



PEARL

**Attention deployment during memorizing and executing complex instructions.**

Apel, Jens K.; Revie, Gavin F.; Cangelosi, Angelo; Ellis, Rob; Goslin, Jeremy; Fischer, Martin H.

**Published in:**  
Exp Brain Res

**DOI:**  
[10.1007/s00221-011-2827-4](https://doi.org/10.1007/s00221-011-2827-4)

**Publication date:**  
2011

**Link:**  
[Link to publication in PEARL](#)

**Citation for published version (APA):**

Apel, J. K., Revie, G. F., Cangelosi, A., Ellis, R., Goslin, J., & Fischer, M. H. (2011). Attention deployment during memorizing and executing complex instructions. *Exp Brain Res*, *214*(2), 249-259. <https://doi.org/10.1007/s00221-011-2827-4>

All content in PEARL is protected by copyright law. Author manuscripts are made available in accordance with publisher policies. Wherever possible please cite the published version using the details provided on the item record or document. In the absence of an open licence (e.g. Creative Commons), permissions for further reuse of content should be sought from the publisher or author.

# Attention deployment during memorizing and executing complex instructions

Jens K. Apel<sup>1</sup>, Gavin Revie<sup>1</sup>, Angelo Cangelosi<sup>2</sup>, Rob Ellis<sup>2</sup>, Jeremy  
Goslin<sup>2</sup>, Martin H. Fischer<sup>1</sup>

1 - School of Psychology, University of Dundee

2 - School of Psychology, University of Plymouth

Please send correspondence to: Martin Fischer

School of Psychology, University of Dundee

Dundee DD1 4HN, Scotland

Tel.: \*44-1382-38-46-12, FAX: \*44-1382-22-99-93

email: [m.h.fischer@dundee.ac.uk](mailto:m.h.fischer@dundee.ac.uk)

ABSTRACT

We investigated the mental rehearsal of complex action instructions by recording eye movements of eleven healthy adults as they looked at several objects on a monitor. Participants were presented with 3-8 consecutive instructions of the form “Move the [object] to [location]”. Instructions were only to be executed after a go signal, by manipulating the objects with a mouse. Participants re-inspected previously mentioned objects while listening to new instructions prior to the go signal. This rehearsal behavior broke down after 4 instructions, coincident with participants’ instruction span, as determined from subsequent execution accuracy. Our results indicate that eye movements during instruction presentation can predict their successful execution.

Keywords: assembly task, eye movements, overt attention, rehearsal, sequential instruction, working memory.

## Introduction

Our ability to follow instructions is limited by our capacity to remember them. For example, instead of assembling pieces of furniture in one go after reading the instruction once, we must frequently refer back to the manual and remind ourselves of the next step in the complex assembly sequence. The manual serves as an external support mechanism that compensates for our limited memory span. If, however, complex instructions are given verbally, there is no opportunity to refer back to them later and we must make an effort to rehearse earlier instruction components even while new ones are given (xxx). How do we do this, and do detailed aspects of our rehearsal behavior predict subsequent performance? Understanding our memory strategies and limitations has important practical implications, and the present paper uses eye movements to study the relation between rehearsal and performance.

Processing information related to an object in the visual environment leads to eye movements towards this object. Cooper (1974) discovered this linguistic cueing effect on eye movements by showing participants a set of objects while listening to a short story. Participants were more likely to look at a lion while hearing the word 'lion', or at a zebra when hearing the word 'Africa'. This pervasive tendency to fixate objects when hearing a related referent in speech has become the basis of the much studied "visual world paradigm" which uses eye movements to assess the speed of speech comprehension. For example, Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy (1995) showed displays with an apple on a towel, another apple, a towel, and a box, and instructed participants to 'put the apple on the towel in the box'. The authors found that eye fixations were equally likely to be directed at both apples until participants heard the words "in the box". Altmann and colleagues (Altmann & Kamide, 1999; Kamide, Altmann, & Haywood, 2003) documented predictive eye fixations towards speech-related objects prior to their naming when a verb refers to a single object (e.g., hearing

Attention deployment during memorizing and executing complex instructions (the verb “eat” induces fixations onto edible objects), thus revealing incremental speech comprehension in real time through overt attention deployment.

