# The Strategic Control of Gaze Direction in the Tower of London Task

**Timothy L. Hodgson and Adnan Bajwa**
Imperial College School of Medicine, London

**Adrian M. Owen**
University of Cambridge

**Christopher Kennard**
Imperial College School of Medicine, London

## Abstract

■ In this paper, we describe a novel approach to the study of problem solving involving the detailed analysis of natural scanning eye movements during the "one-touch" Tower-of-London (TOL) task. We showed subjects a series of pictures depicting two arrangements of colored balls in pockets within the upper and lower halves of a computer display. The task was to plan (but not to execute) the shortest movement sequence required to rearrange the balls in one half of the display (the Workspace) to match the arrangement in the opposite half (the Goalspace) and indicate the minimum number of moves required for problem solution. We report that subjects are more likely to look towards the Goalspace in the initial period after picture presentation, but bias gaze towards the Workspace during the middle of trials. Towards the end of a trial, subjects are once again more likely to fixate the Goalspace. This pattern is found regardless of whether the subjects solve problems by rearranging the balls in the lower or upper visual fields, demonstrating that this strategy correlates with discrete phases in problem solving. A second experiment showed that efficient planners direct their gaze selectively towards the problem critical balls in the Workspace. In contrast, individuals who make errors spend more time looking at irrelevant items and are strongly influenced by the movement strategy needed to solve the preceding problem. We conclude that efficient solution of the TOL requires the capacity to generate and flexibly shift between control sets, including those underlying ocular scanning. The role of working memory and the prefrontal cerebral cortex in the task are discussed. ■

## INTRODUCTION

During natural behavior, we move our gaze around a complex visual environment actively searching for information relevant to current motivations and goals. A number of studies, which have recorded eye movements during cognitively demanding tasks, have found that particular activities are associated with specialized gaze-shifting strategies (Hayhoe, Bensinger, & Ballard 1997; Land & Furneux, 1997; Carpenter & Just, 1984; Groner, Walder, & Groner, 1984; Yarbus, 1967; Buswell, 1935). In this work, we describe the natural gaze movements, which occur during a task used in the neuropsychological assessment of problem solving and "executive" function in brain-damaged patients.

The Tower-of-London (TOL) task was developed with the aim of testing the subtle deficits in behavior, which are observed following frontal-lobe damage in humans (Shallice, 1982). The test involves the presentation of two different arrangements of colored discs or balls (Figure 1). The subject's task is to rearrange the first array of balls (referred to in this paper as the Workspace) so that it matches the second array of balls (the Goalspace) using the minimum number of moves possible. The positioning of the balls is constrained to the location of three pegs or pockets in each half of the display. Because of this, complex problems demand that the sequence of moves is carefully planned in advance before attempting the first move. Failure to engage in advanced planning of the sequence will result in initial moves blocking subsequent ball moves. Owen et al. (1995) have also described a "one-touch" version of the TOL in which the incentive for individuals to plan solutions in advance is enhanced still further. In this variant of the task, subjects are required to inspect the problems visually and then make a single motor response to indicate how many moves would be required to reach an ideal solution. In this way, the one-touch task isolates the cognitive planning component of the test by demanding the internal planning of solutions without actually executing the appropriate moves.

The brain areas involved in planning solutions to TOL problems have been studied extensively using

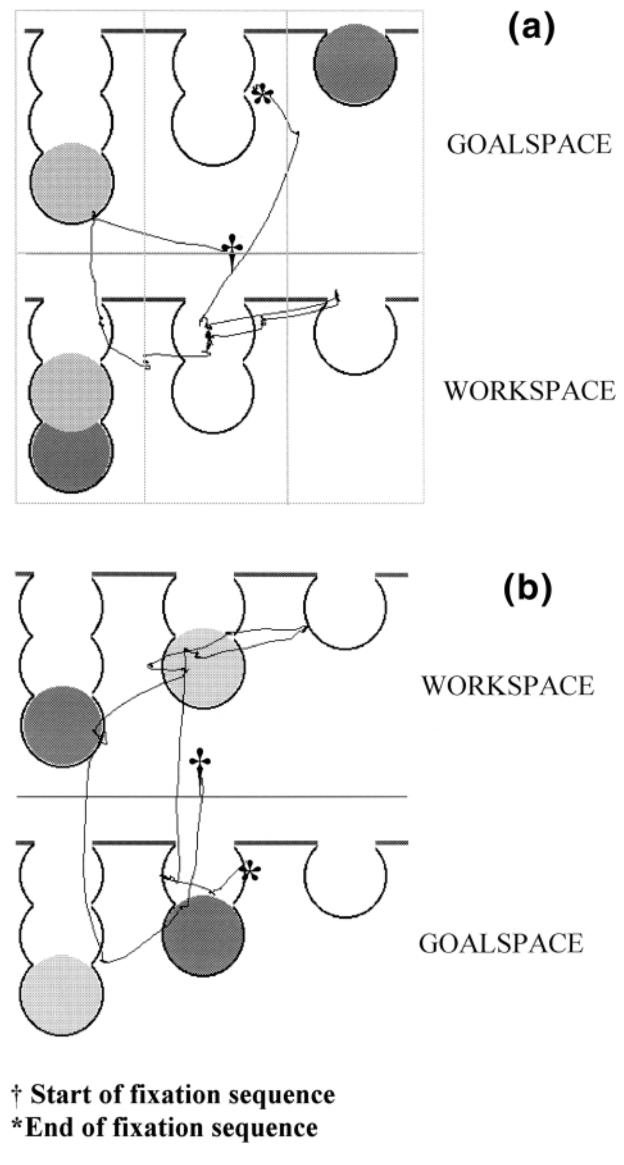

**Figure 1.** Example *X-Y* gaze position plots, superimposed over the relevant TOL pictures from Experiment 1. (a) Data from a single trial for a "Downstairs" subject who was instructed to rearrange the lower arrangement of balls to match the upper (analysis grid superimposed). (b) An "Upstairs" subject who had to plan the moves required to rearrange the upper balls to match the lower balls. Fixations tended to land on the goalspace towards the beginning and end of the trials, regardless of whether this was located in the upper or lower visual fields.

neuropsychological and neuroimaging techniques. In normal subjects, the time taken to plan successful solutions, as well as the number of errors made in executing the final sequence, increases with the minimum number of moves required to solve problems. Shallice (1982) found that patients with frontal cortex pathology required an increased number of moves to solve the TOL. This deficit could not be attributed to a simple visuospatial abnormality, as the same patients successfully completed a block-copying task, which involved reconstructing an arrangement of blocks on an unconstrained workspace and therefore lacked the requirement to plan moves in advance. Other research has confirmed Shallice's findings of frontal cerebral cortex involvement in the TOL (Owen, Downes, Sahakian, Polkey, & Robbins 1990), as well as revealing a slowing of the time taken to plan solutions in Parkinson's disease patients, implicating dopaminergic pathways in the processes underlying cognitive planning (Owen et al., 1995). Meanwhile, functional neuroimaging studies have provided support for neuropsychological findings by identifying the brain regions that become active during cognitive planning in normal subjects (Dagher, Owen, Boecker, & Brooks, 1999; Baker et al., 1996; Owen, Doyon, Petrides, & Evans, 1996; Morris, Ahmed, Syed, & Toone, 1993). Selective activation has been described while subjects perform complex TOL problems in the mid-dorsolateral-prefrontal cortex, a region which has been associated with the manipulation of task relevant information in working memory (Owen et al., 1996; Funahashi, Bruce, & Goldman-Rakic, 1989).

The role of working memory in the TOL and problem solving is a major issue. The Working Memory model proposed by Baddeley (Baddeley, 1988; Baddeley & Della Sala, 1996) predicts that the limiting factor restricting problem-solving performance is the processing capacity of a central working memory system. Planning solutions to more complex problems requires the maintenance of a memory for the planned sequence of moves, as well as current behavioral goals and problem states. In contrast, other accounts suggest that the limitations of performance are not primarily due to constraints imposed by working memory resources, but arise from difficulties in selecting between competing behavioral goals (e.g., Allport, Styles, & Hseih, 1994).

