INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6” x 9” black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI
A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600
RICE UNIVERSITY

INTEGRATING EXPECTED SEARCH TIME AND TARGET DETECTION PROBABILITY IN VISUAL SEARCH STRATEGIES: THE UNDERVERVALUING OF SEARCH TIME

by

KIMBERLY DONNER O'BRIEN

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE

DOCTOR OF PHILOSOPHY

APPROVED, THESIS COMMITTEE:

David M. Lane, Director
Associate Professor of Psychology

Kenneth R. Laughery
Henry R. Luce Professor of Engineering Psychology

Eva J. Lee
Professor of Human Performance and Health Sciences

John W. Brelsford
Professor of Psychology

Houston, Texas
March, 1996
ABSTRACT

Integrating Expected Search Time and Target Detection Probability in Visual Search Strategies: The Undervaluing of Search Time

by

Kimberly Donner O'Brien

In trouble-shooting, subjects choose the order in which they test hypotheses. This choice often involves a trade-off between the time it takes to test a hypothesis and the probability that the hypothesis is correct. In visual search, subjects may have to choose which of two displays they are going to search first. The trade-off is similar in that the display most likely to contain the target may take longer to search. A series of experiments investigated whether subjects are efficient in the way they trade-off the size of a visual display and the probability it contains the target. Subjects were presented with information about the size of each of two displays and their respective probabilities of containing the target. The task was to choose which display to search first and then conduct the search as rapidly as possible. The first experiment showed that subjects did follow specific patterns of display choice, and that subjects have a strong tendency to overweight probability information. Experiment 1 also showed that subjects generally fail to take into consideration the effect of highlighting the target. When a
target is highlighted, the size of the display has a negligible effect on search time and therefore should make the size of the display unimportant in choosing which display to search. The display choice of most subjects was uninfluenced by target highlighting. Experiment 2 used a wider range of display sizes to see if that would make display size more salient and possibly induce subjects to pay more attention to it. As in Experiment 1, subjects paid too much attention to the probability that a display contains the target and not enough attention to the time it takes to search the display. Three methods of training subjects to consider the time to test a hypothesis when choosing hypotheses were evaluated in Experiment 4. Even though the surface characteristics of the training materials and the visual search task differed, training did improve subjects' performance on the visual search task. The relative neglect of the time it takes to test hypotheses may be a general phenomenon.
ACKNOWLEDGEMENTS

I hope that the members of my committee, David Lane, Ken Laughery, Eva Lee, and Jim Pomerantz, understand that the time they gave was always appreciated, and that the comments they provided were always illuminating. I would like to thank John Brelsford for agreeing to serve as a committee member towards the end of this dissertation project.

I would also like to acknowledge the members of the Human Factors Program at Rice University for their input and general encouragement. A special thanks is extended to Doris Malone, who looked after my interests while I was away from the university.

And my parents, my brothers, my sister, and my husband must be thanked because without their support and unceasing optimism this project would have been much more difficult to realize.
## TABLE OF CONTENTS

INTRODUCTION ...........................................................................................................1
  Visual Search of Alphanumeric Displays ...............................................................3
  User Search Strategies .........................................................................................4
  The Present Experiments ......................................................................................11

EXPERIMENT 1 ...........................................................................................................14
  Method ..................................................................................................................14
    Subjects ................................................................................................................14
    Experimental Design .............................................................................................14
    Stimuli ..................................................................................................................16
    Procedure ..............................................................................................................20
    Apparatus .............................................................................................................21
  Results ..................................................................................................................21
    Overall ................................................................................................................21
    Choice Order: Determination of Optimal Search ..............................................22
    Choice Order: Choice Patterns ..........................................................................27
    Choice Time .........................................................................................................33
    Questionnaire Responses ....................................................................................36
    Summary of Experiment 1 Results ....................................................................36

EXPERIMENT 2 ...........................................................................................................38
  Method ..................................................................................................................39
    Subjects ................................................................................................................39
    Experimental Design .............................................................................................39
    Stimuli ..................................................................................................................40
    Procedure ..............................................................................................................40
    Apparatus .............................................................................................................40
  Results ..................................................................................................................40
    Overall ................................................................................................................40
    Choice Order: Determination of Optimal Search ..............................................41
    Choice Order: Choice Patterns ..........................................................................43
TABLES


2. Comparison of Estimated Search Times and Probability to Size
   Ratios: Experiment 1 ................................................................. 26

3. Percent of Correct Group Choices in Conflict Trials: Experiment 1 ..... 30

4. Number of Subjects Using Strategies in Non-Directed and Directed Trials: Experiment 1 ............................................................. 33

5. Comparison of Estimated Search Times and Probability to Size
   Ratios: Experiment 2 ................................................................. 42

6. Percent of Correct Group Choices in Conflict Trials: Experiment 2 ..... 44

7. Number of Subjects Using Strategies in Non-Directed and Directed Trials: Experiment 2 ............................................................. 45

8. Percent of Subjects Using Specific Strategies in Experiments 1 and 2 ................................................................. 46

9. Number of Subjects Using Strategies by Training Condition .......... 56

10. Mean Correct Responses to .2/.8 and .4/.6 Trials by Condition ...... 57
FIGURES

1. Example of equation choice display from Ashby (1988) .................. 7
2. Group choice display ........................................................................ 17
3. Display of Group T ........................................................................... 18
4. Answer/Feedback display .................................................................. 19
5. Error display ...................................................................................... 20
6. Choice time interaction of highlight and decision type ...................... 48
7. Interaction of training condition and group size in .4/.6 conflict trials ... 58
8. Interaction of training condition and group size in .2/.8 conflict trials ... 59
INTRODUCTION

Search for a target is a pervasive and important task. People search for lost items, search for information in newspapers and other print material, and search for answers to problems. The order of their search is crucial to the effectiveness of their search. In such activities as fault diagnosis, hypothesis testing, and visual search, the order in which items are checked or tested affects the time needed to complete the task.

An efficient strategy is to start the task with the easiest-to-test subtasks or hypotheses: this approach could be called a "cheap check." As an example, assume that a car headlight is not working. Although there are several ways in which the problem could be approached, it would be more time-efficient to investigate this problem by initially trying relatively simple solutions. The "cheap check" strategy would begin by making sure that the headlight switch had been operated correctly. If this was not the solution to the problem, the headlight bulb could be checked to ensure that is not broken or burnt out. The fuses could be checked, followed by a check of the connection between the bulb socket and wires. Only after these have been checked would the more intensive task of tracing the electrical wires for problems be attempted.

The ability to discover the optimal strategy may also be applied to visual display search. Two components are important in optimizing search times: the presentation of the information in the display, and the perceptions and decisions of the searcher. Decades of research on information display format have demonstrated the display characteristics that influence the time needed to extract
information from the display. These characteristics include interline spacing, display density, font characteristics, and highlighting (e.g., color, blinking, bolding, underlining).

However, in many cases, the search for a target is not merely perceptual, as it is in the case of a highlighted, blinking target. Many display-search tasks in work situations present the user with a discrete number of information displays. Over time, the characteristics of these displays are learned, and factors such as location knowledge emerge, requiring that the search task be investigated as a cognitive, decision-making task as well as a perceptual task.

The information display format research is discussed in the first section of this Introduction, followed by a discussion of the other component of some search tasks: cognitive decision-making strategies and the pursuit of optimal search strategies. Following these summaries, a series of experiments on the ability to use basic format characteristics of information displays (such as location constancy, group size, and target highlight) to guide search is described. A guiding question throughout these experiments is whether subjects are able to take advantage of common, effective display characteristics to produce the fastest possible search. Research-derived interface standards and guidelines have improved interfaces, but recent research in computer program debugging has raised the question of whether better format may make a search time difference given poor performance on even simple decision-making tasks.
Visual Search of Alphanumeric Displays

Early search research focused on performance under basic properties of the visual display, such as interline spacing (Kolers, Duchnicky, & Ferguson, 1981), letter case (Poulton & Brown, 1986), display density (Callan, Curran, & Lane, 1977), and item alignment (Bailey, 1982). These were mostly legibility issues, the study of which was dictated by the need to present the information in at least a readable format.

Once these readability issues were addressed, research focused on higher-level display factors. Grouping of display items was one focus, the issues ranging from the number and size of groups to whether information should be grouped according to semantic association or function (Dodson & Shields, 1978).

The information derived from display search studies was combined and given an applied use in the research of Tullis (1983, 1984). The most basic and universal of these factors were manipulated and search times were measured. A regression model including overall density, local density, number of groups, and size of groups was determined to best predict user search time and has been used, in the form of a computer program (Tullis, 1986a), to evaluate the layout of visual displays.

The Tullis prediction model assumes that search is serial and otherwise unguided. In some search tasks and under some information displays, this assumption may be true. However, there is a class of search tasks and of information displays in which the underlying structure of display information may guide the user via semantic organization and location constancy of items. Of
particular interest in the present series of experiments is the concept that search for a target may be guided by previous exposure to the display: the user has learned the location of an item in a display and search is guided by that location knowledge.

This caveat was also noted by Burns, Warren, and Rudisill (1985) in the discussion of their results. In an experiment comparing information display layouts and the performance of naive and expert subjects, expert subjects searched current and reformatted versions of the same information equally well. The current version of the displays, although "poorly" formatted compared to the ergonomically designed reformatted displays of the same information, were familiar to the expert subjects. According to Burns, Warren, and Rudisill, these current displays could be quickly searched by experts because of their familiarity with the displays: these expert subjects "were able to call upon their semantic knowledge about the displays in answering the experiment questions."

**User Search Strategies**

As described above, the search for a target in an information display is not always a purely perceptual task. In many cases, the search for information on an alphanumeric display is a decision making task. In searching a display, the operator must choose which area of the display to search initially. These choices may be based on display characteristics such as readability, semantic grouping, and target location constancy. These choices may also be based on behavior, such as stereotypical reading behavior (left to right) and target expectancies. It
cannot be assumed that people will use the most efficient strategy or make the most efficient decisions when performing the search choice task. Decision making performance, according to Wickens (1984), "probably lies somewhere midway between chance and optimal" (p. 114).

The emphasis on efficient performance is an important concern in the decision making area. Decisions can be described as correct or incorrect, and they can be described as costly or even dangerous. The less than stellar decision making performance of subjects has been well documented: the case studies of incorrectly made decisions, predictions, and diagnoses are numerous (Wickens, 1984). Wickens cites examples of faulty decision making in areas as diverse as automotive repair, medical diagnostic testing, flight school student selection, military action, and nuclear incidents. The financial costs of incorrectly made decisions associated with these examples are millions of dollars spent on unnecessary auto repair, unnecessary medical tests, and pilot education and training for students who do not graduate.

The time spent searching the wrong part of a visual information display is also an unnecessary cost. Given the frequency of visual display search tasks, the use of inefficient choice decisions and strategies can incur both great time and financial costs.

