

Adaptive Gaze and Hand Coordination while Manipulating and Monitoring the Environment in Parallel

Jolande Fooken^{1,2*}, Roland S. Johansson^{3†}, and J. Randall Flanagan^{2†}

¹Department of Psychology and Centre for Cognitive Science, Technical University of Darmstadt, Darmstadt, Germany

²Department of Psychology and Centre for Neuroscience Studies, Queen's University, Kingston, Ontario, Canada

³Department of Medical and Translational Biology, Umeå University, Umeå, Sweden

*Corresponding author: jolande.fooken@tu-darmstadt.de

† Equal contribution

Credit authorship contribution statement

Conceptualization (JF, RSJ, JRF), Methodology (JF, RSJ, JRF), Validation (JF, RSJ, JRF), Formal analysis (JF, RSJ, JRF), Resources (RSJ, JRF), Data Curation (JF, RSJ), Writing – Original Draft (JF, RSJ, JRF), Writing – review & editing (JF, RSJ, JRF), Visualization (JF, RSJ, JRF), Funding acquisition (JF, RSJ, JRF)

Declaration of Competing Interest

The authors have no competing financial interests to disclose.

Acknowledgments

This work was supported by a Deutsche Forschungsgemeinschaft (DFG) Research Fellowship to JF (FO 1347/1-1), the Swedish Research Council Project (grant 22209 to RSJ), the Canadian Institutes for Health Research (grant RGPIN/05944-2019 to JRF), and the Natural Science and Engineering Research Council of Canada (grant 156173 to JRF).

Abstract

1 In real-world tasks, visually guided manual actions are often performed in parallel with visual
2 monitoring of the environment, creating competition for gaze. This study examined how
3 participants coordinated gaze and hand movements while performing a manipulation task—
4 grasping a small ball with either fingertips or tweezers and depositing it into a slot—while
5 concurrently tasked with monitoring a display for probabilistically occurring letter changes.
6 Gaze fixations away from the text display were exclusively directed to the ball and slot, and
7 were temporally linked to key contact events (ball contact and slot entry). In fingertip trials,
8 action-task fixations, when present, were brief and could support directing the hand using
9 peripheral vision. In tweezer trials, action-task fixations were more frequent and helped both
10 directing the hand with peripheral vision and guiding it with central vision. Participants
11 identified and leveraged the temporal regularities in the monitoring task to reduce competition
12 for gaze resources between tasks. Specifically, they adjusted both their gaze behaviour and hand
13 movement timing to lower the likelihood that action task fixations would coincide with letter
14 changes. These findings demonstrate how individuals integrate sensorimotor demands with
15 environmental statistics to manage competition for gaze when concurrently acting on and
16 monitoring the environment.

Introduction

17 Gaze fixations, occurring between eye movements, play a vital role in both perceiving the world
18 and in planning and controlling actions (Hayhoe, 2017; Kowler, 2011; Land, 2006; Yarbus,
19 1967). Although research has extensively examined gaze control in action and perception tasks
20 independently, real-life scenarios often demand concurrent performance of visually guided
21 actions and visual perception tasks (Fooken et al., 2023; Land and Furneaux, 1997), leading to
22 a competition for gaze resources. For instance, at a dinner party, diners use their gaze to control
23 their manual actions, such as handling objects, while also using their gaze to survey their
24 surroundings and engage in conversations. Under such circumstances, one would expect that
25 gaze would only be directed to the action task when the support of gaze is most important.

26 To our knowledge, no prior research has delved into the control policies and strategies
27 governing gaze allocation when there are competing demands for gaze from manual actions and
28 environmental monitoring. To address this gap, we designed an experiment where participants
29 simultaneously performed an object manipulation task and a visual monitoring task that
30 required central vision. The manipulation task involved repeatedly grasping a small ball situated
31 on a platform and inserting it into a slot in a vertical tube, from which the ball returned to the
32 platform (Fig. 1A). To manipulate the visuomotor demands, participants were required to
33 perform the task using both their fingertips and tweezers. At the same time, participants
34 monitored a text display for randomly occurring letter changes. Successful ball drops were
35 rewarded, while failures to detect letter changes incurred penalties.

36 This study aimed to examine three novel hypotheses regarding how gaze allocation might be
37 optimized during the concurrent execution of an action task and a perception task that compete
38 for gaze. First, we hypothesized that the prevalence and function of gaze fixations directed to
39 the action task would depend on the specific task phase and the visuomotor demands, influenced
40 by the end-effector employed (fingertips and tweezers). Second, we hypothesized that, by
41 observing the temporal statistics of relevant visual events in the environment, individuals would
42 anticipate when events that require central vision will occur and use this information to allocate
43 gaze more efficiently between tasks. Third, we hypothesized that individuals would reduce
44 competition between tasks for gaze resources by adjusting the timing of the action task and,
45 consequently, the timing of required action task fixations. The rationale and motivation for these
46 hypotheses are elucidated below.