A number of computerized models of problem solving have been constructed, which are capable of performing the TOL and other related tasks (Dehaene & Changeux, 1997; Andersen, 1993; Polson & Jeffries, 1982; Newell & Simon, 1972). One particularly influential model has been the Soar production system (Newell, 1990). Importantly, there is no theoretical limit on "working memory" capacity in this model. Soar has no difficulty in maintaining current goals, the current problem state and other task relevant information on-line during problem solving. However, the program does run into difficulty when there is no clear choice of which move is the most beneficial to make in a given context. Under these conditions, the program creates a temporary subgoal to resolve the conflict. This situation arises during the TOL task when moving a ball directly into its target location blocks subsequent moves. Correct solution of these problems necessitates a "shunting" maneuver to place the ball in a temporary (i.e., subgoal) location.

The manner in which models like Soar might be implemented within biological systems remains to be specified. In particular, it is unclear how the unrestricted working memory capacity available to problem solving machines can be reconciled with the limited short-term memory capacity of humans. Despite extensive functional neuroanatomical studies of the TOL task, there have been relatively few detailed behavioral investigations of problem-solving performance in normals (Ward & Allport, 1997, are a notable exception).

The importance of such studies are emphasized by the "embodied cognition" perspective recently outlined by a number of authors (Ballard, Hayhoe, Pook, & Rao, 1997; Boden, 1997; Clark, 1997; Churchland, Ramachandran, & Sejnowski, 1994; Kirsh & Maglio, 1994; Brooks, 1991). According to this view, problem solving essentially involves the coordination of motor behavior, which modulates sensory input from the environment in order to reduce the computational complexity of the problem faced by the organism. A simple example, which neatly illustrates the point, involves the board game Scrabble® (Clark, 1997; Kirsh, 1995). Those familiar with the game will recall that while pondering potential words to play, we shuffle and rearrange the letter tiles in front of us. This behavior is not extraneous to getting a high score, but serves to prompt new word associations, reducing the load placed on internal resources by the letter conundrum (Kirsh, 1995). An upshot of this active component of cognition is that the detailed analysis of behavioral strategies (e.g., hand and eye movements) during cognitive tasks may provide a window onto what have traditionally been thought to be hidden mental processes.

In this paper, we use a novel approach to the study of problem solving in which we analyze the pattern of natural scanning eye movements made by subjects while performing the "one-touch" TOL task. We initially proposed that subjects would inspect problems in a strategic manner and that it would be possible to identify discrete components of gaze control corresponding to the acquisition of the stimulus configuration and elaboration of problem solutions. We contrasted this with an alternative model of gaze control in which the eyes do not scan strategically and selectively, but instead constantly sample and resample salient information in the visual field. If this later model were correct, eye-movement patterns should be independent of the instructions given to subjects while viewing TOL problems and should depend solely on the salient features of the visual stimulus.

## EXPERIMENT 1

### Results

#### *Errors*

Error rates in the experimental block were very low (<1% of problems overall) and errors were not analyzed further.

#### *Total Response Times*

The total response (i.e., "planning") times were analyzed using a two-way analysis of variance (ANOVA) with subject group (Upstairs/Downstairs) and sequence length (one, two, or three) as factors. As expected, response times increased with sequence length ($F(2,14) = 36.74, p < .0001$. Mean response times for sequence lengths one, two, and three: 1,569, 2,457, 3,365 msec, respectively). There was no significant effect of subject group on total response times.

#### *Fixation Dwell Time*

The total fixation time per trial was calculated for each sector of the 3 × 2 analysis grid. These values were then analyzed using a three-way ANOVA with subject group (Upstairs/Downstairs), visual field (upper/lower) and sequence length (1, 2, or 3) as factors. This analysis revealed a significant three-way interaction between-subject group, visual field, and sequence length ($F(2,12) = 26.18, p < .0001$). As the difficulty of the problems increased, subjects spent proportionately longer viewing the respective Workspace, regardless of whether this was in the upper or lower visual field. However, time spent fixating the Goalspace remained approximately constant as the difficulty of the planned sequence increased (Table 1).

More detailed analysis of this interaction revealed that it was due to a modulation in the total number of fixations made (moves × location × subject group: $F(2,12) = 11.40, p < .005$), as well as a weaker effect on the duration of individual fixations between saccades (moves × location × subject group: $F(2,12) = 4.63, p < .05$).

We also examined how total time fixating the three horizontal sectors varied with move difficulty. This analysis revealed that as the complexity of the problems increased fixations became more biased towards the central location (moves × location $F(4,24) = 9.63, p < .001$). A bias towards the left-hand location was also observed, reflecting the increased number of locations on the left side of the problem pictures (Table 2).

Another analysis examined whether subjects spent longer looking at ball locations compared to empty locations. More time per trial was spent fixating single-ball locations relative to empty locations. Almost exactly twice as long was spent fixating locations containing two balls compared to pockets with a single ball ($F(2,14) = 25.16, p < .0001$) (Table 3).

#### *Temporal Order of Fixations During a Trial*

We examined the probability of fixations landing on the Goalspace ($P$(Goal)) within different 500-msec time bins from the onset of the picture stimulus. This analysis showed that for simpler problems, fixations were equally likely to fall on the Goalspace and the Workspace arrangement throughout the course of a typical trial.

**Table 1**. Mean Time Per Trial (msec) Spent Fixating the Upper and Lower Visual Fields Versus Subject Group and Minimum Number of Moves (Experiment 1)

| | *Upstairs* | | | *Downstairs* | | |
|---|---|---|---|---|---|---|
| *Visual Field* | *One* | *Two* | *Three* | *One* | *Two* | *Three* |
| Upper field | 1158 ± 220 | 2054 ± 342 | 3163 ± 383 | 852 ± 247 | 1242 ± 381 | 1223 ± 331 |
| Lower field | 738 ± 216 | 1225 ± 235 | 962 ± 189 | 728 ± 123 | 942 ± 63 | 2172 ± 190 |

However, for three-move problems, a bias was observed towards fixating the Goalspace at the beginning and the end of the trial, but during the middle of the trial, fixations more commonly landed on the Workspace ($F(6,42) = 3.94, p < .005$) (Figure 1 and Figure 2a).

In order to confirm that this pattern of eye movements was task dependent, a second analysis assessed the probability of the Upstairs and Downstairs groups fixating the upper visual field ($P$(upper)). This showed a clear dissociation between the Upstairs and Downstairs groups (time × group, $F(3,36) = 4.61, p < .005$) and a bias towards fixating the Goalspace at the beginning and end of the trial independent of whether the subjects were in the Upstairs or Downstairs group (Figure 1 and Figure 2b).

### *Saccadic Shifts Between Locations*

Finally, we examined the probability of transformations in gaze position between different regions of the display. The probability of a given fixation being followed by another fixation at a grid location horizontally adjacent to it ($P$(horizontal shift)) was calculated and analyzed according to whether the two fixations occurred in the upper or lower visual field. This analysis showed that the probability of making a horizontal shift in fixation increased with the length of the problem sequence, and that there was a dissociation in the proportion of these eye movements measured between visual fields and subject groups. More horizontal transformations in fixation were recorded in the Workspace area of the display, independent of whether this was in the upper or lower visual field (subject group × visual field × sequence length: $F(2,12) = 4.87, p < .05$) (Table 4; Figure 1).

Finally, the probability of making a vertical shift in fixation between the upper and lower visual field was assessed relative to sequence complexity. No significant increase in the probability of making a vertical shift in gaze between the two hemifields was observed as problem difficulty increased ($F(2,12) = 1.42$).

**Table 2**. Mean Time Per Trial Fixating in Each of the Horizontal Sectors in Experiment 1

| *Sequence Length* | *Left* | *Middle* | *Right* |
|---|---|---|---|
| One | 436 ± 91 | 648 ± 102 | 163 ± 75 |
| Two | 449 ± 105 | 1280 ± 176 | 529 ± 75 |
| Three | 868 ± 153 | 1776 ± 373 | 346 ± 72 |

## Discussion

For both Upstairs and Downstairs subject groups, the total fixation time on the Workspace increased with the complexity of the problem. In contrast, the total time per trial spent looking at the Goalspace remained approximately constant. Changes in the proportion of fixations landing in the two regions of the display also occurred over the course of a trial. For three-move problems, an initial period spent inspecting the Goalspace was followed by a bias towards fixating the Workspace. Towards the end of a trial, more fixations were found to land on the Goalspace once again.

The qualitative reversal in the distribution of gaze between the Upstairs and Downstairs subject groups confirms that efficient solution of TOL problems involves the coordination of an appropriate gaze-shifting strategy. The nature of this strategy suggests that solving three move TOL problems proceeds in several discrete phases. An initial assessment of the problem is followed by a solution elaboration phase in which different operations are rehearsed and assessed. Finally, there is a verification phase, during which an internal representation of the planned solution is compared to the goal configuration. In contrast, problems requiring one or two moves proceed via a direct comparison between the upper and lower arrangements of balls.