Two examples from recent research focus attention on the ability (or inability) to use task or display information to make time-efficient choices. Both of these studies (Ashby, 1988; Dammon, 1992) investigated the use of the "cheap check" strategy in the context of a hypothesis testing (or debugging) task. The
basic task was the same in both studies. Subjects were given a set of five algebraic equations which varied in the number of terms (from 1 to 5). One of the five equations was incorrect, and subjects were required to solve the equations and find the incorrect equation as quickly as possible without risking accuracy.

Subjects were told that they could test the equations in any order they wished, and that each of the equations had an equal likelihood of being the incorrect equation, regardless of equation complexity (i.e., number of terms).

In order to accomplish the task, subjects choose one of the five equations to test by clicking a "Test" button to the left of the equation (Figure 1). A display repeating the equation and showing the values for the terms would then be presented. Subjects solved the equation with the provided terms and indicated whether the equation was correct or incorrect by clicking either the "Correct" or the "Incorrect" button on the display. If the equation was correct, the subject would return to the five equations, and choose a different equation for testing. Checkmarks were placed by equations that had been tested as reminders for the subjects (and to negate the need to test the equations in an sequential, orderly fashion).

Subjects selected and checked the equations until the incorrect equation was identified. Following the identification of the incorrect equation, a different set of five equations was presented, and the process continued until all twenty sets of equations had been completed.
Test >>>
5a + 2b + 3c + 2d = 46

Test >>>
5a + 2b + 3c + 2d + 3e = 37

Test >>>
5a + 2b = 29

Test >>>
5a = 10

Test >>>
5a + 2b + 3c = 22

Figure 1. Example of equation choice display from Ashby (1988).

In Experiment 1 of Ashby, 1988, sixteen college students performed the
tasks as described above. Ashby had hypothesized that subjects would use one
of three strategies for testing the equations: optimal (equations were tested from
fewest to most number of terms), sequential (testing progressed from the top
equation to the bottom equation, or from bottom to top), and not discernible (a
lack of a clear, apparent testing pattern).

Ashby found that the use of an optimal strategy in these tasks began
slowly, but increased with the number of sets of equations completed. Only one of
the sixteen subjects used the optimal strategy in the first set of equations: by the
twentieth set of equations, eleven of the sixteen subjects used the optimal
strategy when choosing equations for testing.

Of interest to Ashby was the discovery that five of the sixteen subjects did not complete this relatively simple task optimally. Instead of starting their check with the simplest equation (with only one term in the computation), and basing their choices on the number of terms in the equations, some subjects were tackling larger, more time-consuming equations, often creating more work for themselves, and requiring more time to test the equations.

Ashby also noted a propensity for subjects to use a particular choice strategy continually: subjects with no discernible strategy in the early trials often ended up using the optimal strategy in later trials, while subjects using a sequential strategy in early trials continued using that strategy throughout the trials.

The second experiment extended the task of Experiment 1 to be slightly more complex. The same experimental procedure was used with the exception that more than one of the five equations could be incorrect (specifically, there were two incorrect equations per set). As subjects checked for incorrect equations, they were told when they had located one of the incorrect equations and reminded that there was another incorrect equation in the set. Subjects were informed when the second incorrect equation in the set was identified, and they were subsequently advanced to the next set. Nine college undergraduates took part in the second experiment.

Although the task was only slightly more complex than that used in Experiment 1, the results of Experiment 2 were dramatically more negative. The
percentage of subjects who used the optimal strategy on the last problem was significantly lower than in Experiment 1: 11% in Experiment 2 compared to 69% in Experiment 1. As in Experiment 1, subjects who started using the sequential testing strategy in the early trials continued to use the strategy for the remainder of the problems.

Ashby attributed the difference in results between Experiment 1 and Experiment 2 to the ability of those in the simpler task to move from a diffuse testing strategy (or strategies) to the optimal strategy. As seen in Table 1, the percentage of subjects using no discernible strategy decreased as the trials progressed in Experiment 1 (most of the subjects moving to the optimal testing strategy), while those using no discernible strategy in Experiment 2 actually increased in number as the trials progressed.

Table 1. Percent of Subjects using Strategies at the Beginning and End of the Experimental Trials: Ashby (1988)

<table>
<thead>
<tr>
<th>Testing Strategy</th>
<th>Experiment 1</th>
<th>Experiment 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Trial 1</td>
<td>Trial 20</td>
</tr>
<tr>
<td>Optimal</td>
<td>12.50</td>
<td>68.75</td>
</tr>
<tr>
<td>Sequential</td>
<td>50.00</td>
<td>18.75</td>
</tr>
<tr>
<td>No Strategy</td>
<td>37.50</td>
<td>12.50</td>
</tr>
</tbody>
</table>

Although the task did increase in complexity from Experiment 1 to Experiment 2, the basic task of Experiment 2 is not an unduly strenuous one. The ability to realize that the time needed to ascertain the correctness of a 4-term
equation (e.g., $4a + 2b + 3c + 2d = 27$, where $a = 2$, $b = 1$, $c = 3$, and $d = 4$) compared to the time needed to solve a 1-term equation (e.g., $4a = 12$, where $a = 3$) should be apparent to subjects, if not initially, than after a few trials of such calculations. This reasoning holds regardless of the number of equations in the set that are incorrect.

The most time-efficient testing process is to test the simple equations (those with the least number of terms) before testing more complex equations. With a one out of five chance that any of the equations would be the incorrect equation, there was often no need to test all five before finding the incorrect equation. In many cases, solving the simplest equation first would reveal the incorrect equation, resulting in a relatively quick trial time. The fact that 31% of the subjects in Experiment 1 and 89% of the subjects in Experiment 2 did not realize the best way to quickly finish the task was surprising: especially considering that these college undergraduates had 20 trials in which to discover and apply the optimal strategy.

Ashby's unexpected findings were supported by Dammon (1992). In a replication and extension of Ashby's thesis, Dammon was interested in the learning and transfer of the use of the "cheap-check" strategy.

Dammon's first experiment investigated the effect of training on the ability of subjects to adopt the optimal strategy for the same set of tasks. The Experimental group performed twenty trials of the task used in Ashby's Experiment 1, in which there was one incorrect equation in each set of five equations. Following this "training", these subjects completed the tasks as they
were in Ashby's Experiment 2, in which there were 2 incorrect equations per set of five equations. A Control group performed the same total number of trials (40); all of the trials had two incorrect equations (as in Ashby's Experiment 2).

Subjects in the Experimental group learned the optimal strategy: 75% of subjects used the optimal strategy in Trial 1, and the percentage of subjects using the optimal strategy on the 20th trial increased to nearly 92%. The performance of Dammon's Control group, however, replicated the difficulty of uncovering the optimal strategy in the slightly more complex task (with two incorrect equations in the set of problems) as in Ashby's Experiment 2. Although 58% of the Control group subjects used the optimal strategy in Trial 1 (a possibly inflated number that may have included sequential strategy users, since the optimal choice in Trial 1 was also the top-most equation), not one of the subjects used the optimal strategy on the 20th trial.

The present series of studies looks at the same basic issue, the ability to discover and apply an optimal choice strategy, but in a slightly different paradigm. Instead of a hypothesis-testing exercise used in Ashby and Dammon, these studies focus on the ability to perform optimally in a commonly-performed task: searching for and extracting information from alphanumerical information displays.

The Present Experiments

The present set of experiments attempts to determine how subjects use display characteristics in their initial choice of which of two display groups to search first. Specifically, subjects were shown displays in which two display
groups varied in the probability of containing the target, the number of items in the
group, and whether the target was highlighted. As in the hypothesis testing
research, the goal of this present set of experiments is the description of the
ability to find and use the optimal strategy. For the purpose of this study, optimal
search is defined as the initial choice of the group with the shortest expected
search time.

A logical subject would consider both the probability that the target is in the
group and the time necessary to search the group before choosing which section
to search initially. Obviously a subject should chose the higher probability group
when it can be more quickly searched than the other group. Given a more
complicated set of display characteristics, subjects should trade-off the difference
in group probabilities and the difference in group sizes. For example, when
presented with the choice between a 4-item group with a 40% likelihood of
containing the target and a 16-item group with a 60% chance of containing the
target, a logical subject would choose to start with the former group. The latter
group is 1.5 times more likely to contain the target but has four times as many
items: a logical subject would chose on the basis of the time needed to search
(reflected by the group size).

The optimal search choice strategy is also contingent on the presence, or
absence of target highlighting. Highlighting was included as an experimental
variable because it has been proven to be an effective target location cue. It is
also an often-applied display factor, generally taking the form of reverse video,
blinking, color, boxing, or underlining. Color especially has been shown to
enhance performance on search tasks (Christ, 1975).

Highlighting is an effective cue in these display search tasks because it demands attention and separates the target (the highlighted item) from noise (non-highlighted display elements). A highlighted target obviates the need for a serial search of the display items, significantly reducing the time needed to locate a target. In the case of the valid application of highlight on the target item, the search for the target within a display group is nearly immediate; the absence of highlighting necessitates a serial search of the display group. In terms of optimal group choice strategy, highlight should affect choice in the following manner: When the target is highlighted, a logical subject would always chose the group with the highest likelihood of containing the target since the time to search a group is essentially unaffected by size.

Highlighting should be an aid to search that becomes apparent to subjects after only a few instances of searching highlighted and non-highlighted groups. The study sought to determine whether subjects are aware of the effectiveness of the highlight cue for their target search and whether they take advantage of this highly-effective cue in their search choice strategy.

There are several ways in which subjects could be inefficient. Subjects could overemphasize one piece of information (e.g., target location probability, group size), choose inconsistently, or exhibit individual preferences (e.g., choice of top-most or left-most group).

Can operators integrate the information present in a visual display to optimally direct their search? The following series of experiments manipulate
expectancies and costs of search, describe user search strategies, and compare actual to optimal strategies.

EXPERIMENT 1

Method

Subjects. Twenty-five Rice University graduate and undergraduate students participated as subjects. All subjects had normal, or corrected-to-normal, vision.

Experimental Design. Three within-subject independent variables were manipulated: the probability the target is in a particular group, group size, and type of search. The probabilities associated with each of the two groups summed to 1.0 and the following pairs of probabilities were used: 0/1.0, .2/.8, .4/.6, and .5/ .5.

Group size was manipulated by varying the number of items (4, 12, 16) per group. Items were four-letter non-words. Trial items were randomly selected from a sample of 96 English non-words (listed in Appendix A).

Type of search was manipulated by adding highlighting to the target item on half of the trials (representing the directed condition). A distractor item in the non-target group was also highlighted. In the non-directed search condition, all items were presented in standard font.

The trials were constructed to present all combinations of the conditions. One hundred and forty-four combinations were possible: 8 probabilities presented in the left group (0, .2, .4, .5, .5, .6, .8, 1.0) x 3 group sizes presented in the left
group (4, 12, 16 items) x 3 group sizes presented in the right group (4, 12, 16 items) x 2 types of search (directed, non-directed). Three replications of these combinations were tested, resulting in 432 trials per subject. The order of trials was randomly determined for each subject.

The number of .5/.5 trials (54) was doubled so that all pairs of probabilities would have the same number of trials (108). For example, the .2/.8 combination and its converse, the .8/.2 combination, resulted in 108 trials (54 trials each) with that particular ratio. Doubling the number of .5/.5 trials was identical to adding 54 converse trials (.5/.5 trials) to the "original" .5/.5 trials.