Previous “visual world” studies have investigated eye movements in response to single instructions. The current study extends this approach towards more complex instructions. Similar to the work by Chambers and colleagues (Chambers, Tanenhaus, & Magnuson, 2004; Chambers & Juan, 2008), we investigated spatio-temporal aspects of eye behavior as participants listened to how they had to move objects within a grid on the screen. However, more than one object had to be moved per trial, and several instruction components had to be kept in memory before any action could be executed. Do participants attend to each object only once, or do they re-fixate previously mentioned objects? We reasoned that systematic re-fixations might reflect mental rehearsal which, in turn, might predict subsequent execution performance. Moreover, the time spent looking at an object might predict the probability of its correct placement.

### **Method**

Participants were presented with multi-component instructions like ‘Move the mug to square 5, then move the hammer to square 3, then move the pan to square 8’ (3 instruction components; see Figure 1). After hearing all instruction components, they had to execute these in the correct order with a mouse on a computer screen. We were interested in three main questions: (1) How well are complex instructions executed? (2) How do participants attend to objects and locations during encoding of the complex instructions (3) Does attention deployment during encoding predict performance during execution?

#### **Participants**

We tested a convenience sample of eleven students at the University of Dundee (9 female, one left-handed, age range 19-46 years, all fluent in English).

## Attention deployment during memorizing and executing complex instructions

### Apparatus

Stimuli were presented on a 17-inch CRT monitor running at 100 Hz and with a resolution of 1024 x 768 pixels. A desktop-mounted EyeLink® 1000 eye tracker recorded the position of participants' dominant eye at 1000 Hz and with 0.25° spatial accuracy. A chin rest was located 62 cm from the screen to stabilize participants' head. The experiment was created using Experiment Builder and data were analyzed using Data Viewer (both SR Research Ltd., Canada) and Matlab®.

### Stimuli

Twenty-four items were used, each consisting of a 3x3 numbered black grid on a white background containing four colored objects (see Figure 1). The array was 532 pixels across (16.5° of visual angle). Each square within the grid had a length of 169 pixels (5.5 ° of visual angle). The same four objects were presented in each item; a hammer, an umbrella, a pan, and a mug. Two start configurations appeared equally often: hammer in square 8 or 1, umbrella in square 6 or 9, pan in square 2 or 3, and mug in square 5 or 4. The handle of each object appeared equally often on the left and right side and there were equally many left- and right-side handles in each item. All objects fitted snugly inside the squares.

Insert Figure 1 here

Twenty-four auditory instructions were recorded from a Scottish male native speaker and assembled into lists of 3-8 instruction components (see Appendix). An example instruction with three components is "*Move the umbrella to square nine, then move the mug to square five, then move the pan to square three.*" Each component size was presented four times in one of four randomized instruction lists. Very long (seven or eight component) instruction lists did not appear in the first three trials.

### Task and Procedure

## Attention deployment during memorizing and executing complex instructions

Participants gave written informed consent and completed the Miles Test (Miles, 1930) and the Porta Test (Crovitz & Zener, 1962) to determine ocular dominance. Using a “two out of three” criterion, only the dominant eye was tracked (nine right, one left eye). The chinrest height was adjusted to be comfortable and the eye tracker was calibrated using a 9-dot matrix, followed by validation. Once calibration was accurate to less than 0.5 ° of visual angle the testing began with a fixation dot located at the centre of the screen. Once participants fixated this dot the grid with objects would appear. The auditory assembly instructions began 500 ms after the grid onset and ended with a beep (40 ms, 2000 Hz) that was the go signal to execute all instruction components in order by dragging and dropping the objects with the mouse. Participants pressed a button on a button-box to indicate when they had finished responding for that trial. The first two trials were practice trials using instruction lists of lengths one and two, respectively. When it was clear that participants understood the task they proceeded with the experimental items, else they received more practice. Participants were offered a short break in the middle of the experiment and were recalibrated before resuming testing. At completion of testing participants were debriefed and paid. The experiment took approximately 30 min.

