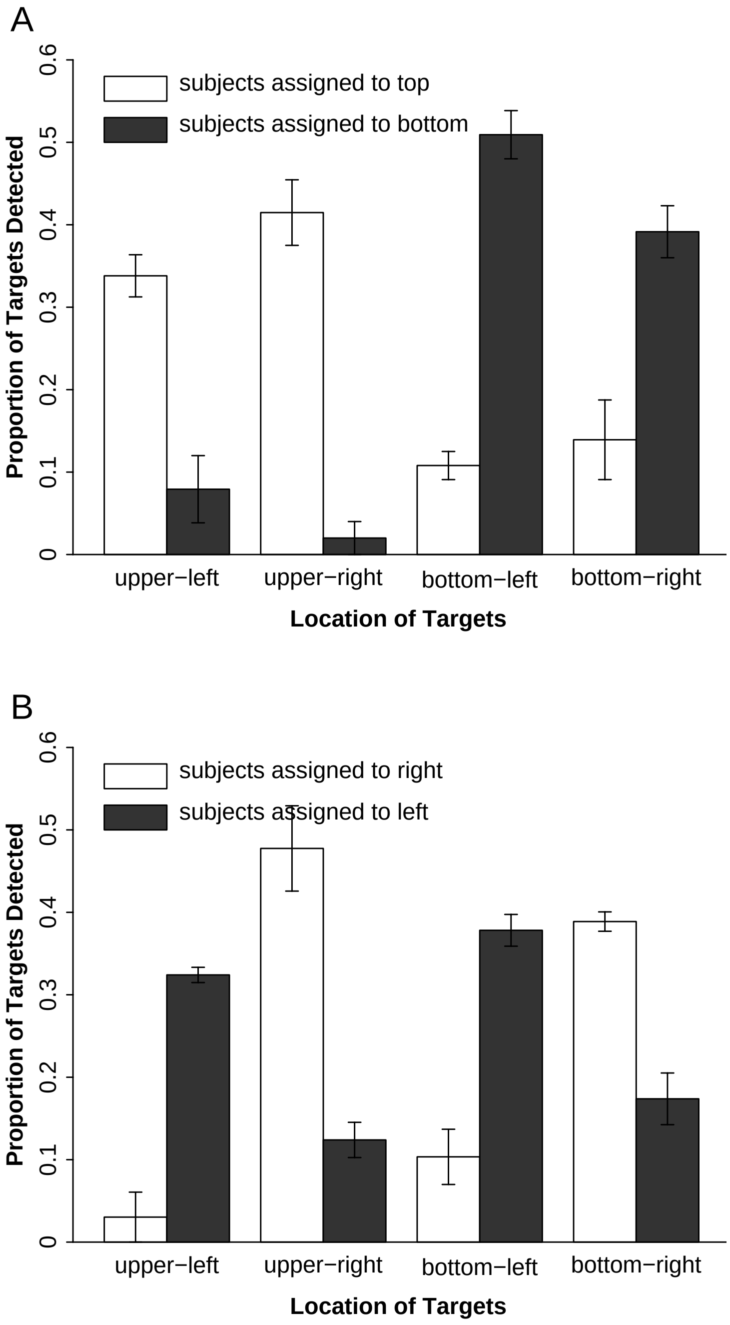

Figure 5: The black dot dash lines represent the RTs of experiment 1: the black dots search experiment. And the solid lines represent the RTs of experiment 2: the color dots search experiment. The error bars represent one standard error.