"Non-choice" trials (i.e., when the probability of one of the groups was 1.0) were used to determine the time needed by subjects to search the 4-, 12-, and 16-item groups.

The order of probability and size information was controlled: one-half of the subjects were shown group choice displays with the probability information presented above the size information and one-half of the subjects were shown the reverse order of information. Assignment of subjects to these conditions was random.

On each trial, the assignment of the target to either the left or the right group was determined in accordance with the slated probability. In addition, the location of the target within a search group was randomly determined.

Three dependent measures were collected: initial group choice, choice time, and search time.

The dependent variable of primary interest, group choice, was the record of
which of the two groups was initially chosen for search. Choice time was the
record of the time needed to indicate the choice of group for initial search, and
was defined as the elapsed time between trial onset and selection of the group.
Search time was the time between presentation of the selected group and a
"Found" response.

In the event that the subject did not locate the target in the initial group,
secondary group choice was recorded, as well as search time for that group.

*Stimuli.* A set of several displays constituted one trial and followed the
sequence: group choice display; displays of one, or both, of the search groups;
answer/feedback display; and, when necessary, error display.
**Figure 2.** Group choice display.

The group choice display (Figure 2) presented the trial target and the trial information. A bolded target (as shown in Figure 2) indicated that the target would be highlighted in the upcoming trial. Two areas, representing the two search groups, were positioned below the trial target: the areas were labeled "T" and "Y." The "T" and the "Y" labels were associated with the top-row, side-by-side keyboard buttons that the subjects pressed in order to choose a group. The probability and size information (i.e., items) associated with each group was placed in the corresponding areas.
Figure 3. Display of Group T.

Group displays repeated the target and presented only the items for the selected group. Group items were arranged in one column, with line numbers placed to the left of each group. Figure 3 shows the display of Group "T."

The answer/feedback display (Figure 4) allowed subjects to type the line number of the target and to monitor their search time for each trial.
Target on line #: 7  (Press the Return key after you have typed your answer.)

Your search time was 1.116 seconds.

**Figure 4.** Answer/Feedback display.

In the event of false alarms and misses, an error display was presented, describing the type of error committed (Figure 5).
ERROR! The target was not present in that group.

False Alarm.

Press the Return key when you are ready to begin the next trial.

**Figure 5.** Error display.

*Procedure.* Subjects were initially given experimental instructions (Appendix B): the instruction text was accompanied by illustrations of the particular screen elements. Twenty practice trials following the instructions allowed subjects to learn the sequence of, and the response to, experimental displays.

Following the presentation of the group choice display, subjects indicated the group they wished to search by pressing the corresponding keyboard button (i.e., "T" or "Y"). The group items were then displayed. The subject searched for the target and pressed any alphanumeric key to indicate that the target had been found. A press of the Return key allowed the subject to search the other group, and a press of any key indicated that the target had been found in that group.
Following a "Found" response, subjects typed the line number of the target. The time needed to search the group (or groups) for that trial was then displayed. Subjects pressed the Return key to begin a new trial.

Short breaks were allowed after one-third, and two-thirds, of the trials were completed. Breaks were signaled by the computer. Following the final trial, subjects completed a written questionnaire (Appendix C) describing their choice of search group and their search for the target.

Subjects were instructed that accuracy and speed of search was important. They were not instructed to make the choice of which display to search quickly. Minimizing speed of search was emphasized throughout the experiment by displaying search times after each search. No feedback on choice time was given.

**Apparatus.** Presentation of the instructions, practice trials, and experimental trials, and the collection of data, were done using the HyperCard programming environment (version 2.1). Experimental sessions were conducted using a Macintosh LC computer with a 12-inch monitor.

**Results**

**Overall.** On average, subjects made a search error (either a false alarm or a miss) on 3.2 percent of the trials (range = 0.5 to 9.0 percent). A false alarm occurred when the subject indicated that the target had been found while searching the non-target group; a miss occurred when the subject pressed the Return key to search the other group while searching the target group. Subjects gave an incorrect target line response in less than 1 percent of the trials (mean
percent = .96, range = 0 to 3.1 percent). Incorrect answer trials were deleted from
search time analyses.

Choice Order: Determination of Optimal Search. In order to compare the
efficacy of search with an "optimal search," it is necessary to know the speed with
which the subject can search a group of a particular size. The "certainty trials"
(i.e., where one of the search groups had a 1.0 probability of containing the target)
were used as timing trials for this purpose. Since subjects were certain that the
target was contained in the group, relatively pure search times based only on size
of group (and the presence of highlight) were collected.

Overall, the median search times for 4-, 12-, and 16-item groups in the
non-directed condition were 1.10, 1.93 and 2.50 seconds, respectively. In the
directed condition, the median search times for 4-, 12-, and 16-item groups were
0.80, 0.90 and 0.97 seconds, respectively.

Faced with two groups from which to choose, there are two expected
search times that can be determined: an expected time based on a search of the
left-most group prior to a search of the right-most group, and an expected time
assuming a search of the right-most group prior to a search of the left-most group.

Both expected search times are based on the probability of the target being
in the group and the median time needed by each subject to search the groups.

The expected search times (EST) are calculated in the following manner:

Variables

a = left group probability

b = right group probability (where a + b = 1.0)
\[ x = \text{median search time of left group} \]
\[ y = \text{median search time of right group} \]

**Initial search of the left group**

\[ \text{EST (left group)} = ax + b(y + x), \]

**Initial search of the right group**

\[ \text{EST (right group)} = by + a(x + y). \]

The optimal choice for that trade-off is defined as the group with the shorter of the two expected search times.

The following is an example of the computation of expected search time. The median search time for Subject 5 to search a 4-item group was 1.28 seconds and the time needed to search a 12-item group was 2.21 seconds (non-directed searches). To determine the optimal search for Subject 5 in the .2(4)/.8(12) condition (where the left display group has four items and a probability of .2 and the right display group has twelve items and a probability of .8), the following steps are taken:

**Variables**

\[ a = .2 \text{ = left group probability} \]
\[ b = .8 \text{ = right group probability} \]
\[ x = 1.28 \text{ seconds = median search time of left group (4-item group)} \]
\[ y = 2.21 \text{ seconds = median search time of right group (12-item group)} \]

**Initial search of the left group**

\[ \text{EST (left group)} = (.2)(1.28) + (.8)(2.21 + 1.28) = 3.05, \]
**Initial search of the right group**

\[
\text{EST (right group)} = (.8) (2.21) + (.2) (1.28 + 2.21) = 2.47.
\]

In this case, the expected search time for the initial search of the right group [.8(12)] was shorter (2.47 seconds) than the expected search time for the left [.2(4)] group (3.05 seconds). The optimal choice for this condition/trade-off is considered to be the initial search of the .8(12) group.

Table 2 presents expected search times for non-directed .2/.8 and .4/.6 trials. When the ratio of probability to size is greater than one, the expected search time is lower for the “right” group; when the ratio is less than one, the expected search time is lower for the “left” group.

The definition of optimal group choice based on expected search times was supported by actual subject performance. Actual search performance on a trial type was derived by applying the above calculations. There was general agreement concerning which initial choice resulted in faster estimated search. In a small number of instances, the subject's estimated search time for the 'non-optimal' choice was shorter than the estimated search time for the 'optimal' choice. This was restricted to only 3 of the 36 trial types: .4(4)/.6(12), .4(4)/.6(16), .4(12)/.6(16). In addition, within those trial types, most subjects' expected optimal and actual optimal search choices were the same. The proportion of subjects not in agreement in the above trial types were, respectively: 3/25, 2/25, 5/25.

This was a manipulation check: these estimated search times should have supported the optimal search argument. In addition, in the directed trials, the initial
search of the higher probability group (i.e., .6, .8, 1.0) always resulted in faster estimated search times, supporting the optimal strategy of searching higher probability groups when the target was highlighted. This was supported, without exception, by the actual subjects' search times.
Table 2. Comparison of Estimated Search Times and Probability to Size Ratios:

**Experiment 1**

<table>
<thead>
<tr>
<th>Condition [left probability(left size)/right probability(right size)]</th>
<th>Ratio of Probability to Size</th>
<th>Expected Search Time (seconds): Left Group</th>
<th>Expected Search Time (seconds): Right Group</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Probability faster conditions</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.2(12)/.8(16)</td>
<td>4:1.33</td>
<td>4.03</td>
<td>2.93</td>
</tr>
<tr>
<td>.2(4)/.8(12)</td>
<td>4:3</td>
<td>2.79</td>
<td>2.25</td>
</tr>
<tr>
<td>.4(12)/.6(16)</td>
<td>1.5:1.33</td>
<td>3.53</td>
<td>3.33</td>
</tr>
<tr>
<td>.2(4)/.8(16)</td>
<td>4:4</td>
<td>3.20</td>
<td>2.76</td>
</tr>
<tr>
<td><strong>Size faster conditions</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.4(4)/.6(16)</td>
<td>1.5:4</td>
<td>2.69</td>
<td>2.99</td>
</tr>
<tr>
<td>.4(4)/.6(12)</td>
<td>1.5:3</td>
<td>2.39</td>
<td>2.48</td>
</tr>
<tr>
<td><strong>No conflict conditions</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.2(16)/.8(4)</td>
<td>4:25</td>
<td>3.47</td>
<td>1.68</td>
</tr>
<tr>
<td>.2(12)/.8(4)</td>
<td>4:33</td>
<td>2.96</td>
<td>1.58</td>
</tr>
<tr>
<td>.2(16)/.8(12)</td>
<td>4:75</td>
<td>4.14</td>
<td>2.52</td>
</tr>
<tr>
<td>.2(4)/.8(4)</td>
<td>4:1</td>
<td>2.12</td>
<td>1.41</td>
</tr>
<tr>
<td>.2(12)/.8(12)</td>
<td>4:1</td>
<td>3.63</td>
<td>2.42</td>
</tr>
<tr>
<td>.2(16)/.8(16)</td>
<td>4:1</td>
<td>4.54</td>
<td>3.03</td>
</tr>
<tr>
<td>.4(16)/.6(4)</td>
<td>1.5:25</td>
<td>3.23</td>
<td>2.19</td>
</tr>
<tr>
<td>.4(12)/.6(4)</td>
<td>1.5:33</td>
<td>2.72</td>
<td>1.98</td>
</tr>
<tr>
<td>.4(16)/.6(12)</td>
<td>1.5:75</td>
<td>3.73</td>
<td>3.02</td>
</tr>
<tr>
<td>.4(4)/.6(4)</td>
<td>1.5:1</td>
<td>1.88</td>
<td>1.65</td>
</tr>
<tr>
<td>.4(12)/.6(12)</td>
<td>1.5:1</td>
<td>3.22</td>
<td>2.82</td>
</tr>
<tr>
<td>.4(16)/.6(16)</td>
<td>1.5:1</td>
<td>4.04</td>
<td>3.53</td>
</tr>
</tbody>
</table>
Choice Order: Choice Patterns. This section focuses on the actual choice patterns of the subjects. The approach to the analysis of the choice data, especially in regard to the focus on individual subject patterns, is similar to that taken in Siegler's classic 1976 study, Three Aspects of Cognitive Development. Siegler's main thesis was that "children whose performance conformed to different rules would display dramatically different patterns of successful and unsuccessful predictions" (page 487) on a series of balance scale problems. Siegler therefore sought to classify individual subjects on the basis of the rules they used rather than rely on summary group statistics that can obscure psychological processes. As in the present study, subjects had to trade off two attributes. In the balance scale problem, the trade-off is between the amount of weight and the distance from the center of the scale.