Subjects were found to spend longer fixating the Workspace as problem complexity increased, but there was no similar increase in the time spent fixating the

**Table 3**. The Effect of the Contents of Each Location on the Mean Total Dwell Times in Experiment 1 (Single-Move Problems)

| *Location Type* | |
|---|---|
| Single ball | 281 ± 34 |
| Double ball | 554 ± 25 |
| Empty | 100 ± 77 |

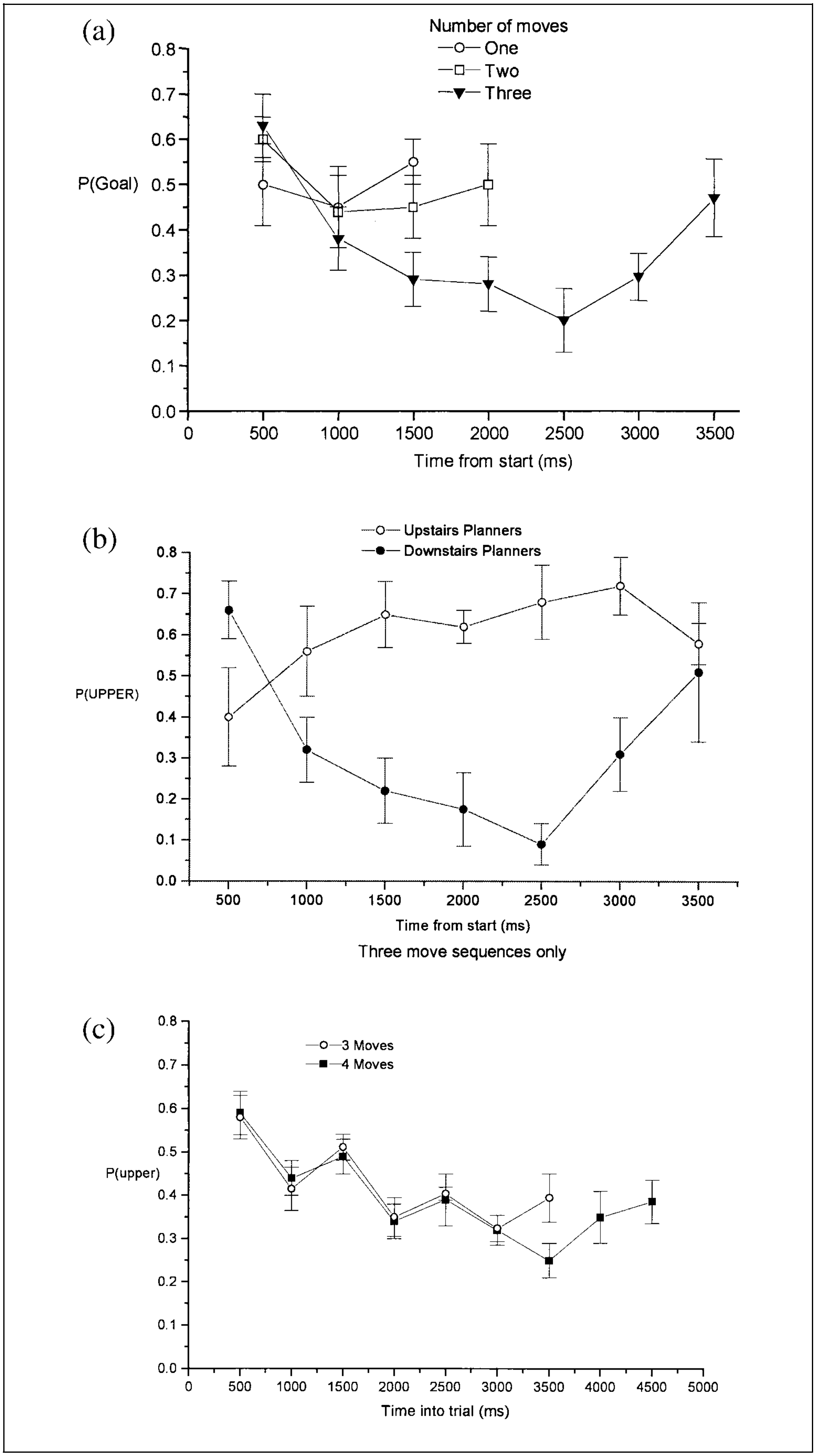

**Figure 2.** Graphs showing the probability of fixations landing in the upper and lower fields during 500-msec time bins from the onset of the problem pictures. (a) Probability of fixations landing on the Goalspace, Experiment 1 (all sequence lengths collapsed across subject group). (b) Probability of fixations landing in the upper visual field for both subject groups, Experiment 1 (three-move problems only).
(c) Probability of fixations landing in the upper field, Experiment 2.

**Table 4**. Probability of a Given Fixation Being Followed by a Horizontal Shift in Fixation Between Sectors in the Analysis Grid (Three-Move Problems, Experiment 1)

| | *Subject Group* | |
|---|---|---|
| *Visual field* | *Upstairs* | *Downstairs* |
| Upper | 0.11 ± 0.04 | 0.32 ± 0.03 |
| Lower | 0.23 ± 0.04 | 0.14 ± 0.01 |

Goalspace. One interpretation of this observation is that viewers use direct fixation to acquire task relevant information from the Goalspace and then hold this information in a memory buffer during solution elaboration (Hayhoe et al., 1997; Land & Furneaux, 1997). An alternative explanation for this effect is that the Goalspace is constantly monitored in parafoveal vision via covert attention, even when it is not being fixated directly. We cannot distinguish between these alternatives without further experimentation (e.g., by increasing the spatial separation between the two arrangements of balls), although studies of block copying, which have manipulated the color of items located away from the point of gaze imply that covert attention is not used in this manner (Hayhoe et al., 1997). The finding that return fixations to the Goalspace occurred at the end of trials also suggests that subjects did not rely exclusively on either parafoveal vision or memory of the goal arrangement. Instead, subjects place little faith in the accuracy of nonfoveal representations of current goals, preferring instead to refixate the Goalspace to confirm that problems had been correctly solved at the end of each trial.

Some quantitative differences were observed between Upstairs and Downstairs groups. For example, the increase in Workspace fixation time with problem difficulty was more pronounced for the Upstairs group (Table 1). It is difficult to interpret these differences because Downstairs subjects were given additional practice using a computer touch-screen version of the TOL task, for which balls are actually moved on screen and problems are always solved in the Downstairs manner. In order to avoid an interference effect between these two tests, the Upstairs group was not exposed to this version of the task. This difference in practice regimes may explain the quantitative dissociation between subject groups. A more speculative explanation is that the lower and upper visual fields may have different functional specializations for coordinating actions within peripersonal space and searching extrapersonal visual space, respectively (Previc, 1990). However, even if practice conditions were matched, comparison between the two groups is complicated by the fact that the difficulty of TOL problems is noncommutative with respect to transformations between the two ball arrangements (Ward & Allport, 1997). Reversing the arrangements of balls in the Goal and Workspace can affect the difficulty of problems