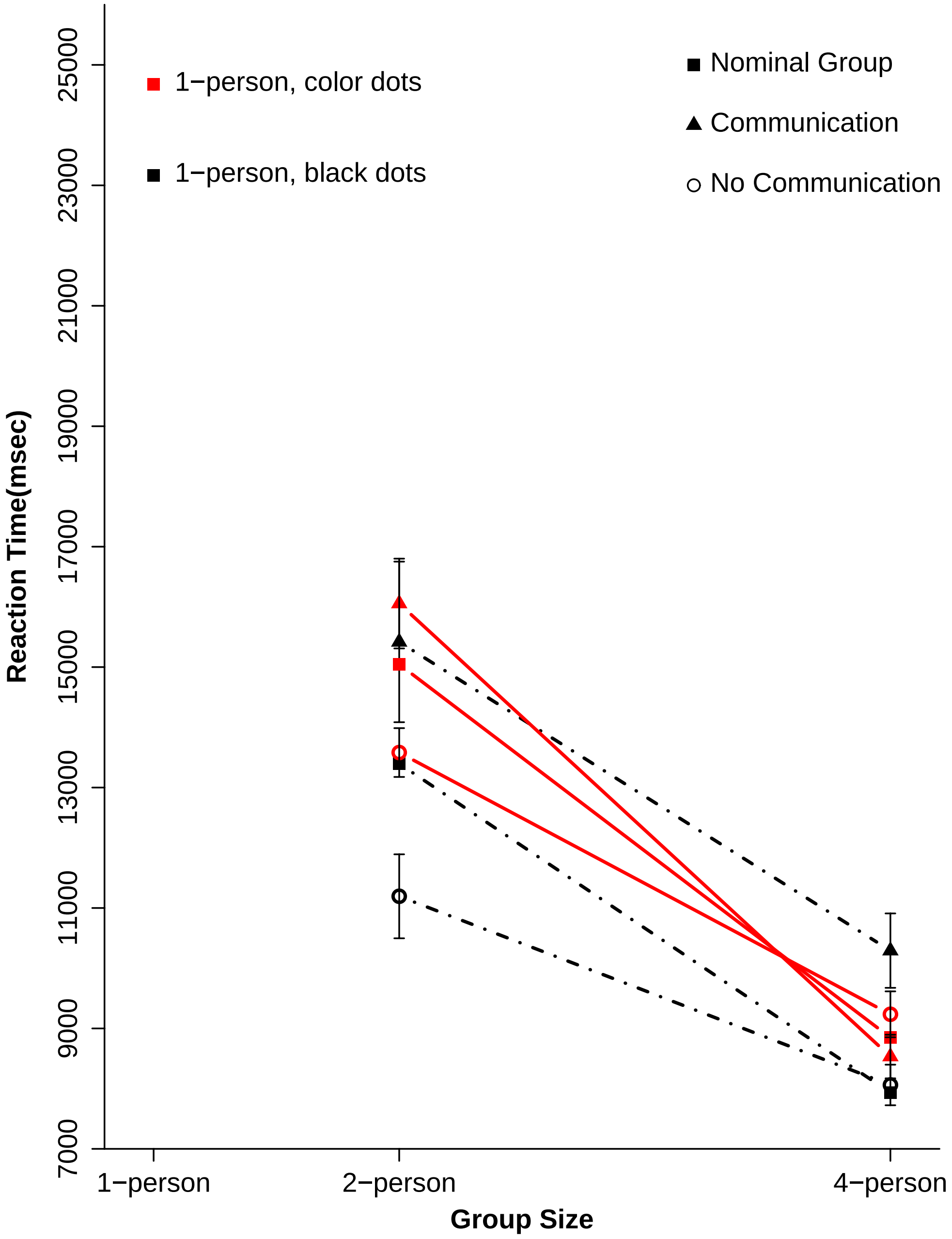

Figure 6: The locations of the targets found by the members of the three 4-person groups who split the search scene by quadrant. The different type bars represent the members who were assigned to different quadrant.

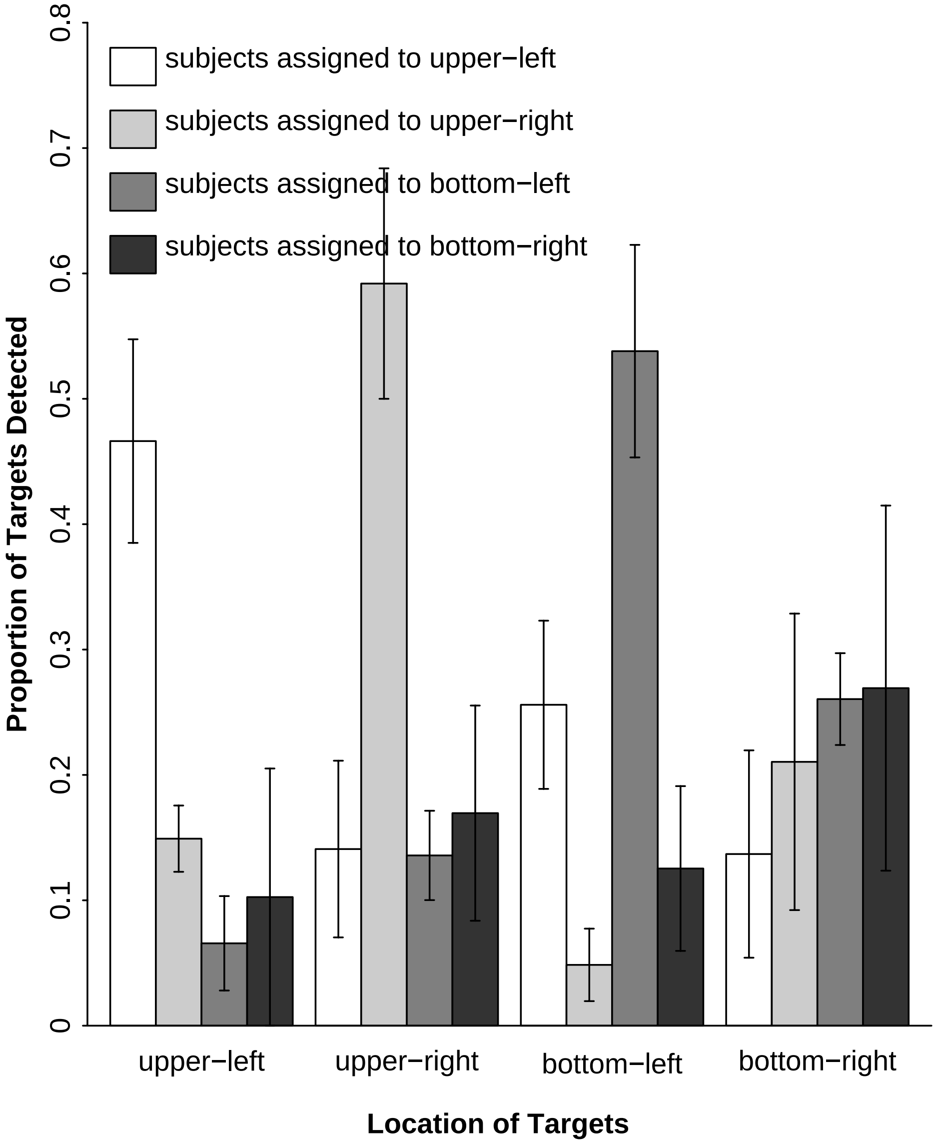

Figure 7: A sample of the search stimuli used in experiment 2.