Siegler discovered four "Rules" that could guide children's behavior on these balance scale problems. Some children considered only a single dimension (usually weight); others gave some consideration to weight and distance, but overvalued one dimension, using the other mainly as a tie-breaker; others considered both weight and distance dimensions, but were not able to combine them optimally; some children knew the optimal multiplication rule for combining weight and distance.

The basic test was as follows: metal weights were placed on predetermined pegs on the balance scale and the child was asked to predict which side of the scale would go down, or whether the scale would balance. There were six Problem Types: balance (equal weights, equal distances from the
fulcrum), weight (unequal weights, equal distances from the fulcrum), distance (equal weights, unequal distances), conflict-weight (conflict trials in which the side with the greatest weight should go down), conflict-distance (conflict trials in which the side with the weight farthest from the fulcrum should go down), and conflict-balance (conflict trials in which the scale should balance).

Siegler's hypothesis was that children categorized to be in different Rule groups would demonstrate different patterns of prediction to these Problem Types. For example, children using one of the rules would be able to predict the balance behavior for balance, weight, and distance problems correctly, but their inability to form rules to handle discrepant weight/distance cues would result in chance responding to the conflict problem types.

Siegler's analytical approach maps well to the needs of the present set of experiments. Siegler's classification of how children may respond to the balance scale predictions (his "Rules") is similar to the classification of how subjects in the present experiments may respond to the group choice trials (as listed below). This type of analysis allows error patterns to be investigated and reported, and also allows one to investigate and report summaries of incorrect and correct responses. In many analyses, it is more informative and meaningful to describe responses from individual subjects than to provide values averaged across subjects. This was seen as crucial in the present study (as well as in Siegler's): averaging responses to the trial types (or to the balance problems) across subjects would lose a great amount of information and could be misleading. The averaging of qualitative responses could hide the fact that some individuals were
acting consistently, albeit differently from others.

In the present study, each subject had been exposed to 6 replications of a specific trade-off (e.g., .6/.4, no highlight). When the subject chose one of the two groups for initial search in either 5 (83%) or 6 (100%) of the 6 trade-offs, this was categorized as a consistent choice. This consistent choice was categorized as Optimal or Non-optimal, depending on whether the subject's consistent choice matched the optimal search choice for that trade-off.

In some cases, the subject did not respond consistently to a particular trade-off (e.g., the subject initially chose one of the groups on 2 of the 6 trials, choosing the other group on the other 4 trials). In this event, the response to those conditions was categorized as Inconsistent.

Table 3 presents group choice performance under the trade-off conditions. Subjects were adept at choosing the group with the faster expected search time when the trade-off strongly favored one group: however, in more complicated trade-offs, especially those favoring the choice of the smaller group, subjects were less efficient in selecting the optimal group.

The search choice patterns of subjects in the .4/.6 trials are of particular interest since the .4/.6 trials comprised a set of trade-off choices in which the optimal choice was not limited to the higher probability group and in which a more difficult trade-off was required. In the .2/.8 trials, the .8 group always resulted in the faster estimated choice, regardless of group size. The .5/.5 trials were a special case: the required trade-off was relatively simple, since only group size varied in those trials. However, in the .4/.6 trials, the .6 and the .4 group both
resulted in the faster estimated time, depending on the associated group size.

Table 3. Percent of Correct Group Choices in Conflict Trials: Experiment 1

<table>
<thead>
<tr>
<th>Condition</th>
<th>Percent of Subjects Who Chose Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Correct</td>
</tr>
<tr>
<td>NON-DIRECTED TRIALS</td>
<td></td>
</tr>
<tr>
<td>Probability faster conditions</td>
<td></td>
</tr>
<tr>
<td>.2(12)/.8(16)</td>
<td>100</td>
</tr>
<tr>
<td>.2(4)/.8(12)</td>
<td>84</td>
</tr>
<tr>
<td>.4(12)/.6(16)</td>
<td>60</td>
</tr>
<tr>
<td>.2(4)/.8(16)</td>
<td>92</td>
</tr>
<tr>
<td>Size faster conditions</td>
<td></td>
</tr>
<tr>
<td>.4(4)/.6(16)</td>
<td>32</td>
</tr>
<tr>
<td>.4(4)/.6(12)</td>
<td>16</td>
</tr>
<tr>
<td>DIRECTED TRIALS (all Probability faster conditions)</td>
<td></td>
</tr>
<tr>
<td>.2(12)/.8(16)</td>
<td>100</td>
</tr>
<tr>
<td>.2(4)/.8(12)</td>
<td>88</td>
</tr>
<tr>
<td>.4(12)/.6(16)</td>
<td>64</td>
</tr>
<tr>
<td>.2(4)/.8(16)</td>
<td>96</td>
</tr>
<tr>
<td>.4(4)/.6(16)</td>
<td>52</td>
</tr>
<tr>
<td>.4(4)/.6(12)</td>
<td>52</td>
</tr>
</tbody>
</table>

The following classifications of individual subjects' choice patterns are possible:

- optimal: chose optimal group in all trade-offs
- complete probability: chose higher probability group, regardless of group size, in all trade-offs
• strong probability: chose optimal group in most trade-offs, but predominantly used or incorrectly emphasized probability in some trade-offs

• strong group size: chose optimal group in most trade-offs, but incorrectly emphasized size in some trade-offs

• inconsistent: used a strategy, correctly or incorrectly, in most trade-offs but did not always use the strategy in trade-offs in which it was appropriate.

Subjects could also be categorized as using an Inconsistent strategy if they demonstrated a pattern of choosing inconsistently within a trade-off condition (e.g., choosing the .4(16) group twice and choosing the .6(12) group in the remaining four trials in the .4(16)/.6(12) condition). If the subject’s responses revealed a pattern of inconsistent choice across a number of conditions, this response pattern was categorized as Inconsistent.

As a test of this categorization scheme, a simulation was done to approximate the choice patterns that would result from a strategy of random group choice. In order to simulate a subject who is simply choosing a group at random, the experimenter performed the Experiment 1 trials, randomly choosing a group for initial search. In order to ensure the randomness of group choice, digits in a random order table guided the choice of group: on the occurrence of an odd-numbered digit, Group T was chosen initially, and on the occurrence of an even-numbered digit, Group Y was chosen initially. The experimental trials were completed in this fashion. Thirty such simulations were conducted, representing a
hypothetical thirty subjects, following this random choice strategy.

The data from this simulation were analyzed in the same manner as the data from the Experiment 1 subjects (described in this section). All 30 subjects were classified as Inconsistent (for both non-directed and directed trials).

The search strategies of the actual subjects under non-directed and directed conditions demonstrated patterns of choice that were not compatible with a strategy of random choice. As shown in Table 4, 15 of 25 subjects responded in such a manner that their responses to the non-directed trials were consistent and followed a particular strategy pattern (i.e., Optimal, Complete Probability, Strong Probability, Size). Sixteen subjects had consistent responses to the directed trials.

Few subjects chose optimally: two subjects chose optimally in non-directed trials and seven subjects choose optimally in the directed trials. Surprisingly, none of the subjects changed strategy in order to choose optimally in both non-directed and directed trials. For both the non-directed and directed conditions, the remaining subjects tended to choose either inconsistently or based on group probability. Seventeen of the 25 subjects used the same strategy in both non-directed and directed trials.

Subjects' responses to the .5/.5 trials were also investigated. For these analyses, only non-directed trials are reported. In those trials in which the group size was unequal [e.g., .5(4)/.5(16)], 15 of the 25 subjects consistently chose the smaller of the two groups, one subject consistently chose the Right group (the Y group), and nine of the subjects chose inconsistently. In those trials in which the group size was equal [e.g., .5(12)/.5(12)], 10 of the 25 subjects consistently chose
the Left group (the T group), 2 of the subjects consistently chose the Right group (the Y group), and 13 of the subjects chose inconsistently.

Table 4. Number of Subjects Using Strategies in Non-Directed and Directed Trials: Experiment 1

<table>
<thead>
<tr>
<th>Strategy in Directed Trials</th>
<th>Strategy in Non-Directed Trials</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Optimal</td>
</tr>
<tr>
<td>Optimal/Complete Probability</td>
<td>0</td>
</tr>
<tr>
<td>Strong Probability</td>
<td>0</td>
</tr>
<tr>
<td>Size</td>
<td>1</td>
</tr>
<tr>
<td>Inconsistent</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>2</td>
</tr>
</tbody>
</table>

**Note.** In the Directed Trials, a Complete Probability strategy is categorized as an Optimal strategy. Optimal and Near Optimal categories are combined into the Optimal category.

**Choice Time.** The analysis of the choice time data evaluated the information present on the Group Choice display as the subject made a choice of initial group for search. Information Placement was one such piece of information: some subjects were presented Group Choice displays in which the Probability Information was displayed above the Items information, and some subjects were presented Group Choice displays in which the Items information was displayed above the Probability information. The Information Placement condition was not
expected to affect choice time significantly, but it merited analysis as a manipulation check and also because an argument could be made that stereotypical English-language reading behavior (i.e., left to right, top to bottom) might result in Information Placement having an effect on choice time.

Highlight was another piece of information that could (and should) affect choice time. As discussed in previous sections, highlight could be an important, and relatively easy to use, cue to efficient search: when the target was highlighted, the appropriate search strategy is to initially search the group with the higher probability. The choice time analysis will demonstrate whether subjects effectively used highlight as a cue: choice time to highlighted trials should be lower than choice time to trials without highlight if subjects are using the most efficient strategy in highlighted trials.

The last pieces of information displayed to subjects via the Group Choice display were the Probability information and the Items information for each group. In order to organize these pieces of information (and their combinations), trial data were categorized as representing one of three 'decision types.' This Decision Type factor was thought to best represent the trade-off between probability and size information faced by the subject when making a choice. Trials were categorized as either representing Agreement [the probability information and the item information both favored the same group, for example, .2(16)/.8(4)], Disagreement [probability information and item information must be further evaluated since probability and item information favor different groups, for example, .4(4)/.6(16)], or Neutral [probability information and item information
favor neither of the two groups, for example, .5(12)/.5(12)].

It was expected that subjects would respond faster to the Agreement trials than to the Disagreement trials, since the Disagreement trials would require a more rigorous analysis of the trade-off. Choice times to the Neutral trials were expected to fall somewhere between the Agreement and the Disagreement trials: Neutral trials would not be as easy to respond to as the Agreement trials, but the Neutral trials present an easier trade-off condition than the Disagreement trials.