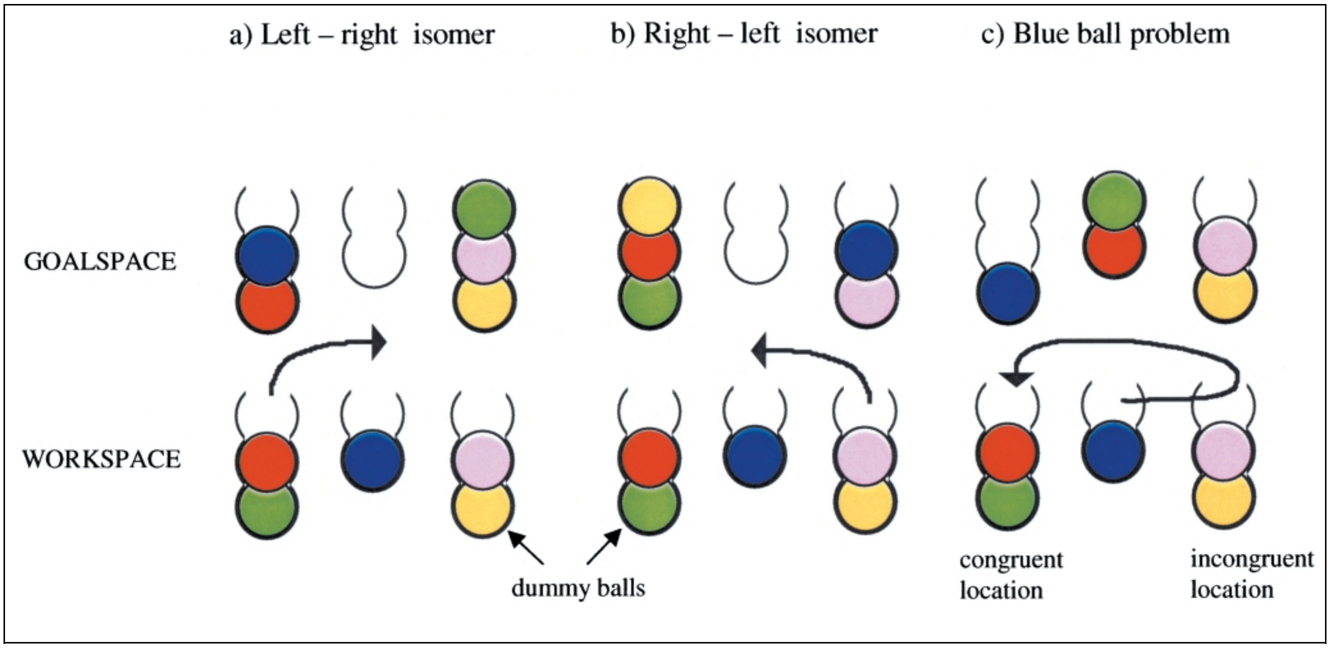

**Figure 3.** Schematic description of the different problem types in Experiment 2. All problems contained five balls in the upper and lower fields and subjects always solved problems in the "Downstairs" manner. The arrangement of balls in the lower visual field was the same in all 24 problems. Only the arrangement of balls in the upper visual field was changed from trial to trial. (a, b) For each problem, there were two "dummy" balls, which were irrelevant to the problem solution, creating left-right and right-left "isomers" of the same problem. (c) "Blue ball problems" were defined as those for which successful solution demanded that the blue ball had to be moved to a temporary, subgoal location before other moves were attempted. Mean gaze times were measured within the left, middle, and right workspace locations for each type of problem.

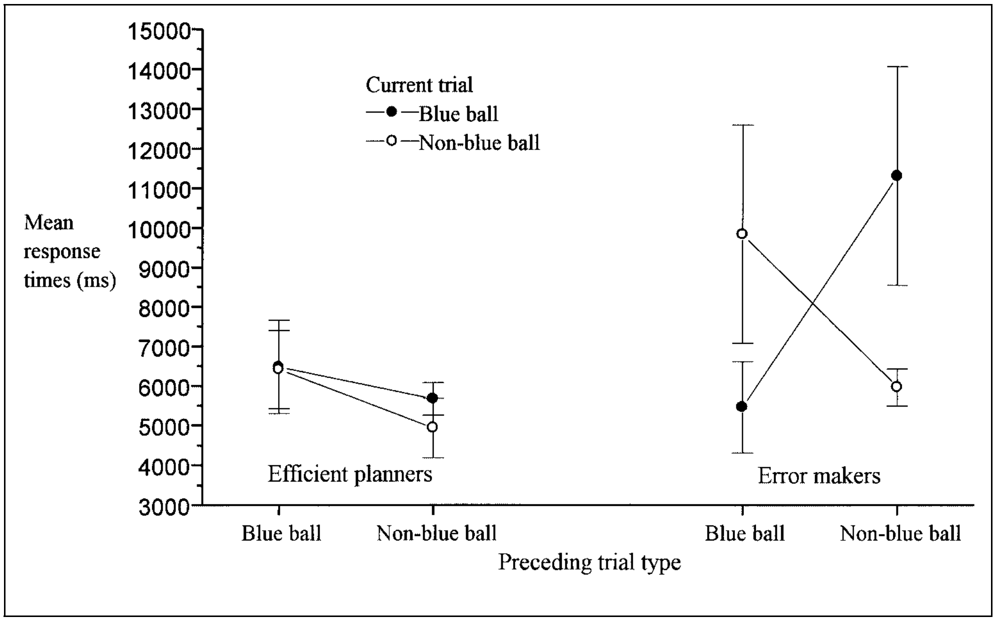

**Figure 4.** Mean response times (Experiment 2) for both subjects groups on Blue ball and Nonblue ball problems, showing the effect of the preceding trial type. Response times in the Error-maker group were increased when trials were preceded by a problem of a different type.

irrespective of the visual field within which they are presented. None of these differences in experimental conditions between subject groups can convincingly explain the marked qualitative reversal in gaze strategies observed dependent upon instruction set.

Aside from vertical shifts in gaze direction between the two visual fields, horizontal transformations in fixation between locations were also observed. These occurred more frequently in the Workspace relative to the Goalspace, indicating that lateral shifts in fixation were correlated in some way with the planning of problem solutions (Figure 1; Table 4). Although we have not been able to establish a one-to-one correspondence between lateral gaze shifts and the ball movements required to solve problems, it seems likely that at least some of these eye movements arise from the operation of visuomotor imagery during the elaboration of problem solutions (Brandt & Stark, 1997). If gaze shifts did reflect cognitive rehearsal of problem solutions, then, as well as being task specific, we would expect gaze strategies to be problem specific and vary according to the actions demanded by particular solutions. The pattern of fixations would also be expected to influence how quickly and accurately the task was performed. These hypotheses were difficult to test using the data collected in Experiment 1 because error rates were very low and gaze direction was found to be strongly modulated by the location of the balls in the display, as well as being centrally biased during more complex problems (Tables 2 and 3).

In contrast to our initial investigation, Experiment 2 used carefully designed stimuli to test whether the distribution of gaze within individual TOL problem pictures could in any way be described as problem dependent. This was achieved by keeping the arrangement of balls in the Workspace constant and manipulating only the arrangement of balls in the Goalspace from problem to problem (Figure 3). In this way, differences between problems in the distribution of gaze on the Workspace could be directly attributed to planning of problem solutions, rather than being attributable to simple differences in the distribution of objects in the display.

## EXPERIMENT 2

### Results

*Errors*

There was no significant difference in the number of errors made by the Practiced and Unpracticed subjects ($F(1,8) = 2.15$, $p = .18$; Practiced: 1.5 ± 0.8; Unpracticed: 4.75 ± 2.5). However, half of the 10 subjects made more than five % errors during the experimental block. In the following analysis, the subjects were divided into two groups according to whether they made more than five % errors in the experimental block. These groups were named Efficient planners and Error makers.

*Total Response Time*

The total response time increased significantly with sequence length ($F(2,16) = 24.65$, $p < .0001$). No

**Table 5.** Mean Total Time Per Trial Fixating Upper and Lower Visual Fields Against Minimum Number of Moves (Experiment 2)

| | *Number of Moves* | | |
|---|---|---|---|
| *Visual Field* | *Two* | *Three* | *Four* |
| Upper | 1428 ± 117 | 2175 ± 224 | 2542 ± 288 |
| Lower | 1765 ± 158 | 3102 ± 480 | 4708 ± 362 |

significant difference in total RT was found between the Practiced and Unpracticed groups ($F(1,8) = 3.36, p > .1$) or Error makers and Efficient planners ($F(1,8) = 1.08, p = .33$). However, an interaction was found between the effect of the current problem type (Blue/Nonblue ball) and the preceding problem type. Response times were increased when a Blue ball trial was preceded by a Nonblue ball trial and vice versa (trial type × preceding trial type interaction: $F(1,6) = 15.53, p < .01$). This interference effect was only found for the Error-maker group and was not evident in the response times of the Efficient planners (trial type × preceding trial type × subject group: $F(1,6) = 11.74, p < .01$) (Figure 4).