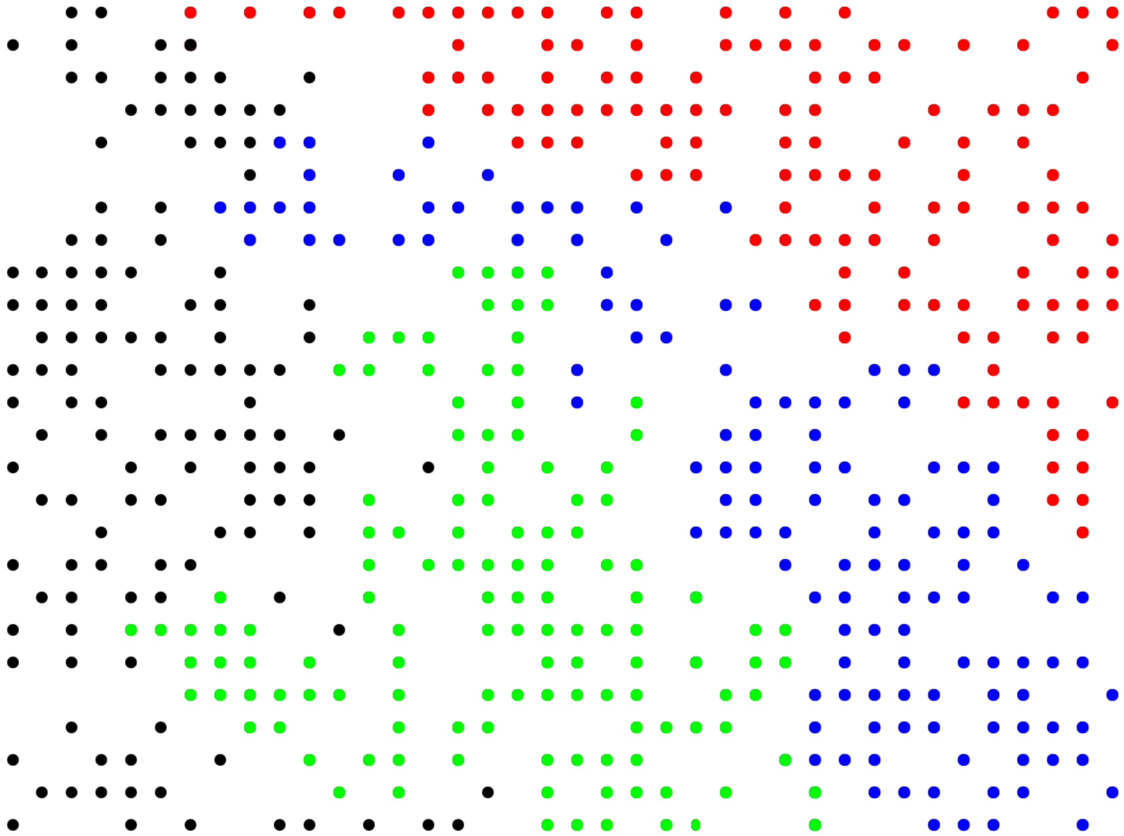

Figure 8: The color of the targets found by the members of the six 4-person groups who divided the search display by dot color. The different type bars represent the members who were assigned to different color.