## Results

This section is structured into three main parts: (1) Execution, (2) Encoding, and (3) Influence of encoding on execution. The execution section reports how fast and accurately the instructions were carried out. The encoding section describes how participants fixated the objects within the grid while they listened to the instructions, both in space and over time. Section three documents that overt attention allocation during encoding predicts how participants subsequently perform in our task. Data were analyzed with analysis of variance (ANOVA) and Bonferroni-corrected post-hoc-tests, and degrees of freedom were Greenhouse-Geisser corrected where appropriate.

### **Execution**

The execution phase started with the acoustic go signal and ended with pressing a button on the button box. This interval increased systematically with complexity: Participants required 6.7-18.9 s for performing between three and eight component instructions,  $F(1.99, 19.87) = 26.1, p < .001$ . Despite this, there was a tendency to attempt no more than about five components, as indicated by the average number of objects picked up and moved per trial (left panel of Figure 2). An ANOVA confirmed a significant main effect of instruction length on number of actions,  $F(5, 50) = 21.7, p < .01$ , with post-hoc analyses showing that the numbers of actions significantly differed from each other only for 3-5 component instructions ( $ps < .02$ , all other  $p > .10$ ).

Insert Figure 2 here

An action was counted as correct when an instructed object was moved to the instructed location. The order of actions was not taken into account because participants always tried to respond in the instructed order. Consistent with the idea of a limited *instruction span*, the main effect of instruction length on proportion of correct actions was highly significant,  $F(5, 50) = 45.2, p < .01$ , as were most pair-wise comparisons (see center panel of Figure 2). If instead the absolute number of correct actions is plotted against instruction length, it becomes clear that participants only correctly executed approximately three instructions (right panel of Figure 2). The main effect of instruction length was significant ( $F(5, 50) = 2.7, p < .05$ ).

### **Encoding**

In order to visualize eye movements towards objects and the goal squares, the locations of participants' eye fixations relative to the time of speech onsets of the critical object names in all instruction components were coded as follows: For each millisecond (ms) between the onset and offset of each instruction, we measured whether



Attention deployment during memorizing and executing complex instructions

participants fixated an object or goal square. The object in the grid matching the object mentioned in an instruction component was the *critical object*. The square to which the object was instructed to be moved was the *goal square*. Figure 3 shows the results of this analysis. During the first, second and, third instruction components, participants fixated the critical objects and goal squares more often than the remaining objects in the grid. Fixations towards a critical object and goal square increased after speech onset of the associated object and square number conveyed by the instructions.

Insert Figure 3 here

Fixations towards objects and goal squares mentioned in instruction components five to eight are less clearly structured. Presumably the main reason for this is that, from the fifth instruction component onwards, objects were mentioned for a second time. Thus, it was impossible to distinguish fixations towards an object associated with its first or second mentioning. Thus, while it was appropriate to present up to eight instruction components in order to determine the instruction span, we limited our eye movement analysis to the first four components, consistent with the limited instruction span (see above).

Three eye movement measures were analyzed. A *fixation* was defined as belonging to a specific instruction component when the fixation onset appeared after the speech onset of the noun denoting the critical object of this instruction component and before the speech onset of the noun of the following instruction component. *Gaze duration* was the sum of fixation durations of all fixations towards on object or goal square executed within a given instruction component. A *transition saccade* occurred when a fixation on a critical object was directly followed by a fixation on the goal square. Tables 1-3 illustrate the number of fixations, gaze durations of fixations towards the critical objects and goal squares, and transition saccades during the first four instruction components. We conducted several analyses on these three eye movement

Attention deployment during memorizing and executing complex instructions measures during each instruction component, separately for critical objects and goal squares. Our results replicate the general tendency to attend to named objects and locations (see Introduction) and reveal, for the first time, a load effect on linguistic cueing of eye movements.