A 2 (Information Placement: Probability over Items, Items over Probability) x 2 (Highlight: No Highlight, Highlight) x 3 (Decision Type: Agreement, Disagreement, Neutral) general linear models (GLM) procedure revealed that the time needed to chose a group for initial search was not affected by Information Placement, $F(1, 23) = 0.01, p < .92$. Mean choice time to the Probability over Items displays was 2.26 seconds and mean choice time to the Items over Probability displays was 2.23 seconds. Likewise, the main effect of Highlight was not significant, $F(1, 23) = 3.06, p < .09$. Mean choice time to the Non-highlighted trials was 2.22 seconds and mean choice time to the Highlighted trials was 2.27 seconds.

There was a significant difference in choice time for the three levels of Decision Type, $F(2, 46) = 6.02, p < .005$. As expected, mean choice time to the Agreement trials (2.13 seconds) was shorter than mean choice time to either the Disagreement trials (2.27 seconds) or the Neutral trials (2.34 seconds). A contrast comparison of means was conducted, comparing the Agreement and Disagreement means to the Neutral mean. Mean choice time to the Agreement
trials was significantly shorter than mean response time to the Neutral trials ($p < .01$). However, the difference in mean choice time between these Neutral trials and the Disagreement trials was not significant at the $p < .05$ level.

None of the interactions (either the 2- or the 3-way interactions) were significant at the $p < .05$ level.

*Questionnaire Responses.* Questionnaire responses echoed the emphasis on probability. All but one subject reported that they combined the probability and size information when choosing between groups; 23 of the 25 subjects reported using the probability information more than the other information in their group choice.

Only 8 of the 25 subjects reported that they followed a particular strategy for highlighted trials: specifically, they chose the higher probability group in trials in which the target was highlighted (the correct strategy for highlighted trials). A Pearson product-moment correlation was computed in order to determine the relationship between this self-report of choosing the higher probability group in highlighted trials and the actual use of this optimal strategy. The correlation was not significant, $r = 0.016$, $p < .94$.

All 25 subjects reported that highlight affected their search of a group (to summarize, subjects stated that highlighted search was quicker because the bold target "popped-out" from the other items).

*Summary of Experiment 1 Results.* The results of Experiment 1 demonstrated that subjects did follow specific patterns of search group choice: a majority of subjects could be categorized as using one of the specific group
choice patterns. Even when there was no benefit to the choice of either of the groups (because probability and size were the same for the two groups), the behavior of subjects could be classified into a particular pattern of choice.

Additionally, the results of Experiment 1 demonstrate that subjects favored a pattern of group choice based on probability. Probability information was overwhelmingly used in group choice, both in objective and subjective data. The majority of subjects used either a Complete Probability or a Strong Probability strategy, and most subjects reported in the questionnaire that they used the probability information more than the other information in their decision of initial group choice.

One explanation to the emphasis on probability is that it is a familiar and salient bit of information. An alternative explanation to the emphasis on probability in group choice, besides the saliency of probability, is that subjects were trying to minimize total search time by consistently choosing a higher probability group. According to this view, subjects could presume that the best way to quickly finish the experimental trials would be to choose one of the two groups immediately, without trade-off consideration of the probability, size, and highlight information. Subjects could either choose a group randomly or choose the higher probability group: choice time would be minimal, and the subject could focus on the search portion of the trial.

As stated previously, subjects were instructed to search quickly and their search times were presented after each search: they were not instructed to make the choice of which display to search quickly. Nonetheless, the hypothesis that
subjects quickly chose the higher probability group to minimize total time was addressed by a small experiment.

To investigate the choice time that would result from a strategy of higher-probability group choice, the experimenter completed a set of experimental trials \((n = 432)\) with the strategy of choosing a group based on a probability strategy: as soon as the Group Choice display was presented, the experimenter choose the group associated with the higher probability. For purposes of this experiment, only trials in which the Probability information was displayed above the Size information (Condition 1) were used.

The median choice time based on this probability choice strategy was 0.5 seconds. The median choice time for subjects in Experiment 1 ranged from 1.18 to 3.55 seconds. (This data is based on the 13 subjects who were randomly assigned to Condition 1 trials.) Even the minimum choice times (i.e., each subject’s shortest choice time, out of 432 trials) were relatively high: the average minimum choice time was 1.10 seconds. Therefore subjects’ choice time was not consistent with a strategy of immediate, higher probability group choice. If subjects were choosing a group using a simplified higher probability choice strategy (or even using a random choice strategy), they would not have taken so long to respond.

**EXPERIMENT 2**

Experiment 1 demonstrated that the majority of subjects chose the higher probability group for initial search, even when inappropriate. Subjects professed
to using, and actually used, a probability-driven initial group choice. One issue
concerning Experiment 1 was that there were not enough conditions in which
group size was more valuable to the choice trade-off than probability.

Experiment 2 was designed to test whether subjects were better able to
chose the optimal group under conditions in which size was made more salient. A
task similar to that used in Experiment 1 was developed, decreasing the
conditions under which the choice of the higher probability group was the optimal
choice, and increasing the conditions under which the choice of the smaller
probability group was the optimal choice.

Method

Subjects. Twenty-six Rice University undergraduate students participated
in exchange for psychology course credit.

Experimental Design. The probability of locating the target in a group,
group size, and type of search were again manipulated, with a few changes. The
.5/.5 probability level was dropped, leaving the following pairs of probabilities: 0/
1.0, .2/.8, .4/.6. Group size was changed to 3, 9 or 27 items per group and 0/1.0
trials were used to determine the time needed to search these groups. The type of
search variable was not altered.

The trials were constructed to present all combinations of the conditions.
Seventy-two combinations were possible: 6 probabilities presented in the left
group (0, .2, .4, .6, .8, 1.0) x 6 left group/right group size combinations (3/9, 3/27,
9/3, 9/27, 27/3, 27/9) x 2 types of search (directed, non-directed). Three
replications of these combinations were tested (only two replications of the 0/1.0 trials were used, for reasons of economy), resulting in 192 trials per subject. The order of trials was randomly determined for each subject.

Initial group choice, choice time, and search time data were collected as dependent variables.

Stimuli. The layout of the trial stimuli was similar to the layout used in Experiment 1, with the exception that the group columns were longer, in order to accommodate the 27-item groups. Instructions for Experiments 1 and 2 differed only in the specific levels of probability and number of items described in the instruction text and in the illustrations.

Procedure. The experimental procedure, including the computer trials and the post-trials questionnaire, was similar to that used in Experiment 1.

Apparatus. Presentation of the instructions, practice trials, and experimental trials, and the collection of data, were done using HyperCard (version 2.1). In order to accommodate the larger group sizes, experimental sessions were conducted using a Macintosh LC computer with a 13-inch monitor.

Results

Overall. The average percentage of trials in which the subject made an error (either a false alarm or a miss) was 2.8 (range = 0.0 to 7.8 percent). Subjects gave an incorrect target line response in less than 1 percent of the trials (mean percent = .93, range = 0 to 4.5 percent). Incorrect answer trials were deleted from search time analyses.
Choice Order: Determination of Optimal Search. The median search times for 3-, 9-, and 27-item groups in the non-directed condition (under 0/1.0 and 1.0/0 trials) were 1.03, 1.78, and 3.65 seconds, respectively. In the directed condition, the median search times for 3-, 9-, and 27-item groups were 0.83, 0.92, and 1.10 seconds, respectively.

As in Experiment 1, expected initial search times for the left group and right group were calculated for each subject across all of the conditions. Again, there was agreement concerning which of the two groups involved in the trade-off resulted in faster estimated search, and these findings were in agreement with the optimal group based on trade-off ratios (Table 5).
### Table 5. Comparison of Estimated Search Times and Probability to Size Ratios: Experiment 2

<table>
<thead>
<tr>
<th>Condition [left probability(left size)/right probability(right size)]</th>
<th>Ratio of Probability to Size</th>
<th>Expected Search Time (seconds): Left Group</th>
<th>Expected Search Time (seconds): Right Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability faster conditions</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.2(3)/.8(9)</td>
<td>4:3</td>
<td>2.48</td>
<td>1.96</td>
</tr>
<tr>
<td>.2(9)/.8(27)</td>
<td>4:3</td>
<td>4.62</td>
<td>3.94</td>
</tr>
<tr>
<td>Size faster conditions</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.4(3)/.6(27)</td>
<td>1.5:9</td>
<td>3.24</td>
<td>4.02</td>
</tr>
<tr>
<td>.2(3)/.8(27)</td>
<td>4:9</td>
<td>3.95</td>
<td>3.80</td>
</tr>
<tr>
<td>.4(3)/.6(9)</td>
<td>1.5:3</td>
<td>2.13</td>
<td>2.18</td>
</tr>
<tr>
<td>.4(9)/.6(27)</td>
<td>1.5:3</td>
<td>3.90</td>
<td>4.29</td>
</tr>
<tr>
<td>No conflict conditions</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.2(27)/.8(3)</td>
<td>4::11</td>
<td>4.46</td>
<td>1.80</td>
</tr>
<tr>
<td>.4(27)/.6(3)</td>
<td>1.5::11</td>
<td>4.24</td>
<td>2.52</td>
</tr>
<tr>
<td>.2(9)/.8(3)</td>
<td>4::33</td>
<td>2.61</td>
<td>1.43</td>
</tr>
<tr>
<td>.2(27)/.8(9)</td>
<td>4::33</td>
<td>4.99</td>
<td>2.46</td>
</tr>
<tr>
<td>.4(9)/.6(3)</td>
<td>1.5::33</td>
<td>2.40</td>
<td>1.78</td>
</tr>
<tr>
<td>.4(27)/.6(9)</td>
<td>1.5::33</td>
<td>4.64</td>
<td>3.18</td>
</tr>
</tbody>
</table>

There were two exceptions to the general agreement between estimated search times and trade-off ratios. In the .2/.8 trials, the estimated search times favored the initial search of the higher probability group regardless of group size, with the exception of the .2(3)/.8(27) condition. Despite the fact that the trade-off ratio favored a choice based on group size, 20 of the 26 subjects had lower
expected search times to the larger .8(27) group.

In the .4(3)/.6(9) choice, the ratio of probability to size (1.5:3) favored the initial search of the smaller group (i.e., the .4(3) group). However, investigation of the expected search times to this choice revealed that not all subjects had shorter median search times to the .4(3) group: fifteen subjects had lower expected search times given the initial search of the .4(3) group, ten subjects had lower expected search times given the initial search of the .6(9) group, and one subject had equal search times to the initial search of both groups.

When the ratio of probability to size favored the initial choice of one group, but the subject’s expected search time, derived from the subject’s actual performance, was shorter for the initial choice of the other group, search strategy classification was based on the subject’s expected search time.

Choice Order: Choice Patterns. The criterion for consistent choice, as in Experiment 1, was an initial choice of one of the two groups in 83% (i.e., 5/6) of the particular trade-off trials.

As in Experiment 1, subjects were not able to choose the group with the faster expected search time satisfactorily under relatively complicated trade-off conditions, especially conditions favoring the choice of the smaller group (Table 6).