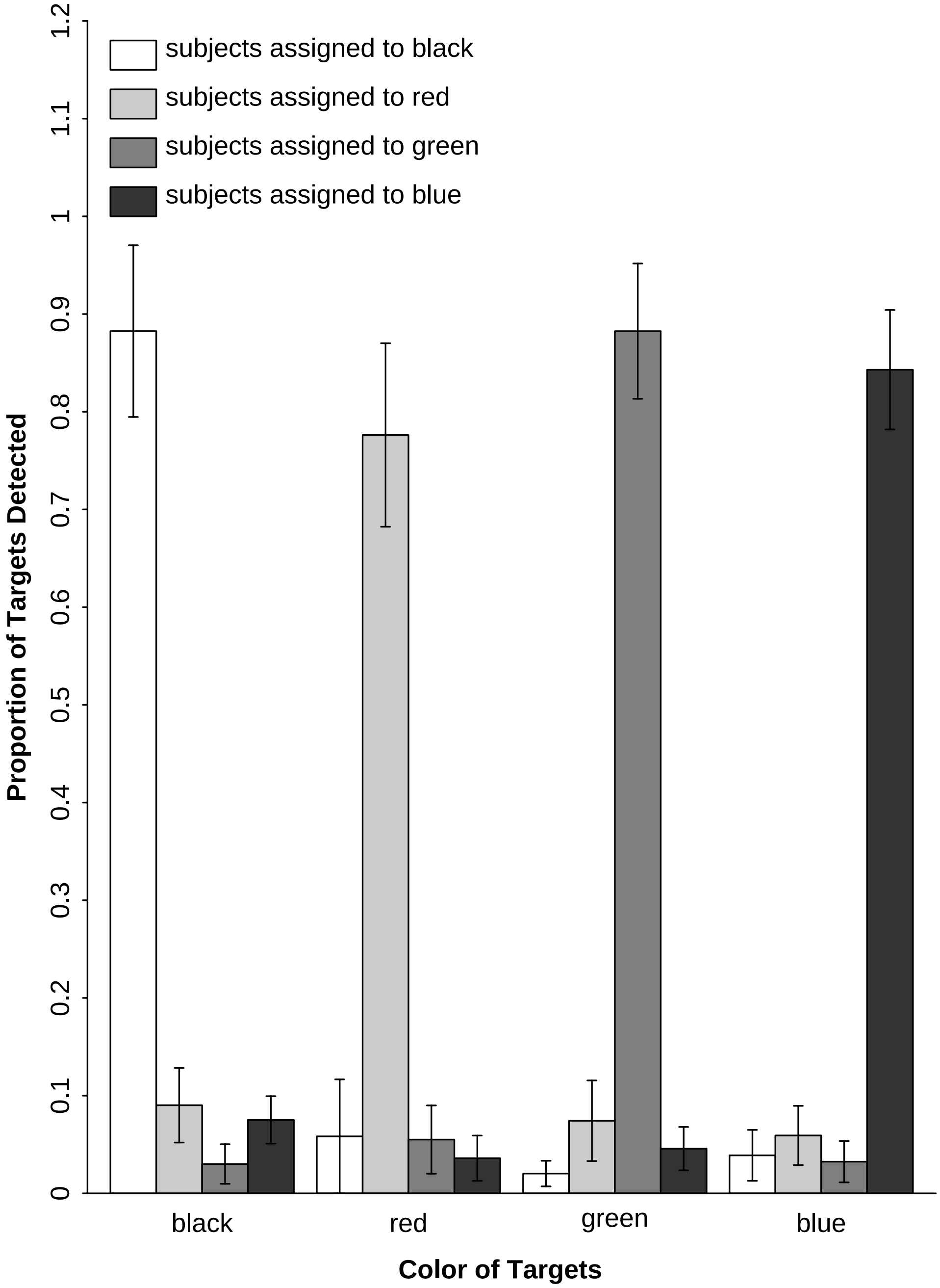

Figure 9: The search stimuli used in experiment 3. The target preview had either 1, 4 or 8 objects. For the 4 target preview, the top 2 objects from left to right were labeled as target 1 and target 2; the bottom 2 objects from left to right were labeled as target 3 and target 4. For the 8 target preview, the top 4 objects from left to right were labeled as target 1, target 2, target 3 and target 4; the bottom 4 objects from left to right were labeled as target 5, target 6, target 7 and target 8. The duration of the target preview followed a "1 second per object" rule. So a 1-object preview was displayed for 1 second, and an 8-object preview 8 seconds. The search display had constant 14 objects.

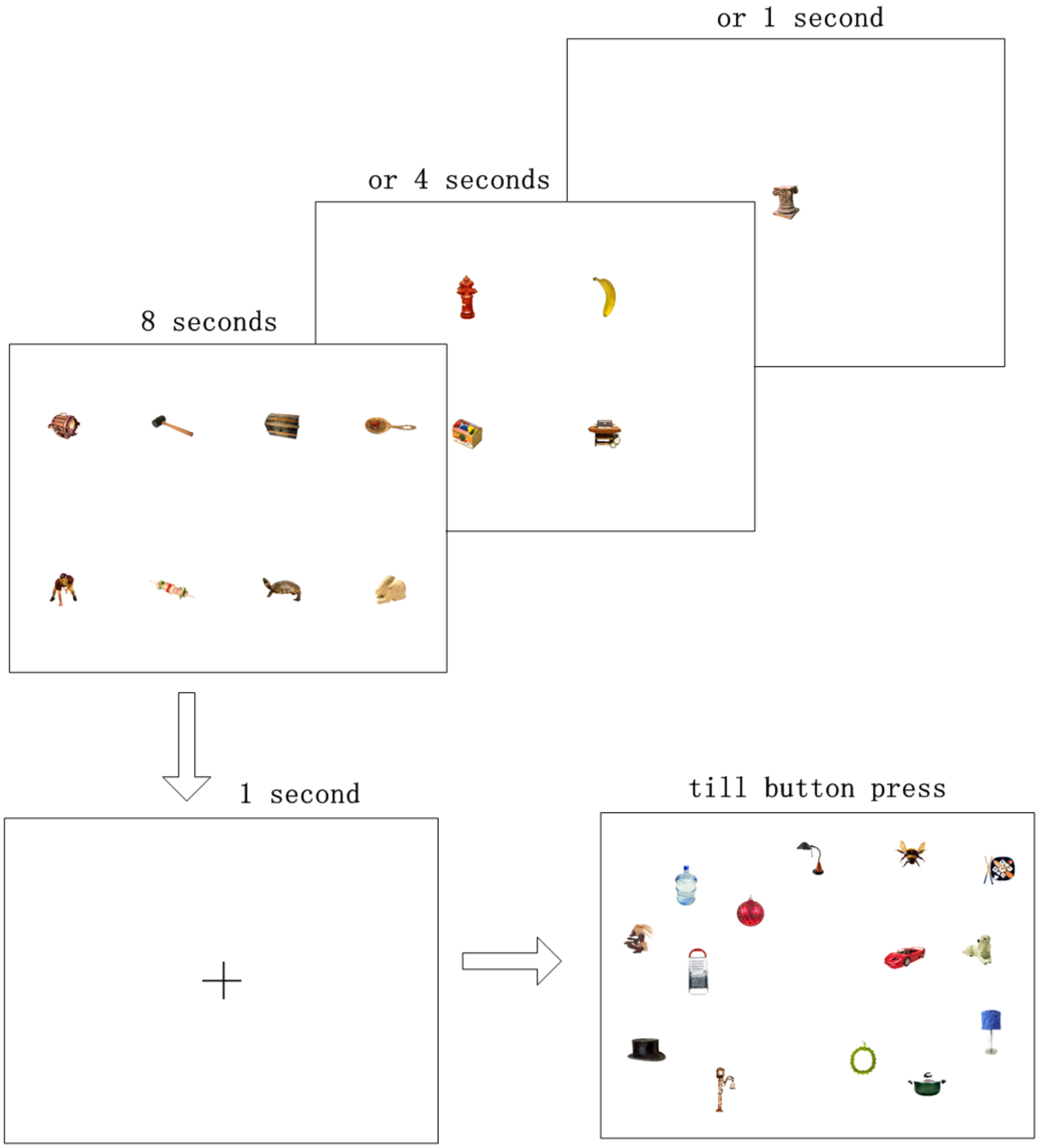

Figure 10: A: The relation between the target preview location and the target found by the five 2-person groups who divided the 4 targets. The white bars represent the members of the 2-person groups who were assigned to target 1 and target 2. The black bars represent the members of the 2-person groups who were assigned to target 3 and target 4. B: The relation between the target preview location and the target found by the five 4-person groups who divided the 4 targets. The different bars represent the members of the 4-person groups who were assigned to target 1, target 2, target 3 and target 4 respectively. The error bars represent one standard error.