Critical objects: During the first three instruction components, participants fixated the critical object more often than the remaining objects (1<sup>st</sup> instruction component:  $F(1.30, 12.97) = 116.00, p < .001$ ; 2<sup>nd</sup> instruction component:  $F(1.67, 16.73) = 53.76, p < .001$ ; 3<sup>rd</sup> instruction component:  $F(2.34, 23.41) = 22.87, p < .001$ ). Post-hoc comparisons between the critical object and all remaining objects were all significant for the first and second instruction component ( $ps < .01$ ). Number of fixations on the critical object in the third instruction component differed significantly from those on the object mentioned in the fourth instruction component ( $p < .01$ ) and marginally differed from the object mentioned in the first instruction component ( $p = .07$ ). Number of fixations did no longer differ between critical and remaining objects during the fourth instruction component ( $F(3, 30) = .64, p > .10$ ).

Results of the gaze duration analysis were very similar. While listening to the first three instruction components, participants fixated the critical objects longer than the remaining objects (1<sup>st</sup> instruction component:  $F(1.08, 10.77) = 69.98, p < .001$ ; 2<sup>nd</sup> instruction component:  $F(1.64, 16.40) = 29.88, p < .001$ ; 3<sup>rd</sup> instruction component:  $F(2.15, 21.52) = 22.13, p < .001$ ). Post-hoc comparisons were significant between the critical object and all remaining objects in the first and second instruction components ( $p$  values  $< .05$ ). During the third instruction component the critical object only differed from the object mentioned during the fourth instruction component ( $p < .001$ ). Gaze duration of eye movements towards critical vs. remaining objects during the fourth instruction did not differ ( $F(3, 30) = 1.04, p > .10$ ).

Insert Tables 1 and 2 here

## Attention deployment during memorizing and executing complex instructions

Goal Squares: Very similar results were also found for numbers of fixations and gaze durations towards the goal squares (see Table 2). Participants fixated the goal squares in the first three instructions more often than most of the remaining squares. Significant main effects of numbers of fixations (1<sup>st</sup> instruction component:  $F(1.28, 12.75) = 91.3, p < .001$ ; 2<sup>nd</sup> instruction component:  $F(1.74, 17.40) = 38.4, p < .001$ ; 3<sup>rd</sup> instruction component:  $F(2.47, 24.73) = 18.6, p < .01$ ) and gaze duration (1st instruction component:  $F(1.10, 10.96) = 41.1, p < .001$ ; 2<sup>nd</sup> instruction component:  $F(1.95, 19.50) = 27.6, p < .001$ ; 3<sup>rd</sup> instruction component:  $F(3, 30) = 13.4, p < .01$ ) were found. Post-hoc comparisons for both numbers of fixations and gaze durations showed that during the first instruction component, the goal squares differed significantly from the remaining squares (all  $p$ s  $< .001$ ). During the second instruction component, the goal square differed from the goal squares mentioned in the third and fourth instructions (all  $p$ s  $< .001$ ). During the third instruction component, the goal square differed from the goal squares mentioned in the fourth instruction component ( $p < .01$ ). Number of fixations and gaze duration did not differ between squares mentioned during the fourth instruction component (number of fixations:  $F(3, 30) = 1.9, p > .10$ ; gaze duration:  $F(3, 30) = 2.3, p > .10$ ).