In the analysis of individual subjects, the classifications that emerged in Experiment 1 (optimal, complete probability, strong probability, strong group size, inconsistent) were again apparent. The only change in the classifications was the addition of a near optimal group, defined as consistent choices in all trade-offs
except for the .4(3)/.6(9) condition. Some flexibility was allowed concerning this condition since the optimal choice was not certain: although the ratio favored the initial search of the .4(3) group, some subjects had faster search times to the initial search of the .4(3) group, whereas some subjects had faster search times to the .6(9) group.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Percent of Subjects Who Chose Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Correct</td>
</tr>
<tr>
<td><strong>NON-DIRECTED TRIALS</strong></td>
<td></td>
</tr>
<tr>
<td>Probability faster conditions</td>
<td></td>
</tr>
<tr>
<td>.2(3)/.8(9)</td>
<td>92</td>
</tr>
<tr>
<td>.2(9)/.8(27)</td>
<td>88</td>
</tr>
<tr>
<td><strong>Size faster conditions</strong></td>
<td></td>
</tr>
<tr>
<td>.4(3)/.6(27)</td>
<td>31</td>
</tr>
<tr>
<td>.2(3)/.8(27)</td>
<td>15</td>
</tr>
<tr>
<td>.4(3)/.6(9)</td>
<td>4</td>
</tr>
<tr>
<td>.4(9)/.6(27)</td>
<td>27</td>
</tr>
<tr>
<td><strong>DIRECTED TRIALS (all Probability faster conditions)</strong></td>
<td></td>
</tr>
<tr>
<td>.2(3)/.8(9)</td>
<td>100</td>
</tr>
<tr>
<td>.2(9)/.8(27)</td>
<td>100</td>
</tr>
<tr>
<td>.4(3)/.6(27)</td>
<td>73</td>
</tr>
<tr>
<td>.2(3)/.8(27)</td>
<td>100</td>
</tr>
<tr>
<td>.4(3)/.6(9)</td>
<td>92</td>
</tr>
<tr>
<td>.4(9)/.6(27)</td>
<td>77</td>
</tr>
</tbody>
</table>
Subjects' search strategies under non-directed and directed conditions are presented in Table 7. As in Experiment 1, few subjects could be classified as using the Optimal strategy: five subjects chose optimally or near-optimally in non-directed trials and fourteen subjects choose optimally in the directed trials. Only one of the subjects was optimal in both non-directed and directed trials. Subjects tended to choose either inconsistently or tended to choose based on group probability (in both non-directed and directed conditions). Sixteen of the 26 subjects used the same strategy in both non-directed and directed trials.

**Table 7. Number of Subjects Using Strategies in Non-Directed and Directed Trials: Experiment 2**

<table>
<thead>
<tr>
<th>Strategy in Directed Trials</th>
<th>Strategy in Non-Directed Trials</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Optimal</td>
</tr>
<tr>
<td>Optimal/ Complete Probability</td>
<td>1</td>
</tr>
<tr>
<td>Strong Probability</td>
<td>2</td>
</tr>
<tr>
<td>Size</td>
<td>1</td>
</tr>
<tr>
<td>Inconsistent</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>5</td>
</tr>
</tbody>
</table>

*Note.* In the Directed Trials, a Complete Probability strategy is categorized as an Optimal strategy. Optimal and Near Optimal categories are combined into the Optimal category.
Although the same strategy classifications were apparent in Experiments 1 and 2, the percent of subjects using specific strategies differed between the two experiments. As shown in Table 8, there were fewer inconsistent choosers and more complete probability choosers in Experiment 2.

### Table 8. Percent of Subjects Using Specific Strategies in Experiments 1 and 2

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Experiment 1</th>
<th>Experiment 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-Directed Trials</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Optimal/Near Optimal</td>
<td>8</td>
<td>19</td>
</tr>
<tr>
<td>Complete Probability</td>
<td>20</td>
<td>38</td>
</tr>
<tr>
<td>Strong Probability</td>
<td>24</td>
<td>19</td>
</tr>
<tr>
<td>Size</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>Inconsistent</td>
<td>40</td>
<td>19</td>
</tr>
<tr>
<td>Directed Trials</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Optimal (Complete Probability)</td>
<td>28</td>
<td>54</td>
</tr>
<tr>
<td>Strong Probability</td>
<td>28</td>
<td>19</td>
</tr>
<tr>
<td>Size</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>Inconsistent</td>
<td>36</td>
<td>19</td>
</tr>
</tbody>
</table>

*Choice Time.* The analysis of the choice time data followed the procedure described in Experiment 1, with the exception that one of the levels of the Decision Type factor, Neutral, was not pertinent to the Experiment 2 analysis, since those .5/.5, equal-size conditions were not used in Experiment 2. A 2 (Information Placement: Probability over Items, Items over Probability) x 2 (Highlight: No Highlight, Highlight) x 2 (Decision Type: Agreement, Disagreement)
analysis of variance (ANOVA) revealed a significant interaction between Highlight and Decision Type, $F(1, 24) = 4.62, p < .04$. A plot of this interaction is illustrated in Figure 6. When trials contained a highlighted target, Decision Type did not affect choice time: mean choice times were similar for Agreement and Disagreement trials when highlight was present. However, the mean choice time to non-highlighted Disagreement trials was significantly higher than choice times for non-highlighted Agreement trials (and higher than times for the highlighted Agreement and Disagreement trials).

Unlike the results of Experiment 1, the main effect of Decision Type was not significant, $F(1, 24) = 0.30, p < .59$ (mean choice time to the Agreement trials = 2.67 seconds, mean choice time to the Disagreement trials = 2.70 seconds).

As in Experiment 1, neither Information Placement [$F(1, 24) = 0.42, p < .52$] nor Highlight [$F(1, 24) = 1.54, p < .23$] was significant. Mean choice time to the Probability over Items displays was 2.54 seconds and mean choice time to the Items over Probability displays was 2.81 seconds. Mean choice time to the Non-highlighted trials was 2.72 seconds and mean choice time to the Highlighted trials was 2.65 seconds.

There were no reliable differences (at the $p < .05$ level) among the remaining interactions.
**Figure 6.** Choice time interaction of highlight and decision type.

Questionnaire Responses. Fifteen subjects reported that they combined the probability and size information when choosing between groups. Eighteen subjects reported using the probability information more than the other information in their group choice.

Twelve of the 26 subjects reported that they followed a particular strategy (in particular, choosing the higher probability group) for highlighted trials. A Pearson product-moment correlation was computed in order to determine the
relationship between this self-report of choosing the higher probability group in highlighted trials and the actual use of this optimal strategy. The correlation was not significant, \( r = -0.226, p < .27 \).

All subjects reported that highlight favorably affected their search of a group, again reporting that the search was quicker because the bold target "popped-out" from the other items.

EXPERIMENT 3

Experiments 1 and 2 demonstrated that the majority of subjects chose the higher probability group for initial search, even when inappropriate. Subjects professed to using, and actually used, a probability-driven initial group choice. The present experiment attempted to discover if this behavior is present in a different paradigm: instead of a computer display search for a target, a different task represents the search situation, with the goal of discovering whether the behavior (the choice of the high-probability group in a trade-off) is present even in a different medium.

A paper-and-pencil task required subjects to choose between two objects for initial search. The problems (Appendix D) were written to approximate the trade-off search situations used in Experiments 1 and 2 (i.e., varying probability and cost of search), but generalized to other search situations. Problems included locating information in a book or finding a target marble located in one of two bags of marbles.
Method

Subjects. The same group of twenty-six Rice University undergraduate students participated in Experiments 2 and 3.

Experimental Design. Seven essay-type questions (listed in Appendix D) were created to assess search choice behavior. In each question, subjects were required to decide between two options. In each case, there was a cost associated with the decision (e.g., the number of marbles in a bag, the number of pages in a book) as well as a probability associated with the decision.

Answers to the Search Strategy Questionnaire were collected as the dependent variable.

Stimuli. The Search Strategy Questionnaire was presented to subjects as a two-page, word-processor-produced form with areas between the questionnaire items for responses.

The questionnaire was completed in a cubicle area, furnished with a desk and a chair, separate from the room in which the computer tasks were completed.

Procedure. The questionnaire followed the completion of the tasks described in Experiment 2. Subjects were given unlimited time to complete the questionnaire.

Results

Subjects correctly answered 5 of the 7 items from the Search Strategy Questionnaire. Items 1 and 5 accounted for most of the incorrect answers: 15 of the 26 subjects incorrectly answered Item 1, and 13 of the 26 subjects incorrectly
answered Item 5. A Pearson product-moment correlation was computed to
determine the degree of association between the ability to answer the trade-off
questions in the Questionnaire and the ability to respond optimally to the trade-off
trials described in Experiment 2. This analysis used the number of correctly-
answered Questionnaire items and the number of optimal responses to the .2/.8
and .4/.6 trial types. The correlation was not significant, \( r = 0.272, p < .18 \).

**EXPERIMENT 4**

Experiments 1 and 2 demonstrated that subjects were surprisingly
inefficient at combining the probability, size and highlight information. Few
subjects chose optimally, and most subjects either relied exclusively on group
probability or were inconsistent in their group choice.

The poor performance on these search tasks is especially distressing
given the sample of subjects: even relatively intelligent college-educated subjects
were inept at these tasks.

The surprisingly inefficient search performance demonstrated in
Experiments 1 and 2, and the time benefit of an optimal search strategy, mandate
an exploration of methods to increase the use of such strategies. Experiment 4
attempts to induce this optimal search behavior through training.

Two training methods were used to emphasize the optimal trade-off of
probability and group size information in the present search task. One method
included trials in which group size was emphasized through exaggeration; the
other method consisted of story problems illustrating the trade-off between
probability and cost of search.

Subjects were exposed to one of the two training methods, and then they performed the non-directed search task used in Experiment 2. The directed, target-highlighted trials used in Experiments 1 and 2 were not used in the present experiment, in order to concentrate on the training of the basic, non-directed search task.

The goal of Experiment 4 is to determine whether these training methods increase the number of optimal searchers. Do the strategies used under the exaggeration of group size training transfer to the more complex tasks used in the previous experiments? Can the description of optimal, time-efficient trade-offs influence the same behavior applied to search? It was expected that exposure to these training methods would result in a larger percentage of subjects categorized as using optimal search compared to subjects given no additional training.

Method

Subjects. Fifty-four Rice University undergraduate students participated as subjects in exchange for psychology course credit. Subjects were randomly assigned to one of three conditions: there were 18 subjects in each condition.

Experimental Design. Three training conditions were used: exaggerated exposure, pedagogic exposure, and control group.

In the exaggerated exposure training condition, prior to performing the search task used in Experiment 2, subjects were initially given instructions and a set of trials in which the search procedure was similar to that used in Experiments
1 and 2, with the exception of the probabilities and group sizes manipulated. A much larger range of group sizes was used: 2, 10, and 50. The probabilities associated with each of the two groups in this condition were .4/.6 and .5/.5. The 50-item group was displayed in two columns with twenty-five items in each column, the 10-item group was displayed in two columns with five items in each column, and the 2-item group was displayed in one column with two items.