### Fixation Dwell Time

As in Experiment 1, the total time per trial spent fixating the Workspace increased with sequence length, but the time fixating the Goalspace did not (visual field × sequence length interaction: $F(2,16) = 18.73, p <$

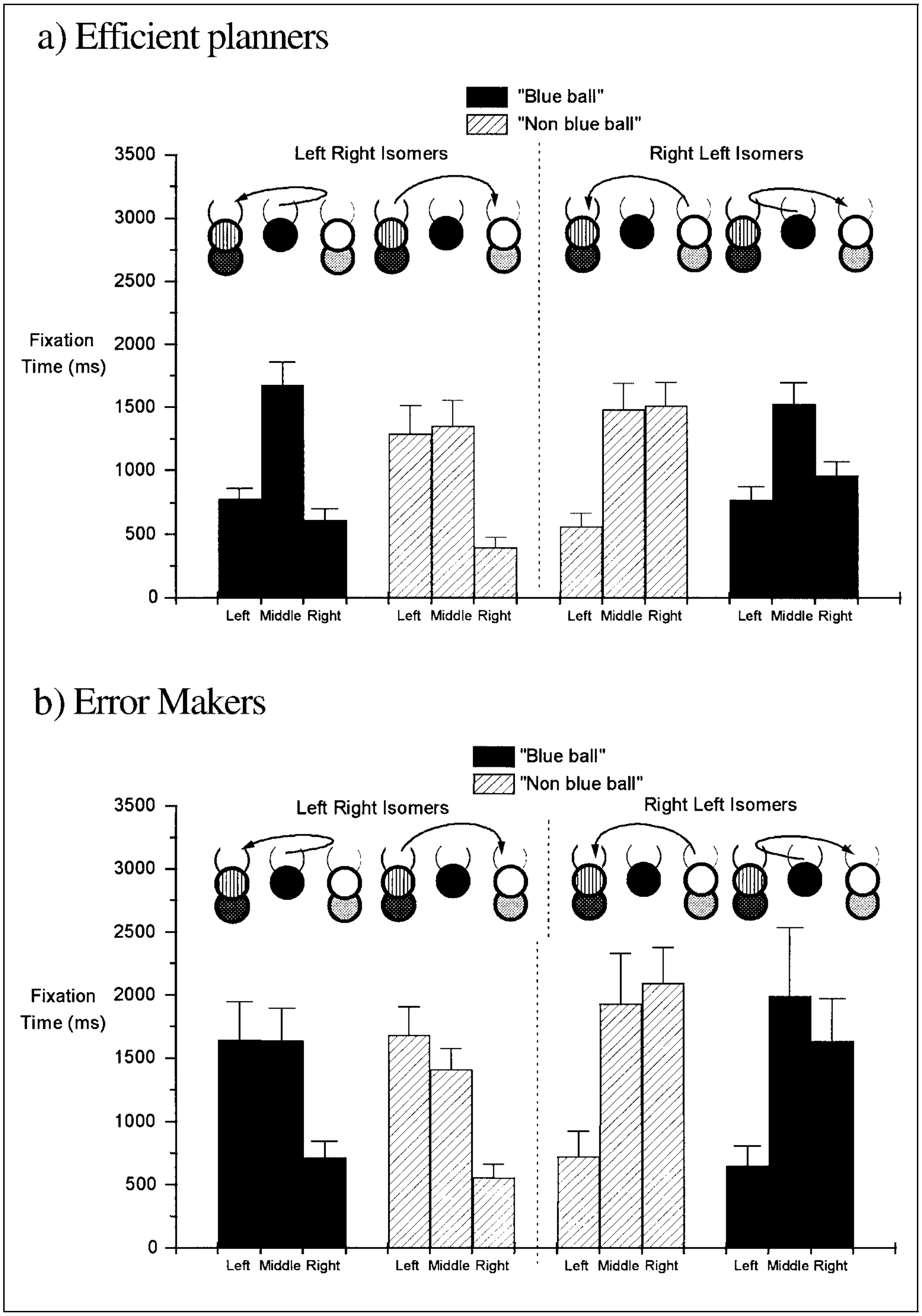

**Figure 5.** Total gaze dwell time per trial in the left, middle, and right locations of the workspace in Experiment 2 for (a) Efficient planners and (b) Error makers. Trials were divided according to the predominant movement direction (left-right and right-left isomers) and then subdivided into Blue ball and Nonblue ball problems. Efficient planners were found to direct gaze towards the blue ball location on Blue ball trials. In contrast, Error makers did not use this gaze strategy.

.0001) (Table 5). There was no effect of subject group (Error makers/Nonerror makers) on the size of this interaction ($F(2,16) = .28$).

### Temporal Order of Fixations

The probability of a given fixation landing on the Goalspace during each 500-msec time bin from picture onset was assessed. As in Experiment 1, the proportion of fixations landing in the Goalspace was highest at the beginning of the trial (three moves: $F(8,72) = 3.31$, $p < .005$; four moves: $F(11,99) = 5.58$, $p < .0001$). However, no clear increase in $P$(upper) was observed towards the end of the trial (Figure 2c). The probability of fixating the Goalspace was slightly higher overall relative to Experiment 1, reflecting more continuous sampling of the Goalspace.

### Fixation Dwell Times Within the Workspace

The distribution of total fixation times within the Workspace were analyzed using a three-way ANOVA of isomer direction (left-right/right-left) by problem type (Blue ball/Nonblue ball) by horizontal location (left, middle, right) for the two subject groups (Error makers and Efficient planners).

Efficient planners showed a significant interaction between dwell time within the three horizontal sectors and isomer type ($F(2,8) = 77.39$, $p < .0001$). Their direction of gaze was biased to the left ball locations for left-right isomers and to the right location for right-left isomers. Fixation times were also more biased towards the blue ball location on Blue ball problems ($F(2,8) = 19.14$, $p < .001$) (Figure 5a). In contrast, although Error makers showed a similar

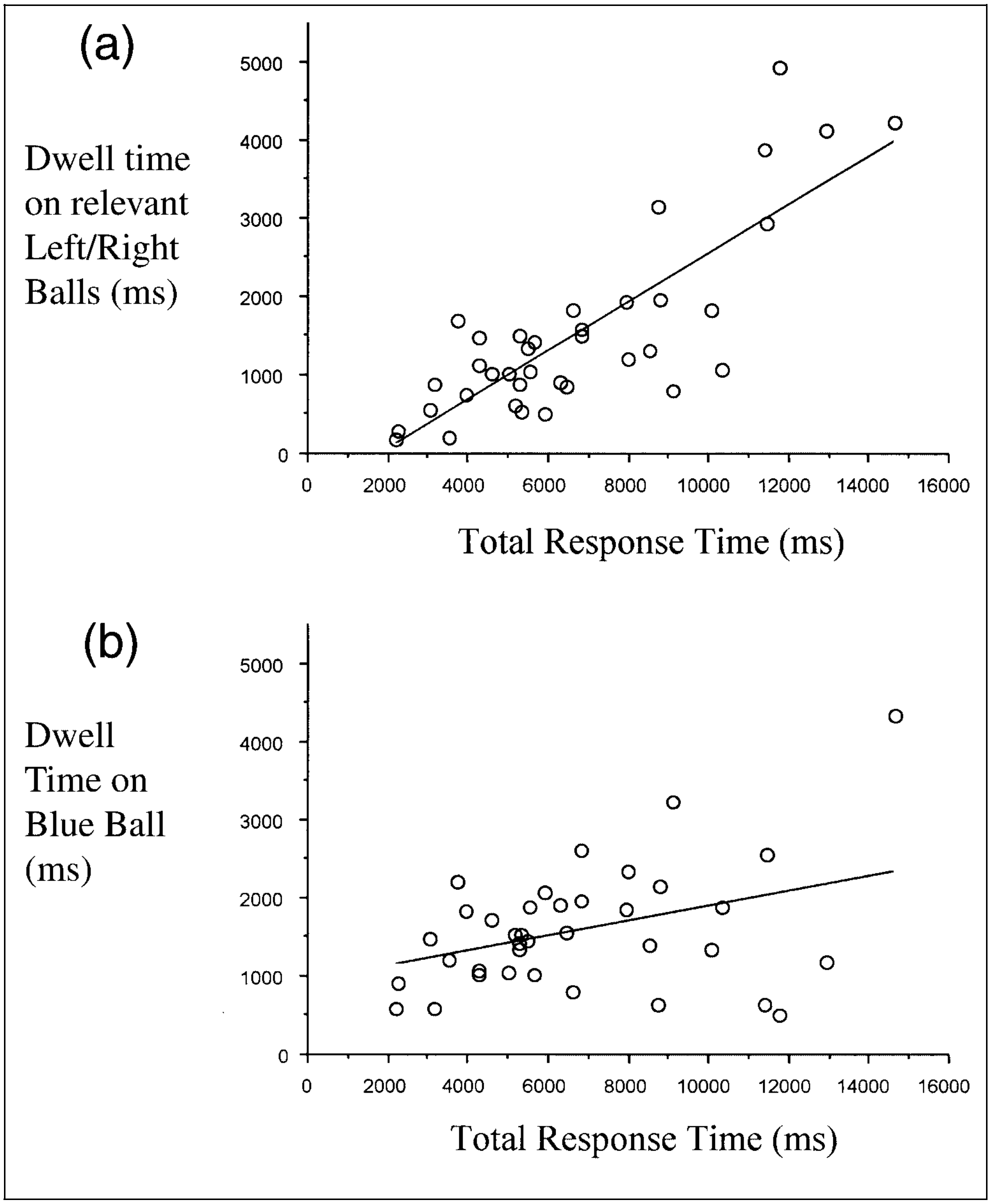

**Figure 6.** Scatter plots for Blue ball problems, with least-squares fit regression line, showing the relationship between total response time and fixation time on (a) the relevant left-right location balls and (b) the blue ball location. Longer response times on Blue ball trials were associated with increased fixation times on the left-right balls, but not on the blue ball location, which was the critical ball to move to solve these problems.