Consider now transition saccades (Table 3). During an instruction component, participants executed more transition saccades between the critical object and goal square for this instruction component compared to saccades between objects and goal squares that were not yet mentioned. A significant main effect for numbers of transition saccades was found for the first instruction component ( $F(1.1, 10.9) = 75.4, p < .001$ ). Post-hoc comparisons showed that transition saccades were executed significantly more often than saccades corresponding to other instruction components (all  $p$  values  $< .001$ ). A significant main effect was also found for the second instruction component ( $F(1.6, 16.3) = 44.7, p < .001$ ). Post-hoc tests showed that the numbers of transition saccades

Attention deployment during memorizing and executing complex instructions during the second instruction component differed from all other saccades (all  $p$ s  $< .01$ ). The same observation held during the third instruction component ( $F(3, 20) = 8.4$ ,  $p < .01$ ), and post-hoc comparisons also showed more transition saccades corresponding to the third vs. fourth instruction component ( $p < .01$ ). Number of transition saccades did no longer differ from number of other saccades during the fourth instruction component ( $F(3, 30) = 1.6$ ,  $p > .10$ ).

Load effect on linguistic cueing: Attention towards critical objects and goal squares clearly declined in a gradual fashion across the first four instruction components. For example, fewer fixations were directed to the object mentioned in the second compared to the first component. This reduced linguistic cueing effect for more complex instructions was present in all eye movement measures: (1) number of fixations and gaze durations on critical objects:  $F(3, 30) = 71.4$ ,  $p < .001$  and  $51.1$ ,  $p < .001$ , respectively; post-hoc comparisons between the critical objects were all significant,  $p$  values  $< .03$  (2) number of fixations and gaze durations on goal squares:  $F(3, 30) = 32.8$ ,  $p < .001$  and  $29.9$ ,  $p < .001$ , respectively; post-hoc comparisons of numbers of fixations and gaze duration between the goal squares were significant (all  $p$  values  $< .05$ ) except the comparison between goal squares of third and fourth instruction components; (3) transition saccades:  $F(3, 30) = 33.5$ ,  $p < .001$ ; all post-hoc comparisons except the comparison between transition saccades of the third and fourth instruction component differed reliably (all  $p$  values  $< .02$ ).

This decline in fixations towards critical objects and goal squares, and also in transition saccades, was due to re-fixating previously mentioned objects while listening to new instruction components. We quantified this effect by comparing attention deployment between previously mentioned and previously unmentioned objects or locations. For instance, while hearing the second instruction component, the number of fixations to the object mentioned in the first component, as well as gaze durations on

Attention deployment during memorizing and executing complex instructions

this first critical object, were significantly larger compared to the third and fourth instruction component (all  $p$  values  $< .01$ ). Similar results held for number of fixations and gaze durations on goal squares (all  $p$  values  $< .01$ ). During the third component instruction, there was a lesser peak of number of fixations and gaze durations on the third critical object and goal square, and a concomitant return of eye fixations to objects and squares mentioned in the first and second instruction components (all  $p$  values  $< .03$ ). For transition saccades, the same results were found again: During the second instruction component, participants executed more transition saccades corresponding to the first instruction components in comparisons to saccades corresponding to the third and fourth instruction component (both  $p$  values  $< .01$ ). Furthermore, during the third instruction component, participants executed more transition saccades corresponding to the first and second compared to the fourth instruction component ( $p$  values  $< .05$ ). This novel result showed that participants, who had already looked at pairs of critical objects and goal squares when they were first mentioned, re-visited these pairs repeatedly; they overtly executing these instruction components again with their eyes, presumably in an attempt to memorize them for subsequent execution. Therefore, transition saccades that were executed between critical objects and goal squares mentioned in previous instruction components were termed *rehearsal saccades* and we termed this effect of returning to previously mentioned critical objects and goal squares the *ocular rehearsal effect*.

We observed this ocular rehearsal effect only for objects mentioned in the first four component instructions, and no systematic fixation preferences for longer instructions. Figure 2 explains the likely reason for this. Note that, during the fourth instruction, fixation probabilities to the three previously mentioned objects are about equal. Thus, a fifth object cannot receive additional fixations without diminishing number of fixations to at least one already mentioned object. This result signals the

Attention deployment during memorizing and executing complex instructions limit of our participants' instruction span and explains the unsystematic fixation pattern during longer instructions.