47 Object manipulation tasks involve a sequence of action phases delimited by kinematic or
48 mechanical events (Flanagan et al., 2006; Johansson et al., 2001). The action phases in our ball-

drop task are illustrated in Fig. 1B. When action tasks are performed in isolation, gaze is almost exclusively directed to targets of action—the ball and slot in our ball-drop task—with the function of gaze changing across action phases. When moving the hand, or object in hand, towards the vicinity of a target—as in the reach and transport phases of our task—individuals typically fixate the target. Such fixations can enable fast, automatic feedback control mechanisms that use peripheral vision and gaze-related signals to *direct* the hand toward the target (Goodale et al., 1986; Saunders and Knill, 2004, 2003). Once the hand gets close to the object, and more slowly approaches it—as in the ball and slot approach phases in our task—gaze may remain on the target, in which case central vision can be used to *guide* the hand, or grasped object, to the target through more deliberate closed-loop feedback control (Ballard et al., 1992; Johansson et al., 2001; Land, 2006). Once the hand or object in hand contacts the target—ball contact and slot entry marking the start of the grasp and slots phases in our task—tactile feedback becomes available. This tactile feedback provides confirmation that contact was achieved. If gaze remains on the target until contact, central vision can also be used to *check* that contact was achieved (Säfström et al., 2014). Within this framework of gaze functions in isolated tasks, the current study addresses whether and how corresponding gaze functions are expressed when there is competition for central vision.

We expected that when the ball-drop is performed concurrently with the visual monitoring task, all fixations directed to the action task would be located at the ball and slot. Regarding our first hypothesis, we predicted that when using the fingertips, action task fixations, if observed, would primarily serve the purpose of *directing* the hand via peripheral vision. We predicted that fixations involved in *guiding* the hand via central vision could be omitted when using fingertips because, once the hand (or ball in hand) arrives at the target, haptic feedback can be used to correct positioning errors. In contrast, we predicted that when using tweezers, gaze fixations, in addition to being involved in *directing* the hand, would also be involved in *guiding* the hand, as tweezers require greater spatial precision, particularly in grasping the ball, and offer limited tactile feedback about the contact state. Note that impaired tactile sensibility is known to increase reliance on visual feedback for object manipulation (Brink and Mackel, 1987; Chemnitz et al., 2013; Jenmalm and Johansson, 1997; Jerosch-Herold, 1993).

Previous work has shown that when concurrently monitoring two locations to detect probabilistic events, individuals can optimize their gaze allocation by learning the temporal regularities of the events at each location and adjusting their gaze accordingly (Hoppe and Rothkopf, 2016). This raises the question of whether people can similarly learn and exploit temporal regularities of environmental visual events when concurrently performing a visual

monitoring task and a visually guided action task. In the monitoring task, letter changes occurred at intervals randomly sampled from a uniform distribution between 1.5 to 6.5 seconds (s). Thus, each letter change was followed by a 1.5 s ‘silent period’, during which no change could occur. After this period, the probability of the next letter change (i.e., the hazard rate) linearly increased. With respect to our second hypothesis, we predicted that participants would exploit this temporal structure to primarily allocate their gaze to the manipulation task during silent periods and shortly afterwards, when the probability of a letter change was low. In contrast to visual monitoring of environmental events, where timing demands on central vision are externally determined, individuals can, in principle, adjust the timing of their own actions and, consequently, the timing of the required action task fixations. For our third hypothesis, we predicted that individuals would mitigate competition for gaze resources between tasks by strategically adjusting the timing of the action task. This hypothesis relies on the assumption that participants not only learn the statistical properties of letter changes in the monitoring task, but also possess knowledge of when and where fixations are required for the action task as it unfolds.