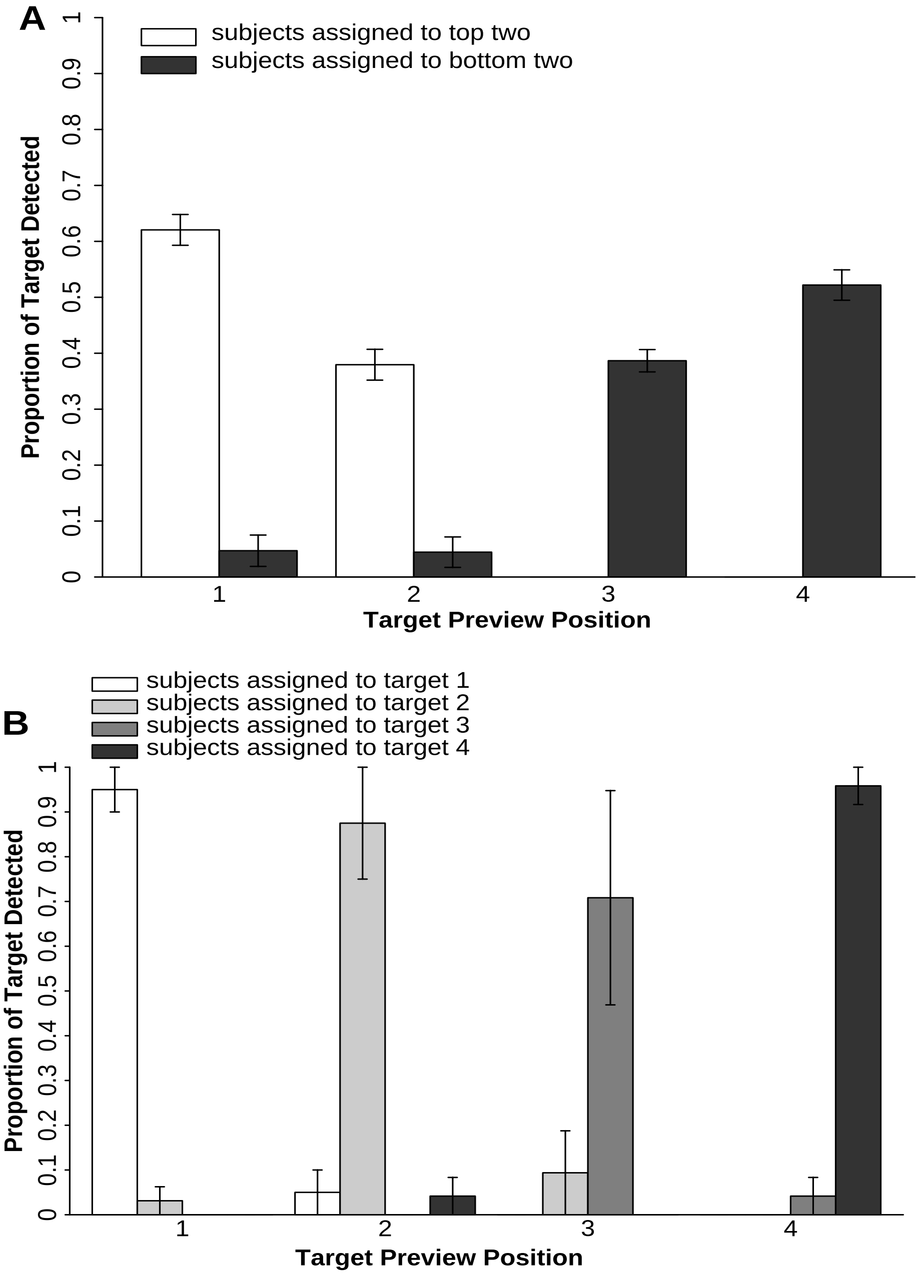

Figure 11: A: The relation between the target preview location and the six 2-person groups who divided the 8 targets. The white bars represent the members of the 2-person groups who were assigned to target 1, 2, 3, 4. The black bars represent the members of the 2-person groups who were assigned to target 5, 6, 7, 8. B: The relation between the target preview location and the six 4-person groups who divided the 8 targets. The different bars represent the different two targets that were assigned to different member.