interaction between problem direction and location ($F(2,8) = 18.15$, $p < .01$), there was no significant difference in their gaze patterns between Blue ball and Nonblue ball problems ($F(2,8) = .34$). Error makers always biased their fixations towards the laterally located balls, even when the blue ball was the critical ball to move to solve that problem (Figure 5b).

### *Relationship Between Total Response Time and Gaze Direction*

The relationship between total response time and the distribution of fixations on the Workspace was investigated. We examined if the time spent fixating the blue ball and nonblue ball locations was correlated with response time for Blue ball problems (using four-move problems only and excluding trials where the subject failed to make a response within the allotted time period). This analysis revealed a strong correlation between response time and the time spent fixating the relevant left-right balls ($R^2 = .67$, $F(1,35) = 70.71$, $p < .0001$). However, a much weaker correlation was observed between the time spent fixating the middle (blue ball) location and total reaction time ($R^2 = .14$, $F(1,35) = 5.78$, $p < .05$) (Figure 6). This result suggests that on long reaction-time Blue ball trials, subjects were concentrating attention inappropriately on the laterally located balls.

### *Lateral Shifts in Fixation*

As in Experiment 1, lateral shifts in eye position between grid locations were more frequent within the Workspace ($F(1,9) = 6.20$, $p < .05$). We did not observe any evidence that subjects followed a stereotypical sequence of eye movements corresponding to a movement plan. However, dissociations were found in the distribution of horizontal eye movements with problem type and subject group. We calculated the proportion of eye movements which shifted fixation between the middle location and adjacent locations in either the congruent or incongruent direction relative to the goal position of the blue ball (Figure 3c). This analysis revealed that efficient planners made more gaze shifts to and from the incongruent location on Blue ball trials (Incongruent shifts: Blue ball, 17% of saccades; Nonblue ball, 12%), but more shifts to and from the congruent location on Nonblue ball trials (Congruent shifts: Blue ball, 25%; Nonblue ball, 40%) (direction × problem type interaction: $F(1,4) = 10.59$, $p < .03$). In contrast, Error makers did not show a significant dissociation between the two types of problem (Incongruent shifts: Blue ball 14%, Nonblue ball 17%; Congruent shifts: Blue ball 24%, Nonblue ball 32%) ($F(1,4) = .58$).

## GENERAL DISCUSSION

Experiment 1 established that there is a task-specific pattern of eye movements associated with TOL problems. Gaze was more likely to be directed towards the Goalspace early and late in the trial reflecting the assessment of relevant goals and the verification of problem solutions, respectively. In contrast, fixations were more likely to land on the Workspace during the middle of the trial and time spent fixating the Workspace increased with problem difficulty. The results of Experiment 2 went further by showing that the distribution of gaze on the Workspace was sensitive to the particular ball moves being assessed by the subject. Efficient planners biased their gaze towards the blue ball location while planning solutions to Blue ball problems and to the correct lateral ball locations on Nonblue ball problems, even though the arrangement of balls being viewed in the Workspace was the same on both types of trial. In contrast, subjects who made errors did not selectively direct gaze towards the critical location on Blue ball problems.

It has been suggested that when individuals are confronted by a particular visual stimulus, they produce a specific sequence of eye fixations or "scan-path," which recurs during recognition of the same stimulus. This account emphasizes the importance of salient stimulus features in guiding eye movements when viewing pictures (Stark & Ellis, 1981; Norton & Stark, 1971). However, in common with other authors, we have observed task-dependent variations in the spatial distribution and sequence of fixations on pictures independent of stimulus characteristics (Ballard, Hayhoe, & Pelz, 1995; Groner et al., 1984; Carpenter & Just, 1978; Yarbus, 1967; Buswell, 1935). It seems difficult to accommodate these results within the original version of "scan-path" theory. Instead, particular visuospatial tasks utilize dedicated eye-movement strategies (Land & Furneax, 1997).

The present results have implications for the high-level control of eye movements, but can an analysis of where people look contribute to our understanding of problem solving? Might it also be possible to identify features of gaze control within the TOL, which are not present in visuospatial tasks, which are less sensitive to anterior cortical pathology?

One issue for which this data has implications is the debate surrounding the constraints imposed on problem solving by limitations in working memory capacity. The time spent fixating the Workspace, but not the Goalspace, was found to increase with problem difficulty. One explanation for this finding is that task relevant information acquired during fixations on the Goalspace is held in memory during solution elaboration. If this were the case, then difficult TOL problems would require more goal information to be maintained in on-line memory. However, limitations in memory capacity would not necessarily disrupt task performance as long as frequent refixations of relevant display areas were possible. In this manner, gaze shifts may construct an "external" memory for the arrangement of balls in the Goalspace, negating the requirement for detailed internal representations of current goals in the task (Hayhoe et al., 1997; Ballard et al., 1995; Churchland et al., 1994; O'Regan, 1992; Ballard, 1991). The observation of more continuous visual sampling of the upper visual field in Experiment 2 relative to Experiment 1 (Figure 3c) suggests that this was exactly what happens when the complexity of TOL problems is increased.

It is also often assumed that advanced planning involves the construction of an internal program for the movement sequence, which is later recalled to control execution of the correct solution. This "plan-as-program" view (Clark, 1997) is difficult to reconcile with models which suggest that working memory constraints are noncritical in determining problem-solving efficiency (Ward & Allport, 1997; Newell & Simon, 1972). However, our data suggests that the classical view of planning may be incorrect. We found no evidence for stereotypical sequences of eye movements corresponding directly to the rehearsal of a fully formed action sequence. However, the finding that efficient planning involves selectively biasing gaze towards the problem-critical balls offers an alternative mechanism through which advanced planning might benefit problem solving. According to this proposal, directing gaze strategically during planning establishes a parsimonious motoric representation of the key features of the problem and facilitates refixation of the same areas of the display during problem solution. Foveating the critical components of the picture would have the effect of shaping the inflow of sensory information to the viewer, biasing other action systems towards selection of the correct behavior without the recall of a complete plan (Ballard et al., 1997). In other words, subjects might be solving problems with their eyes rather than "in their heads."

Of course, the role of mental representations in planning and problem solving should not be understated. One way in which the TOL differs from other visuospatial tasks such as block copying (Hayhoe et al., 1997) is the necessity to manipulate information internally using mental imagery during solution elaboration. Converging behavioral and anatomical evidence indicates that mechanisms of eye movement control and mental imagery are closely linked (Hodgson, Dittrich, Henderson, & Kennard, 1999; Brandt & Stark, 1997; Smyth, 1996; Awh, Smith, & Jonides, 1995; Baddeley, 1988; Selemon & Goldman-Rakic, 1985; Kosslyn, Ball, & Reiser, 1978; Kosslyn, 1988; Kosslyn, Behrman, & Jeannerod, 1995). Consequently, the requirement to maintain an external representation of the world via eye movements might be expected to interfere with the internal imagery processes necessary to elaborate problem solutions in the TOL task. Interestingly, manipulating visuospatial information internally while simultaneously making eye movements appears to be a common feature of a number of tasks sensitive to damage to the prefrontal cerebral cortex, including antisaccades (Walker, Husain, Hodgson, Harrison, & Kennard, 1997) and serial self-ordered search tasks (Morris, Downes, Sahakian, Polkey, & Robbins, 1988; Passingham, 1985). An alternative view is that gaze shifts may actually reduce the load on internal resources during imagery. By partially reproducing the changing pattern of afferent stimulation expected during physical manipulation of the problem, eye movements may allow imagery representations to be "scaffolded" upon sensory representations during cognitive planning (Clark, 1997).