### **Effects of Encoding on Execution**

Our interpretation of rehearsal saccades predicts that their occurrence should be related to successful instruction execution. To test this hypothesis, two analyses were carried out: In the directional analysis, a rehearsal saccade was defined by an eye movement from a critical object towards the goal square. In the non-directional analysis, a rehearsal saccade was defined by *either* a saccade from the critical object to the goal square *or* going in the opposite direction (from the goal square to the critical object). As in our previous eye movement analyses, only the first four instruction components were considered.

Participants executed directional rehearsal saccades during 28.4% (297) of all instruction component; they then correctly executed the associated instruction component in 63.3% (188) of these cases. In comparison, only 49.5% (371 of 750) instruction components were correctly executed when no rehearsal saccade was carried out, a highly significant difference ( $\chi^2(1) = 16.36, p < .001$ ). Similarly, participants executed non-directional rehearsal saccades in 34.7% (363) of all instruction components and then correctly executed the associated instruction component in 60.6% (220) of these cases. In comparison, only 49.6% (339 of 684) instruction components were correctly executed when no rehearsal saccade was carried out, again a significant effect ( $\chi^2(1) = 11.63, p < .01$ ).

### **Discussion**

This study used eye movements to investigate memory rehearsal in a spatial task. We instructed participants verbally to move objects to prescribed locations and varied the complexity of these instructions. We found that eye movements during encoding were systematically directed at both the currently

Attention deployment during memorizing and executing complex instructions mentioned and previously mentioned object-location pairs. These observations constitute linguistic cueing and rehearsal effects on eye movements, respectively. Both effects diminished in their frequency with increasing complexity of the instructions, thus suggesting an intimate link between overt attention deployment and the capacity of spatial working memory. We now discuss these two main findings and their implications for our understanding of the relationship between attention deployment and spatial working memory.

Our first finding replicates standard observations from the “visual world” literature: People look at objects and locations they hear named (Cooper, 1974; Tanenhaus, et al., 1995). This linguistic cueing of overt attention is thought to reflect an inherent connection between linguistic, spatial, and visual representations of objects (e.g. Altmann & Kamide, 2007). A novel aspect of our results is that the effect seems to predict the ability to subsequently remember these external referents for action execution. Specifically, we showed that instructions were more likely to be correctly executed when participants had previously re-visited the instructed object-location pairs. This finding extends our understanding of the linguistic cueing effect. As suggested by Ferreira, Apel, & Henderson (2008), re-fixating previously encoded objects facilitated retrieval of linguistic information related to these objects. In the current study, we showed that re-fixating relevant object also supports the execution of object relevant instructions. However, further research is necessary to determine whether people actively re-fixate relevant object in order to retain information in memory or whether rehearsing information in memory drives the eyes to relevant locations.

Our second finding identifies, for the first time, a relationship between memory span and eye movement activity. We discovered that the probability of revisiting previously named object-location pairs diminished as a consequence of

Attention deployment during memorizing and executing complex instructions the lengthening of the current instruction. This observation can be interpreted in at least two ways. First, participants may forget earlier pairs as they are being replaced by more recently instructed pairs. In this view, the overt attention deployment reflects the limitations of spatial working memory. Alternatively, however, participants simply may not be able to direct their eyes to all the previously named pairs in the time that is available to code the current instruction component. In this view, the oculomotor limitations impose a limit on participants' memory span. An interesting question for future research is to determine whether a central limit in memory capacity prevented the ocular rehearsal to extend beyond four objects, or whether instead the limited time for overt eye movements to occur during the concurrent encoding of further instructions created a motor limit to rehearsal. Support for the latter idea would underline the embodied nature of cognition, according to which our sensory and motor capacities shape our cognitive representations and abilities.