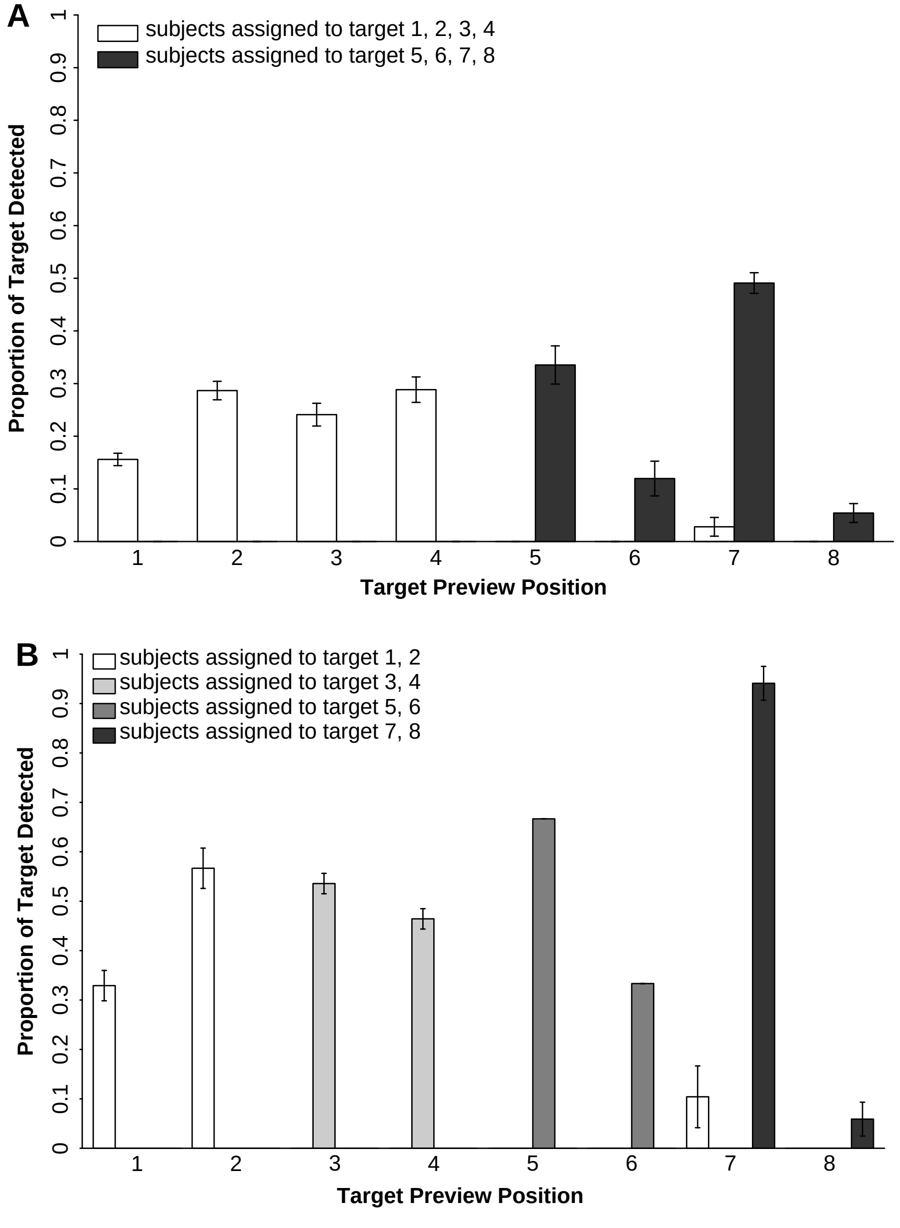

Figure 12: A: The reaction time of the target-present trials in experiment 3. B: The reaction time of the target-absent trials in experiment 3. C: The error rates of the target-present trials in experiment 3. D: The error rates of the target-absent trials in experiment 3. The error bars represent one standard error.

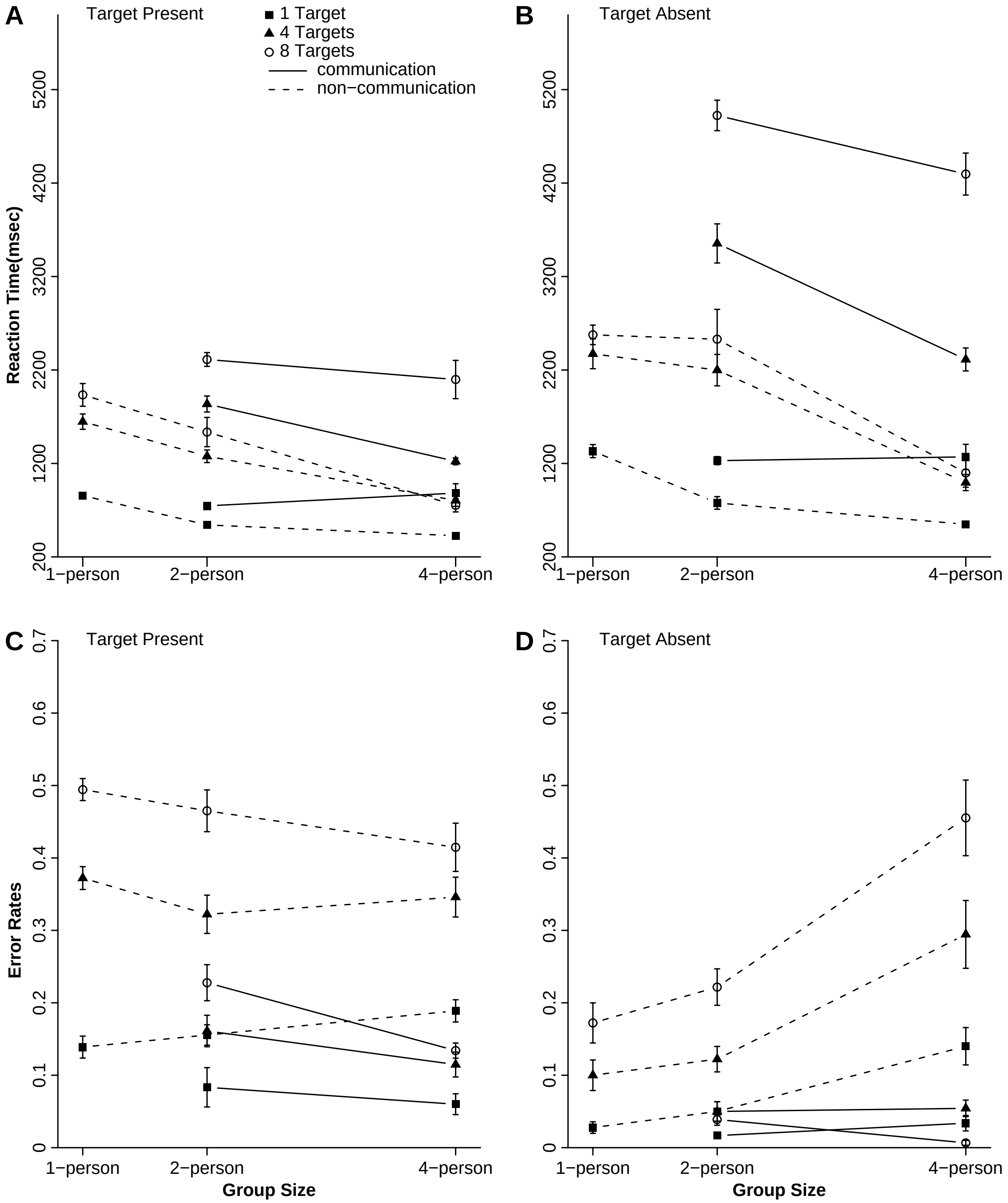

Figure 13: A: The mean sensitivity d' of the groups in experiment 3. The left is the 4-target trials data and the right is the 8-target trials data. The error bars represent one standard error. B: The mean response bias lnβ of the groups in experiment 3. The left is the 4-target trials data and the right is the 8-target trials data. The error bars represent one standard error.