The exaggerated exposure trials were constructed to present all combinations of the conditions. Thirty-six combinations were possible: 4 probabilities presented in the left group (.4, .5, .5, .6) x 3 group sizes presented in the left group (2, 10, 50 items) x 3 group sizes presented in the right group (2, 10, 50 items). Three replications of these combinations were tested, resulting in 108 exaggerated exposure trials per subject. The order of trials was randomly determined for each subject.

In the pedagogic exposure condition, prior to performing the search task used in Experiment 2, subjects were initially given a set of seven story problems focusing on the trade-off of probability and cost in various problem-solving tasks. Feedback to their answers was provided and included the correct answer and an explanation of the problem and its answer.

Pilot testing revealed that the exaggerated exposure trials required an average of seventeen minutes to complete; in order to ensure that the training time was similar in these two training conditions, subjects were given seventeen minutes in which to read the problems and explanations in the pedagogic exposure condition.
In the control group, as in Experiments 1 and 2, subjects were given no training aside from the standard experimental instructions and practice trials, which were the same as those used in Experiment 2.

Initial group choice, choice time, and search time data were collected as dependent variables.

*Stimuli.* The layout of the search task stimuli was similar to the layout used in Experiment 2.

*Procedure.* In the exaggerated exposure condition, subjects were initially given instructions and the set of exaggerated exposure trials. Following these exaggerated trials, subjects performed the search task originally used in Experiment 2. The only change made to the Experiment 2 search task was the deletion of the directed trials, which resulted in 96 total trials. The strategy questionnaire used in the previous experiments was administered following the search task.

In the pedagogic exposure condition, subjects were presented with a problem, selected one of the two answer options, and were then presented with feedback and explanation. Subjects were encouraged to read and understand the explanation, and could review problems and explanations by navigating through the story problem screens. Subjects were allowed to progress through the problems without supervision, but were told that they would be tested on the seven problems. At the end of the seventeen minutes, the experimenter tested the subject on the seven problems. Following these pedagogic trials, subjects were given instructions and practice on the search task originally used in
Experiment 2 (with the directed trials deleted, as in the above condition), and then performed that task. The strategy questionnaire used in the previous experiments was administered following the search task.

Subjects in the control group were given instructions and practice on the search task originally used in Experiment 2 (with the directed trials deleted, as in the above condition), and then performed that task. Following these trials, subjects completed an additional set of the search task trials. The strategy questionnaire used in the previous experiments was administered following the search task.

An additional task in all three conditions required subjects to search for two incorrect algebraic equations from five options. This task was developed by Ashby (1988). This task was done after the subject had completed the strategy questionnaire associated with the search trials.

Apparatus. Presentation of the instructions, story problems, practice trials, and experimental trials, and the collection of data, were done using HyperCard (version 2.1). Experimental sessions were conducted using a Macintosh LC computer with a 13-inch monitor.

Results

Subjects were categorized based on their responses to the .4/.6 trials as in Experiments 1 and 2: the number of subjects using a particular strategy, separated by training condition, is presented in Table 9.
Table 9. Number of Subjects Using Strategies by Training Condition

<table>
<thead>
<tr>
<th>Training Condition</th>
<th>Strategy</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Optimal</td>
<td>Complete Probability</td>
<td>Strong Probability</td>
<td>Size</td>
<td>Inconsistent</td>
<td>Total</td>
</tr>
<tr>
<td>Exaggerated</td>
<td>6</td>
<td>2</td>
<td>4</td>
<td>0</td>
<td>6</td>
<td>18</td>
</tr>
<tr>
<td>Pedagogic</td>
<td>10</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>6</td>
<td>18</td>
</tr>
<tr>
<td>Control</td>
<td>4</td>
<td>7</td>
<td>2</td>
<td>0</td>
<td>5</td>
<td>18</td>
</tr>
<tr>
<td>Total</td>
<td>20</td>
<td>10</td>
<td>7</td>
<td>0</td>
<td>17</td>
<td>54</td>
</tr>
</tbody>
</table>

Analysis of Variance. In order to determine whether training affected choice behavior, analyses were conducted on conflict trials. Conflict trials, as in Experiment 2, were trials in which the probability and the size information favored different groups and required a more rigorous trade-off; non-conflict trials were those trials in which the high probability group was also associated with the smaller group size.

The dependent measure in the following analyses was the number of correct responses to trials (as in Experiment 2, each subject was given 6 replications of each specific conflict trial).

The correct response data were analyzed with a 3 x 3 (Training Condition x Conflict Trial) mixed analysis of variance (ANOVA). The ANOVA was conducted separately on the .4/.6 and .2/.8 conflict trials. The three levels of conflict trials in the .4/.6 trials were .4(3)/.6(27), .4(3)/.6(9), .4(9)/.6(27); the three levels of conflict trials in the .2/.8 trials were .2(3)/.8(9), .2(9)/.8(27), .2(3)/.8(27).

Mean correct responses as a function of condition are presented in Table
10 for both the .4/.6 and the .2/.8 trials.

**Table 10. Mean Correct Responses to .2/.8 and .4/.6 Trials by Condition**

<table>
<thead>
<tr>
<th>Condition</th>
<th>.2/.8 Trials</th>
<th>.4/.6 Trials</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training Condition</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exaggerated</td>
<td>3.65</td>
<td>3.48</td>
</tr>
<tr>
<td>Pedagogic</td>
<td>4.20</td>
<td>4.19</td>
</tr>
<tr>
<td>Control</td>
<td>4.13</td>
<td>2.28</td>
</tr>
<tr>
<td>Display Size</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3/9</td>
<td>5.00</td>
<td>2.22</td>
</tr>
<tr>
<td>9/27</td>
<td>4.22</td>
<td>3.80</td>
</tr>
<tr>
<td>3/27</td>
<td>2.76</td>
<td>3.93</td>
</tr>
</tbody>
</table>

In the .4/.6 trials, there was a main effect of training, $F (2, 51) = 5.13$, $p < .01$. A Tukey comparison of training condition means showed a significant difference between the pedagogic and the control groups ($p < .01$). The main effect of ratio was significant $F (2, 102) = 34.19$, $p < .0001$. A Tukey comparison of ratio means showed a significant difference between the following ratios (all differences at the $p < .01$ level): 3/9 and 3/27, 3/9 and 9/27. The interaction between training condition and ratio was not significant, $F (4, 102) = 0.46$, $p < .77$, and is illustrated for purpose of comparison in Figure 7.
Figure 7. Interaction of training condition and group size in .4/.6 conflict trials.

In the .2/.8 trials, the main effect of training was not significant, $F(2, 51) = 2.85, p < .07$. The main effect of ratio was significant, $F(2, 102) = 15.18, p < .0001$. A Tukey comparison of ratio means showed a significant difference between the following ratios (all differences at the $p < .01$ level): 3/9 and 3/27, 3/27 and 9/27. The interaction between training condition and ratio was significant, $F(4, 102) = 5.09, p < .001$, and is illustrated in Figure 8.
Figure 8. Interaction of training condition and group size in .2/.8 conflict trials.

Ashby Task Analyses. Responses to the Ashby task were investigated in order to determine whether there was a relationship between training conditions and performance on the Ashby hypothesis selection trials.

The dependent measure in this set of analyses was the number of correct responses to the first three Ashby trials (a correct response was defined as the choice of small equations before the choice of larger equations).

The correct response data were analyzed with a 1-way between-subjects
ANOVA. The main effect of training was not significant, $F(2, 51) = 2.02, p < .14$.

The mean number correct for the exaggerated, pedagogic, and control groups were 1.78, 2.39, and 1.78, respectively. A Dunnett's One-tailed T test comparing the two treatment groups to the control group was not significant, $p > .05$.

DISCUSSION

The efficient way to trade-off two display factors basic to search performance, time to search a group and probability of target location, is not apparent to many subjects. Actually, one of these factors, probability, is overly emphasized. Without proper training, as demonstrated in Experiments 1 and 2, subjects rely on probability information in their initial choice of group. As demonstrated by the comparison of actual and expected search times, this is not the optimal strategy. Probability has shown itself to be a potent but sometimes misleading decision cue.

In the integration of expected search time and target detection probability, the size of the group (and the associated expected search time) is markedly undervalued. Few subjects used a size strategy (3 out of 51 in Experiments 1 and 2 non-directed trials, and 4 out of 51 in Experiments 1 and 2 directed trials). Also, few subjects were able to optimally trade-off the probability and size information (7 out of 51 in Experiments 1 and 2 non-directed trials, and 21 out of 51 in Experiments 1 and 2 directed trials). Size was undervalued even when it was the only unequal factor in the trade-off. In those situations in which probability was equal and only group size differed for the two groups, [the .5(4)/.5(12), .5(4)/
.5(16), and .5(12)/.5(16) conditions], the majority of subjects did not use the size information to their advantage. In those trials, only 15 of the 25 subjects chose the smaller of the two groups.

Decision making research has demonstrated the propensity of subjects to fixate erroneously on certain aspects of a decision making problem. The work of Kahneman and Tversky (1972, 1973) and Tversky and Kahneman (1973) in judgment heuristics has uncovered some of the reasoning processes people use in situations in which decisions are made under uncertainty. From their research, we know that reasoning processes may lead to incorrect problem solutions, and these incorrect solutions result from the use of judgment heuristics such as representativeness and availability. Relevant information (such as prior odds or likelihood) is ignored, and judgments are based on representativeness (or similarity) or on the ease with which subjects can retrieve relevant instances from memory.

Probability information may have been favored because it is an easy to use, and familiar, piece of information. People are accustomed to thinking probabilistically, whether it is in terms of the likelihood of rain stated in weather forecasts, the odds that a sports team will win, or the chances of winning in a lottery game. It is possible that subjects were not accustomed to thinking in terms of the time required to search a group, or of the cost of search. The cost of search information may also have required more effort: subjects may have found it easier to compare the group probabilities to each other, relative to the effort involved in estimating how long it would take to search each of the groups and comparing
these estimates.

In the heuristics research and in the present experiment, the overemphasis on certain cues results in sub-optimal decisions. In some circumstances, the sub-optimal solutions produced by heuristic-based reasoning are not a problem: "when the cost of error is not high, applying more effortful, albeit more normative, judgment procedures may not be worthwhile" (Glass and Holyoak, 1986, p. 361).

However, the extra time required when an incorrect display group is initially searched is a legitimate concern. Given this time cost, the reliance on a probability strategy in display search should not be looked on as a reasonable heuristic.

The cost of error in real-world search tasks can be prohibitive. As demonstrated in Experiments 1 and 2, the extra task time required when an incorrect group was initially searched was in the area of seconds, and given the frequency with which operators search the display screen for information, this loss of seconds may be crucial.

The emphasis on sub-optimal search performance and subject errors may seem to be a pessimistic view, but there is reason for the pessimism. According to Evans (1989): "the consideration is not whether the behaviour conforms narrowly to standard logic but whether or not a person reasoning in a particular manner could be expected to make effective and successful decisions" (p. 126). In the present situations, the operators were not making effective decisions, and these decisions affected the time needed to complete the task.