Another way in which the TOL differs from less demanding tasks is the prerequisite to generate and flexibly shift between behavioral strategies or "sets" (Robbins, 1997). Other tasks, such as block copying, can be performed using a repetitive strategy to guide eye movements and behavior (Hayhoe et al., 1997; Land & Furneax, 1997). In contrast, the TOL requires the elaboration of new behavioral strategies for novel problems, as well as the ability to suppress inappropriate strategies, which were successfully applied to solve earlier problems. We found that for Error makers, the strategy required on the previous problem exerted a strong interference effect on the time taken to reach a decision on the subsequent trial (Figure 4). In contrast, efficient planners were able to switch between Blue ball and Nonblue ball problems without incurring a cost in solution time. Analysis of gaze direction on Blue ball trials, showed that errors and long reaction times were accompanied by a failure to selectively direct gaze towards the critical blue ball location (Figures 5 and 6). It was as if on these trials subjects were blind to the behavioral significance of the blue ball, or were unable to inhibit attention from being drawn to other behaviorally salient locations. This blue ball "neglect" is reminiscent of the "goal neglect" proposed to underlie dysexecutive syndrome in patients with frontal cortex pathology, as well as poor task performance in some normal individuals (Duncan, Johnson, Swales, & Freer, 1997; Duncan, Emslie, Williams, Johnson, & Freer, 1996).

We have shown that efficient planning of solutions to TOL problems utilizes specialized gaze strategies. Systematic biases in gaze direction during the course of a trial suggest that problem solving proceeds in three discrete phases corresponding to problem assessment, solution elaboration and solution verification. We have proposed several ways in which strategic gaze shifts act to reduce the working memory load imposed by the task. Our observations also highlight a unique feature of gaze control in the task, namely the requirement to flexibly shift between different control sets or strategies. Further research should examine this mode of gaze control in more detail, as well as assessing the strategic control of eye movements in neurologically impaired populations.

## METHODS

### Experiment 1

### *Task and Stimuli*

Each subject viewed 24 pictures showing two arrays of two colored balls positioned in pockets. Pictures subtended 17° by 20° of visual arc and were displayed using a Macintosh 2ci computer with a 17-in. color computer monitor. Balls could be colored either red and green, red and blue, or pink and yellow. The two arrays of balls were located in the upper and lower visual fields. The leftmost location in each array had space for a maximum of three balls, the middle location two balls, and the right location had space for only one ball (Figure 1). At the start of each trial, a central fixation cross was displayed for 500 msec. For calibration purposes, the subject was asked to look at this cross, which was extinguished simultaneously with the presentation of problem pictures. Subjects were instructed to plan, but not execute the problem solutions (i.e., the "one-touch" TOL task, Owen et al., 1995). Once the subject thought that they had worked out the correct solution to each problem they pressed the mouse key and immediately gave a verbal response to indicate whether the problem solution involved a minimum of either "one," "two," or "three" moves.

### *Subjects*

Eight subjects participated in Experiment 1. Five were male, three were female, aged between 19 and 31 years. All were unfamiliar with the TOL task and had normal or corrected-to-normal vision. Subjects were divided into two groups according to the instructions they were given as to how to solve the task. The Upstairs group were told to plan the sequence of moves required to rearrange the balls in the top part of the display to match the bottom half (i.e., Workspace in the upper visual field, Goalspace in the lower visual field). In contrast, the Downstairs group were required to rearrange the balls in the bottom half of the display to match the top half (i.e., Workspace in the lower visual field, Goalspace in the upper visual field).

### *Procedure*

Prior to the start of the experiment, subjects were given instructions concerning the task indicating how the balls could be moved from one pocket to another and could not be placed directly underneath another ball without moving obstructing balls to an alternative location. It was also emphasized that they should plan out the entire sequence of moves required to solve each problem. The Downstairs group, who were instructed to rearrange the balls in the lower field to match the arrangement in the upper field, were additionally given practice on a touch-screen version of the TOL task (Stockings of Cambridge[1]), without eye movements being recorded. The Upstairs group were tested at a later date and, in order to avoid an interference effect, were not exposed to the touch-screen version of the task, in which the balls are always moved within the lower half of the display. During the eye-tracking session, both groups of subjects were given a block of practice on the task prior to the experimental block and made less than 1% errors in the experimental block.

### *Eye Tracking and Analysis*

Eye movements were recorded using the EyeLink system (Sensorimotoric Systems), a video-based, pupil tracker, with head movement compensation system. Subjects were seated at a comfortable viewing distance in front of the display monitor approximately 60 cm from the computer screen. They were instructed to keep head movements to a minimum and no active restraint of head movements was required to obtain accurate gaze position recordings. Eye movements were analyzed offline using custom software written in C on the Macintosh. Fixations were categorized according to where they landed on a 3 × 2 grid (upper, lower, left, middle, right), which divided the pictures into six sectors of equal area (Figure 1). Fixation duration, *x-y* position, and grid location were outputted to text files for each subject and each TOL problem. Eye-movement traces were visualized by the experimenter and played back at slowed speed superimposed over the picture that was being viewed on that trial. Individual saccades were then identified using a semiautomated procedure, as periods in the eye position signal where the instantaneous, absolute velocity rose above 30°/sec for more than two data samples. Fixations were identified as pauses between saccades longer than 50 msec in duration. The experimenter could reject any fixations, which were contaminated by eye blink or eyelid clipping artifacts.

### Experiment 2

### *Task and Stimuli*

All subjects were instructed to solve the problems in the Downstairs manner, rearranging the lower configuration of balls to match the upper arrangement of balls. Two arrangements of five colored balls (red, blue, pink, yellow, and green) were presented in the upper and lower parts of the display. The leftmost location in each array could contain a maximum of three balls, the middle location contained a maximum of two balls, and the right location had space for three balls. The arrangement of balls in the Workspace was always constant, and only the balls in the Goalspace varied from trial to trial. For each trial, two of the balls were “dummy” balls, which were not relevant to the solution of that problem (Figure 3). This allowed the stimuli to be constructed so that each problem had an accompanying mirror-image problem or isomer. For left-right and right-left isomers, the correct ball moves were predominantly from left to right and right to left, respectively. Problems were always either two, three, or four moves in length. The set of problems selected was based on those used in the Stockings of Cambridge test. The task was otherwise identical to Experiment 1. When interviewed following the test, nine out of the 10 subjects did not notice that the arrangement of balls in the Workspace was constant throughout the experimental block.

Left-right and right-left isomers were further subdivided into Blue ball and Nonblue ball problems. Blue ball problems were defined as those which required a “shunting” maneuver in which the centrally located Blue ball had to be moved to a temporary subgoal location prior to being moved to its final destination (Figure 3). Blue ball problems could not be solved successfully unless the subject realized the importance of this maneuver. In contrast, for Nonblue ball problems, it was the laterally located balls (in either the left or right locations), which were the critical balls to move in order to solve the problem.

### *Subjects*

Ten subjects took part in Experiment 2, aged between 19 and 31 years. Three were male and six were female. All but one of the subjects had not taken part in Experiment 1 or been exposed to the TOL task on a previous occasion. All subjects had normal or corrected-to-normal vision. In order to introduce a range of performance and to provide a control for the differing practice conditions in Experiment 1, four out of the 10 subjects were given no practice on the Stockings of Cambridge, touch-screen task prior to the test. The other six subjects received practice on the touch-screen version of the test prior to the eye-movement recording session. All subjects were presented with a block of practice trials on the one-touch version of the task before performing the experimental block. Different five ball problems were presented in the practice and experimental blocks.

## Acknowledgments

This study was supported by The Wellcome Trust Program Grant Number: 037222/2/98/A. We would like to thank an anonymous reviewer for providing helpful comments on the manuscript and offering some useful suggestions for additional data analysis.

## Note

1. Part of the CANTAB test battery. CeNs Cambridge Cognition.

## REFERENCES

Allport, D. A., Styles, E. A., & Hseih, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In C. Umilta & M. Moscovitch (Eds.), *Attention and performance: XV. Conscious and non-conscious information processing*. Cambridge: MIT Press.

Andersen, J. R. (1993). *Rules of the mind.* Hillsdale, NJ: Erlbaum.

Awh, E., Smith, E., & Jonides, J. (1995). Human rehearsal processes and the frontal lobes: PET evidence. *Annals of the New York Academy of Sciences, 769,* 97–117.