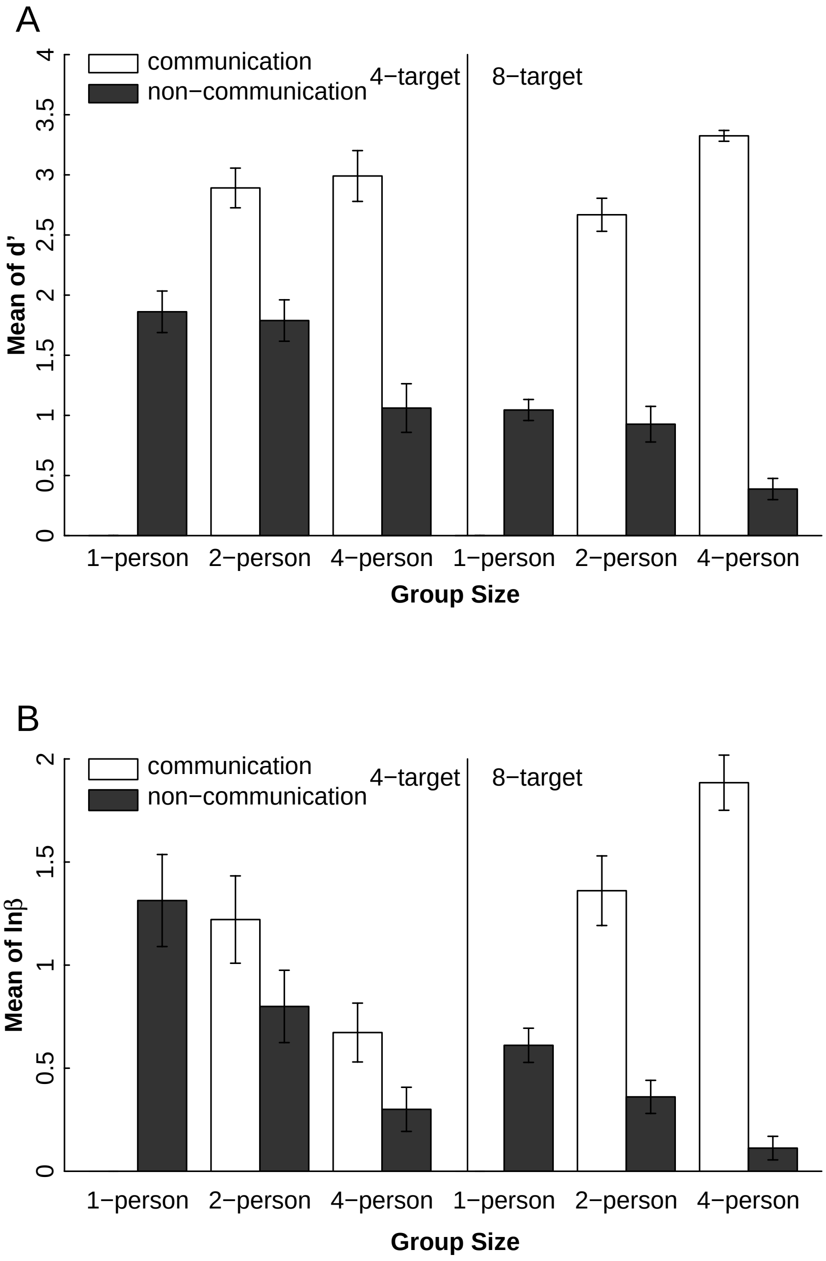

Figure 14: A sample of the search stimuli used in experiment 5.