Effective or not, many subjects' choices follow particular patterns. Subjects
did not choose groups haphazardly or randomly: although some responses to conditions were inconsistent, most subjects chose in such as manner that their pattern of choices fit into rigorous choice categorizations. In addition, patterns were used even when unnecessary. In the conditions in which neither of the two groups was favored (i.e., when group size and group probability were the same for both groups), nearly half of the subjects consistently followed a particular pattern (10/25 chose the Left group and 2/25 chose the Right group).

Regardless of the pessimistic view, the knowledge of the non-optimal strategies displayed in the experiments can still be used constructively. Display designers could take advantage of the probability bias by ensuring that information is consistently located in a specific area on the screen. Training and education could improve group choice decisions: specifically, these methods could emphasize the cost of searching a group and the interaction of group probability and group size.

Training could also strengthen the concept of highlighting as an effective search cue, which seems to be known to most subjects. All of the subjects in Experiments 1 and 2 reported that highlight affected their search of a group, but far fewer subjects (20 subjects in both Experiments 1 and 2) reported that the target highlight information affected which group they searched first (similarly, far fewer subjects used an optimal Complete Probability strategy in the directed trials). Training can emphasize the ability to use highlighted information not only in the search of a target but in the choice of how to search a display. Users should be told that highlighting on a display should not only affect how they search, but
what they search, as well as the order of the search.

The training methods used in Experiment 4 were able to produce modest changes in the number of subjects able to perform optimally: other training methods could be even more beneficial. These research efforts are worthwhile considering the long-term costs of continuing with non-optimal search strategies.
REFERENCES


### APPENDIX A

English Non-Words used as Trial Stimuli (n = 96)

<table>
<thead>
<tr>
<th>cade</th>
<th>dect</th>
<th>bune</th>
<th>vilt</th>
<th>teve</th>
</tr>
</thead>
<tbody>
<tr>
<td>mibe</td>
<td>kuve</td>
<td>dunt</td>
<td>huct</td>
<td>sove</td>
</tr>
<tr>
<td>reze</td>
<td>bape</td>
<td>kelb</td>
<td>hune</td>
<td>murb</td>
</tr>
<tr>
<td>prid</td>
<td>neam</td>
<td>bime</td>
<td>noat</td>
<td>tuse</td>
</tr>
<tr>
<td>mout</td>
<td>fosh</td>
<td>zang</td>
<td>luth</td>
<td>pulk</td>
</tr>
<tr>
<td>lipe</td>
<td>blet</td>
<td>cile</td>
<td>samp</td>
<td>paze</td>
</tr>
<tr>
<td>rimp</td>
<td>hest</td>
<td>fise</td>
<td>hode</td>
<td>cose</td>
</tr>
<tr>
<td>zerd</td>
<td>gake</td>
<td>vute</td>
<td>belk</td>
<td>fant</td>
</tr>
<tr>
<td>rafe</td>
<td>kalt</td>
<td>zumr</td>
<td>hace</td>
<td>lapw</td>
</tr>
<tr>
<td>zote</td>
<td>roke</td>
<td>vone</td>
<td>gine</td>
<td>mact</td>
</tr>
<tr>
<td>vesk</td>
<td>loim</td>
<td>gude</td>
<td>vard</td>
<td>rald</td>
</tr>
<tr>
<td>kofe</td>
<td>pone</td>
<td>flim</td>
<td>kish</td>
<td>krim</td>
</tr>
<tr>
<td>zile</td>
<td>dort</td>
<td>shev</td>
<td>sint</td>
<td>hild</td>
</tr>
<tr>
<td>fuce</td>
<td>cept</td>
<td>sume</td>
<td>tice</td>
<td>gert</td>
</tr>
<tr>
<td>mern</td>
<td>darp</td>
<td>tane</td>
<td>mosh</td>
<td>scod</td>
</tr>
<tr>
<td>nump</td>
<td>gope</td>
<td>dilk</td>
<td>clum</td>
<td>vike</td>
</tr>
<tr>
<td>dren</td>
<td>bown</td>
<td>zhud</td>
<td>peld</td>
<td>thop</td>
</tr>
<tr>
<td>trag</td>
<td>leat</td>
<td>lart</td>
<td>nerb</td>
<td>runk</td>
</tr>
<tr>
<td>ferd</td>
<td>nauk</td>
<td>chib</td>
<td>nide</td>
<td>grem</td>
</tr>
<tr>
<td>pite</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX B

Instructions to Subjects

For the next few minutes, review these instructions. They will explain what you will be doing during the experiment. Don't worry about trying to memorize the instructions: you will get some practice trials before starting the actual experiment.

Your main task in this experiment is to search a display group for a target, and report the line number of the target. Targets will be strings of four-letter non-words; for example, “dilk” and “kuve.” Each trial will have a new target. The target will be displayed in a box at the top of each screen (in case you forget).

At the start of each trial you will see two areas, representing two display groups, Group T and Group Y. One of the two groups will contain the target. There are no surprises in the experiment: there is always one target in each trial, and the target will be in only one of the groups.

You will be given several pieces of information about each group. One of these pieces of information is probability: this tells you the probability that the target is included in a group. The probabilities will be 0, .2, .4, .5, .6, .8 and 1.0. (For example, a .4 probability is the same as a 40% chance.) The probabilities will sum to 1.0 in each trial: for example, if Group T has a .4 chance of containing the target, Group Y must have a .6 chance.

Another piece of information is number of items. Groups will have either 4, 12, or 16 items (i.e., non-words). This information gives you an indication of how many items you may need to scan in order to locate the target.
On half of the trials, the target will be marked by **highlight** (i.e., the text of
the target will be bolded, to make it stand apart from the other items). To indicate
that the target will be highlighted in a trial, the target will be presented in bold text
at the beginning of the trial. Similarly, to indicate that there will be no highlight in a
trial, the target will be presented in regular text.

When there is highlight present, both groups will have one highlighted item:
in the target group, the target will be highlighted, and in the group without the
target, a distractor item will be highlighted.

Use these pieces of information (probability, number of items, target
highlighted/not highlighted) to choose the group that you think contains the target.
After you have made your choice, **press the corresponding key (T or Y) on the**
keyboard. These keys are labeled on the keyboard.

After choosing a group, you will see the group items, and you can search
for the target at this point.

When you have found the target, **press any key on the keyboard (except**
**for the Return key)** as soon as possible to indicate that you have found the
target. Make sure that you know the line number of the target, because as soon
as you press a key, the group display will be replaced by a new screen.

This new screen will have a space for you to type the line number of the
target. If your answer is wrong, you will be told immediately. **Type the line**
**number carefully; a mistyped answer will be seen as an error. **Accuracy, as well
**as speed of search, is important.**

This display will also show the time it took for you to complete your search
for the target. After your search time is displayed, you will automatically proceed to a new trial.

There will be times (unless you are all-knowing) when the group you initially searched does not contain the target. If, after searching the group, you don’t find the target, you can press the Return key to bring you to the other group. This will present the other group, and you can search this group for the target. Signal that you have located the target, as before, by pressing any key (except Return).

If you miss a target, or if you respond "Found" when a target is not present, you will get an error message. You should not make any errors since the task is simple and you should try to be accurate.

Actually, the experiment procedure is fairly easy, since there are only a few displays that you will be presented with:

• the initial display, containing information about Group T and Group Y;
• the search display, presenting the contents of either Group T or Group Y;
• the answer display;
• the error display (if necessary).

When you have completed all of the trials, a message will automatically come up, informing you that you have finished the trials, and instructing you to find the experimenter.

Remember, work as accurately and as quickly as possible.

When you are ready to begin the practice trials, first press the Shuffle button and then hit the Return key.
APPENDIX C

Group Choice and Group Search Questionnaire

**Group Choice**

1) When choosing which group to search first, did you combine the three pieces of information (probability, group size, target highlight)? If so, how did you combine the information?

2) Did you use one piece of information (probability, group size, target highlight) more than others in choosing a group? If so, specify.

**Group Choice/Choice Time**

3) When deciding which group to search first, were some choices easier or harder to make than others? Which choices?

**Search Time**

4) How did you search for the target once the actual group was shown?

**Highlight/Direct Search**

5) When you were told the target was highlighted, did that affect which group you searched first? If so, how?

6) Did highlight affect how you searched a group? If so, how?
APPENDIX D

Search Strategy Questionnaire

1) There are 12 coins in Bag A and 16 coins in Bag B. One of these coins is gold, and the remainder are copper. You are trying to find the gold coin. There is an equal chance that any of the coins is the one that you are trying to find. Which bag would you search first? Why?

2) There are two bags of marbles in front of you, each containing the same number of marbles. You are looking for a black marble. There is a 60% chance of finding the black marble in Bag X, and a 40% chance of finding the black marble in Bag Y. Which bag do you start looking through first? Why?

3) You are given two bags of black marbles. One bag (Bag A) holds 12 black marbles and the other bag (Bag B) holds 16 black marbles. An unbiased judge flips an unbiased coin, with the plan to replace a black marble from one of the bags with a gold marble, depending on the coin toss. If the coin toss results in "heads", the gold marble will replace one of the black marbles in Bag A. Likewise, if the coin toss results in "tails", the gold marble will replace one of the black marbles in Bag B. A black marble from one of the bags is replaced by the gold marble based on the coin toss: there is a 50/50 chance that the gold marble is in Bag A or Bag B. Your task is to find the gold marble. Which of the bags would you open first? Why?

4) You are driving your car through an unfamiliar part of the city. You've reached a busy four-way intersection and decide to make a turn. You know that you should turn left or right, but you can't remember which direction is correct. Do you make a left turn or a right turn? Why?
5) You are working on a take-home, open-note exam. You are in a hurry, since you are running out of time. The exam question requires you to find a specific equation. You have seen that specific equation before, but cannot exactly remember where. You are sure that it is in one of two notebooks (Notebook 1 and Notebook 2). You are fairly confident that the equation is in Notebook 2 (you guess that there is a 70% chance that the equation is in Notebook 2, and a 30% chance that the equation is in Notebook 1). Notebook 1 has 50 pages and Notebook 2 has 200 pages. Which notebook do you scan first? Why?

6) You are nearly finished working on the exam. You are on your last question, and time is limited. This last exam question requires you to find a specific equation. You have seen that specific equation before, but cannot exactly remember where. You are sure that it is in one of two books (Book 1 and Book 2). You are fairly confident that the equation is in Book 1 (you guess that there is a 70% chance that the equation is in Book 1, and a 30% chance that the equation is in Book 2). Book 1 has 1000 pages and Book 2 has 500 pages. Both books have a table of contents and an index. Which book do you scan first? Why?

7) You are doing a search task in which you know that the target is always on line #5 of the group list. You have a choice between two groups for your search, A and B. In the correct group, the target is on line #5, and in the wrong group, a distractor is on line #5. One of the groups is more likely to contain the target than the other. How would the size of the list affect your choice of which group list to first search?