A limitation of our study is the reference to previously mentioned objects and locations when instruction complexity exceeded four components. This raises the possibility that performance limitations do not reflect either oculomotor or memory capacity limits, but also the additional demands arising from pairing a given object with more than one goal location. Although our main findings of oculomotor rehearsal and its link to performance is based on instructions of up to four components and are thus not contaminated by this confound, future work should use distinct objects and locations for each instruction component to address this possibility.

Our results extend previous work showing that *covert* attention is deployed to locations that we try to keep in working memory (Awh, Jonides, & Reuter-Lorenz, 1998; Awh & Jonides, 2001). In that work, a detection task was used to



Attention deployment during memorizing and executing complex instructions cleverly infer the deployment of attention through performance savings. In our work, we documented *overt* deployment of attention through eye movement registration. This result confirms previous work showing that overt attention (Lawrence, Myerson, & Abrams, 2004) influence working memory related tasks. Moreover, we found a limit in the number of objects that could concurrently be rehearsed, and this limit predicted performance of participants in subsequent execution.

In conclusion, we showed that overt attention is directed towards multiple objects during encoding of complex instructions to mentally simulate the execution of these instructions. Fixation towards relevant object pairs and performance strongly correlate.

## REFERENCES

- Altmann, G.T.M., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, 73, 247-264.
- Altmann, G.T.M., & Kamide, Y. (2007). The real-time mediation of visual attention by language and world knowledge: Linking anticipatory (and other) eye movements to linguistic processing. *Journal of Memory and Language*, 57, 502-518.
- Awh, E., & Jonides, J. (2001). Overlapping mechanisms of attention and spatial working memory. *Trends in Cognitive Sciences*, 5(3), 119-126.
- Awh, E., Jonides, J., & Reuter-Lorenz, P. A. (1998). Rehearsal in spatial working memory. *Journal of Experimental Psychology: Human Perception and Performance*, 24, 780-790.
- Chambers, C., & G., Juan, V.S. (2008). Perception and presupposition in real-time language comprehension: Insights from anticipatory processing. *Cognition*, 108, 26-50.
- Chambers, C. G., Tanenhaus, M. K., & Magnusson, J. S. (2004). Actions and affordances in syntactic ambiguity resolution. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 30(3), 687-696.
- Cooper, R.M. (1974). The control of eye fixation by the meaning of spoken language. *Cognitive Psychology*, 6, 84-107.
- Crovitz, H.F. (1961). Differential acuity of the two eyes and the problem of ocular dominance. *Science*, 134, 614.
- Hayhoe, M. M., Bensinger, D. G., & Ballard, D. H. (1997). Task constraints in visual working memory. *Visison Research*, 38, 125-137.

Attention deployment during memorizing and executing complex instructions

- Kamide, Y., Altmann, G.T.M., & Haywood, S.L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language, 49*, 133-156.
- Lawrence, B.M., Myerson J., Abrams, R.A. (2004). Interference with spatial working memory: An eye movement is more than a shift of attention. *Psychonomic Bulletin & Review, 11 (3)*, 488-494.
- Miles, W.R. (1930). Ocular dominance in human adults. *Journal of General Psychology, 3*, 412-430.
- Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science, 268* (16. June), 1632-1634.

### **AUTHORS' NOTES**

This work was funded by EPSRC under the Cognitive Foresight Program (EP/F028598/1). We thank SR Research for their prompt and reliable software support.

## FIGURE CAPTIONS

Figure 1. Example of the experimental display showing objects and grid.  
Not to scale.

Figure 2. Responses as a function of instruction complexity. Left panel: Number of executed actions. Middle Panel: Proportion of correctly executed actions. Right panel: Number of correct actions.

Figure 3. Proportions of fixations to objects depending on the order of mentioning in the instruction. For example, the line labeled *Critical object 1* shows the proportion of fixations towards the critical object that was mentioned in the first instruction component. The long vertical lines denote the speech onsets of the instruction components. Short vertical lines, marked by N1-N8, denote the speech onsets of the critical nouns.

Attention deployment during memorizing and executing complex instructions



Apel et al., Figure 1