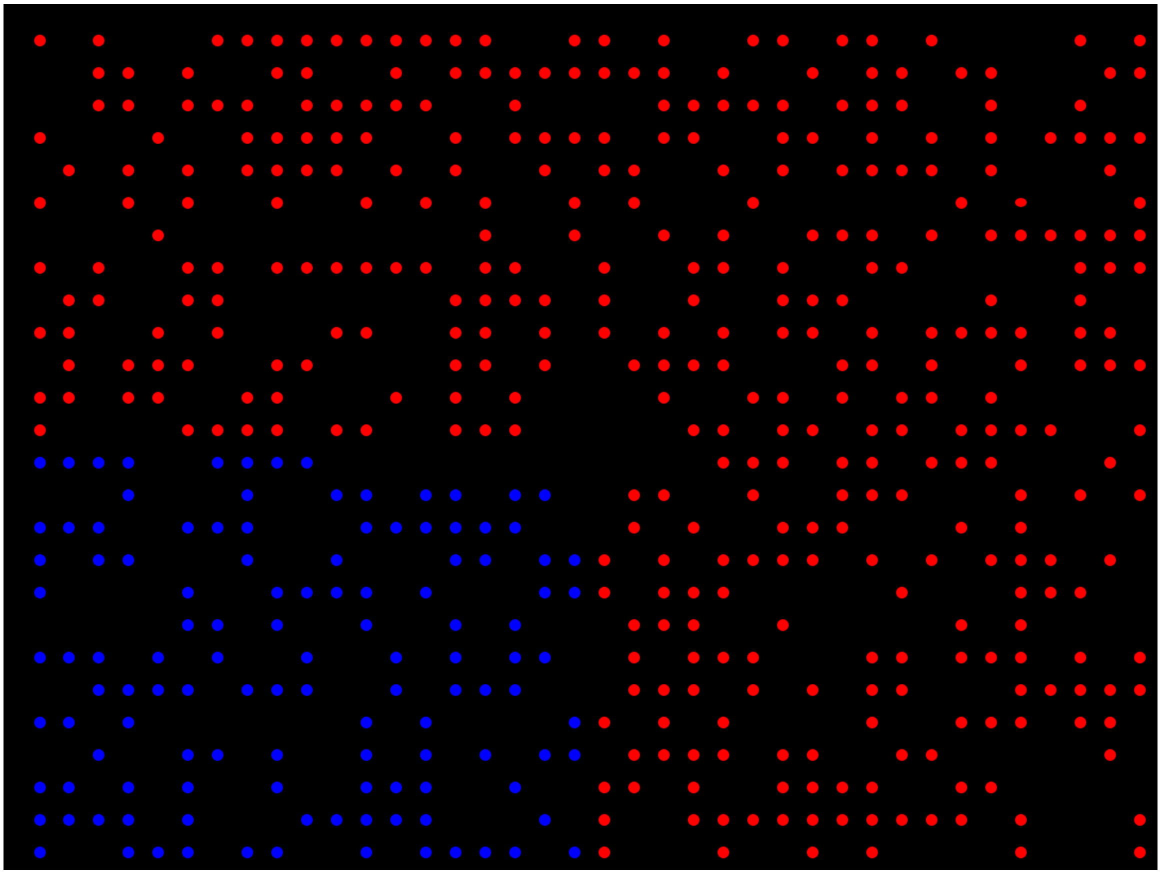

Figure 15: The scatterplots of the fixations of the dyads in experiment 5. The red triangles represent the fixations of the participants who were assigned to smaller dots group, and blue circles represent the fixations of the participants who were assigned to bigger dots group.

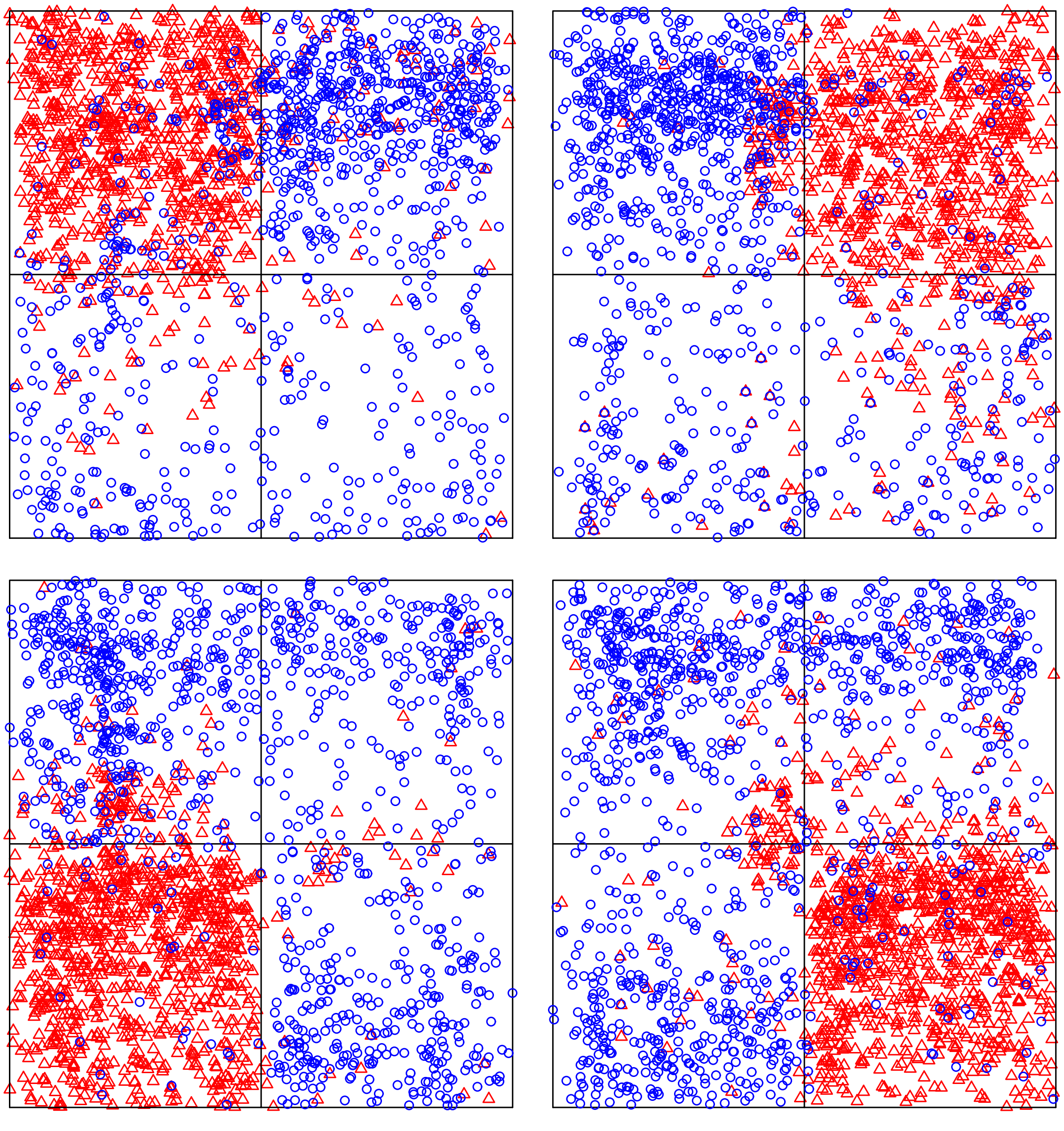

Figure 16: The average gaze distance of the dyadic searchers in experiment 5. The initial 2 groups of bars represent the gaze distance averaged over the whole experiment. The third group of bar represent the gaze distance averaged over from 10 seconds later of the start to the end of each trial of the experiment.