Baddeley, A. (1988). *Working memory.* Oxford: Oxford Science Publications.

Baddeley, A., & Della Sala, S. (1996). Working memory and executive control. *Philosophical Transactions of the Royal Society of London, Series B, 351,* 1397–1404.

Baker, S. C., Rogers, R. D., Owen, A. M., Frith, C. D., Dolan, R. J., Frackowiak, R. S. J., & Robbins, T. W. (1996). Neural systems engaged in planning: A PET study of the tower of London task. *Neuropsychologia, 34,* 515–526.

Ballard, D. H. (1991). Animate vision. *Artificial Intelligence, 48,* 57–86.

Ballard, D. H., Hayhoe, M. M., & Pelz, J. B. (1995). Memory representations in natural tasks. *Journal of Cognitive Neuroscience, 7,* 66–80.

Ballard, D. H., Hayhoe, M. M., Pook, P. K., & Rao, R. P. N. (1997). Deictic codes for the embodiment of cognition. *Behavioural and Brain Sciences, 20,* 723–767.

Boden, M. A. (1997). Being there: Putting brain, body, and world together again—Clark, A. (Book Review). *Nature, 386,* 237–238.

Brandt, S. A., & Stark, L. W. (1997). Spontaneous eye movements during visual imagery reflect the content of the visual scene. *Journal of Cognitive Neuroscience, 9,* 27–38.

Brooks, R. (1991). Intelligence without reason. In *Proceedings of the 12th international joint conference on artificial intelligence*. San Francisco: Morgan Kauffman.

Buswell, G. T. (1935). *How people look at pictures.* Chicago: University of Chicago Press.

Carpenter, P. A., & Just, M. A. (1978). Eye fixations during mental rotation. In J. W. Senders, D. F. Fisher, & R. A. Monty (Eds.), *Eye movements and the higher psychological functions* (pp. 115–133). Hilsdale, NJ: Erlbaum.

Churchland, P. S., Ramachandran, V. S., & Sejnowski, T. J. (1994). A critique of pure vision. In C. Koch & J. Davis (Eds.), *Large-scale neuronal theories of the brain* (pp. 23–60). Cambridge: MIT Press.

Clark, A. (1997). *Being there: Putting brain, body and world together again.* Cambridge: MIT Press.

Dagher, A., Owen, A. M., Boecker, H., & Brooks, D. J. (1999). Mapping the network for planning: A correlational PET activation study with the tower of London task. *Brain, 122,* 1973–1988.

Dehaene, S., & Changeux, J. (1997). A hierarchical neuronal network for planning behaviour. *Proceedings of the National Academy of Science, U.S.A. 94,* 13293–13298.

Duncan, J., Emslie, H., Williams, P., Johnson, R., & Freer, C. (1996). Intelligence and the frontal lobe: The organization of goal-directed behavior. *Cognitive Psychology, 30,* 257–303.

Duncan, J., Johnson, R., Swales, M., & Freer, C. (1997). Frontal lobe deficits after head injury: Unity and diversity of function. *Cognitive Neuropsychology, 14,* 713–741.

Funahashi, S., Bruce, C. J., & Goldman-Rakic, P. S. (1989). Mnemonic coding of visual space in the monkey dorsolateral prefrontal cortex. *Journal of Neurophysiology, 61,* 331–349.

Groner, R., Walder, F., & Groner, M. (1984). Looking at faces: Local and global aspects of scanpaths. In A. G. Gale, F. Johnson (Eds.), *Theoretical and applied aspects of eye movement research* (pp. 523–533). Amsterdam: North-Holland.

Hayhoe, M. H., Bensinger, D. G., & Ballard, D. H. (1997). Task constraints in visual working memory. *Vision Research, 38,* 125–137.

Hodgson, T. L., Dittrich, W., Henderson, L., & Kennard, C. (1999). Eye movements and spatial working memory in Parkinson's disease. *Neuropsychologia, 37,* 927–938.

Kirsh, D. (1995). The intelligent use of space. *Artificial Intelligence, 72,* 1–52.

Kirsh, D., & Maglio, P. (1994). On distinguishing between epistemic and pragmatic action. *Cognitive Science, 18,* 513–549.

Kosslyn, S. M. (1988). Aspects of a cognitive neuroscience of mental imagery. *Science, 240,* 1621–1626.

Kosslyn, S. M., Ball, T. M., & Reiser, B. J. (1978). Visual images preserve metric spatial information: Evidence from studies of image scanning. *Journal of Experimental Psychology Human Perception and Performance, 4,* 47–60.

Kosslyn, S. M., Behrman, M., & Jeannerod, M. (1995). The cognitive neuroscience of mental imagery. *Neuropsychologia, 33,* 1335–1344.

Land, M. F., & Furneax, S. (1997). The knowledge base of the oculomotor system. *Philosophical Transactions of the Royal Society of London, Series B, 352,* 1231–1239.

Morris, R. G., Ahmed, S., Syed, G. M., & Toone, B. K. (1993). Neural correlates of planning ability: Frontal lobe activation in the Tower of London test. *Neuropsychologia, 31,* 1367–1378.

Morris, R. G., Downes, J. J., Sahakian, B. J., Polkey, C. E., & Robbins, T. W. (1988). Planning and spatial working memory in Parkinson's disease. *Journal of Neurology, Neurosurgery, and Psychiatry, 51,* 757–766.

Newell, A. (1990). *Unified theories of cognition.* Cambridge: Harvard University Press.

Newell, A., & Simon, H. A. (1972). *Human problem solving.* Englewood Cliffs, NJ: Prentice-Hall.

Norton, D., & Stark, L. W. (1971). Scanpaths in saccadic eye movements while viewing and recognizing patterns. *Vision Research, 11,* 929–942.

O'Regan, J. (1992). Solving the "real" mysteries of visual perception: The world as an outside memory. *Canadian Journal of Psychology, 46,* 461–488.

Owen, A. M., Downes, J. D., Sahakian, B. J., Polkey, C. E., & Robbins, T. W. (1990). Planning and spatial working memory deficits following frontal lobe lesions in man. *Neuropsychologia, 28,* 1021–1034.

Owen, A. M., Doyon, J., Petrides, M., & Evans, A. C. (1996). Planning and spatial working memory examined with positron emission tomography (PET). *European Journal of Neuroscience, 8,* 353–364.

Owen, A. M., Sahakian, B. J., Hodges, J. R., Summers, B. A., Polkey, C. E., & Robbins, T. W. (1995). Dopamine-dependent fronto-striatal planning deficits in early Parkinson's disease. *Neuropsychology, 9,* 126–140.

Passingham, R. E. (1985). Memory of monkeys (*Macaca mulatta*) with lesions to prefrontal cortex. *Behavioural Neuroscience, 99,* 3–21.

Previc, F. H. (1990). Functional specialization in the lower and upper visual field in humans: Its ecological origins and neurophysiological implications. *Behavioral and Brain Sciences, 13,* 519–575.

Robbins, T. W. (1996). Dissociating executive functions of prefrontal cortex. *The Philosophical Transactions of the Royal Society of London Series B-Biological Sciences 351,* 1463–1470.

Selemon, L. D., & Goldman-Rakic, P. S. (1985). Longitudinal topography and interdigitation of cortico-striatal projections in the rhesus monkey. *Journal of Neuroscience, 5,* 776–794.

Shallice, T. (1982). Specific impairments of planning. *Philosophical Transactions of the Royal Society of London, Series B, 298,* 199–209.

Smyth, M. M. (1996). Interference with rehearsal in spatial working memory in the absence of eye movements. *Quarterly Journal of Experimental Psychology, 49A,* 940–949.

Stark, L. W., & Ellis, S. R. (1981). Scanpath revisted: Cognitive models direct active looking. In D. F. Fisher, R. A. Monty, & J. W. Senders (Eds.), *Eye movements: Cognition and visual perception* (pp. 193–226). Hillsdale, NJ: Erlbaum.

Walker, R., Husain, M., Hodgson, T. L., Harrison, J., & Kennard, C. (1997). Saccadic eye movement and working memory deficits following damage to human prefrontal cortex. *Neuropsychologia, 36,* 1141–1159.

Ward, G., & Allport, A. (1997). Planning and problem-solving using the five-disc Tower of London task. *The Quarterly Journal of Experimental Psychology, 50A,* 49–78.

Yarbus, A. (1967). *Movements of the eyes.* New York: Plenum.