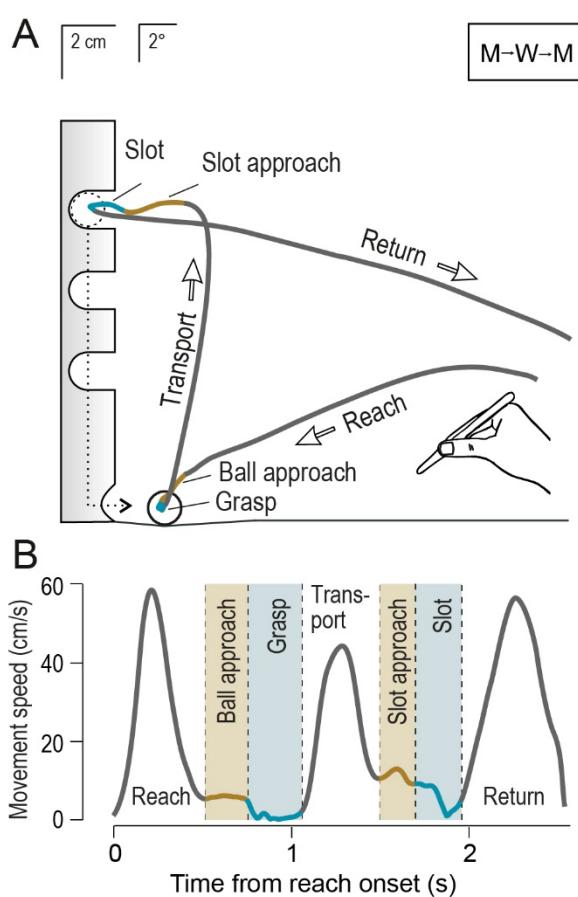


Figure 1. Apparatus and action phases in the ball-drop task. (A) Experimental setup from the participant's perspective showing the path of the end-effector tip during a single tweezer trial. Movements were executed in a plane parallel to the participant's coronal plane, positioned 40 cm from the participant's eyes. (B) Corresponding velocity profile of the end-effector. (A, B) The task was divided into 7 consecutive action phases separated by distinct kinematic or contact events (see Methods). (1) *Reach phase*: Begins when the hand departs from its starting position and is characterized by a bell-shaped velocity profile. (2) *Ball approach phase*: Begins at a minimum (or inflection point) in the velocity profile. (3) *Grasp phase*: Begins upon first contact with the ball. (4) *Transport phase*: Begins when the ball is lifted from the platform and characterized by a bell-shaped velocity profile. (5) *Slot approach phase*: Begins at a minimum (or inflection point) in the speed profile and is characterized by slow movement. (6) *Slot phase*: Begins when the tips of the end-effector holding the ball enter the slot. (7) *Return phase*: Begins when the ball is released and exhibits a bell-shaped velocity profile. The trial concludes when the hand returns to its starting position.

Results

98 Eleven participants performed, at their own pace, 30 consecutive trials of the ball-drop task in
99 each of four experimental conditions. Participants performed the task using either their
100 fingertips or tweezers, either as a standalone activity (referred to as ‘single task’ conditions) or
101 concurrently with the visual monitoring task (referred to as ‘dual task’ condition). In each ball-
102 drop trial, participants reached for and grasped a ball positioned on a platform adjacent to the
103 base of a vertical tube. They then transported the ball to one of three slots within the tube,
104 inserted it, and released it before returning their hand to its starting position. After the ball was
105 released, it descended through the tube and returned to its starting position on the platform.
106 One second after the ball returned to the start position (or was already located in the start
107 position in the first trial), an auditory signal instructed the participant about which slot to use.

108 Figure 1A shows the end-effector movement path during a tweezer trial, from the participant's
109 perspective. The motion occurred in a plane parallel to the participant's coronal plane, situated
110 in front of the body. Each ball-drop trial was divided into seven consecutive action phases,
111 delimited by distinct kinematic or mechanical contact events (see Figure 1B). The visual
112 monitoring task involved detecting changes in a letter displayed on a text screen positioned in
113 the upper right quadrant of the participant's scene (Fig. 1A; for details see Methods).
114 Participants received rewards for successful ball drops and were penalized for missing letter
115 changes, which were signaled by an auditory tone and visual feedback on the display.

116 We first examine gaze and end-effector coordination in the single task conditions performed
117 using either fingertips or tweezers. The conditions provide baselines for comparison with the
118 corresponding dual task conditions.

Single task conditions

119 Figures 2 A-B show gaze and end-effector paths during representative trials with the fingertips
120 and tweezers, respectively. In both fingertip and tweezer trials, participants typically fixated the
121 ball while reaching for it, shifted their gaze to the slot around the time the ball was grasped, and
122 maintained the fixation of the slot until around the time the ball was dropped. Gaze then returned
123 to the ball's start position at the tube's base. Figures 2 C-D show for each end-effector, the
124 average speed of the tip of the end-effector (top panels) and the instantaneous probabilities of
125 gaze fixating the ball at its start position or the target slot, as a function of time (bottom panels).
126 Note that these figures are based on aggregate data from all trials and participants. To temporally
127 align trials while preserving phase-specific information, we normalized the duration of each
128 phase in each trial to the median phase duration across all trials within each condition.

129 In both fingertip and tweezer trials, participants predominantly fixated the ball throughout the
130 reach phase, though gaze was occasionally located at the slot during the reach phase of fingertip
131 trials. Similarly, during the transport phase, participants primarily fixated the slot. The timing
132 of the gaze shift from the ball to the slot differed significantly between fingertip and tweezer
133 trials ($t_{10} = 13.24; p < 0.001; d = 3.99$). In fingertip trials, the shift occurred just before contact
134 (-0.06 ± 0.05 s; mean \pm sem), typically during either the late reach phase or the ball-approach
135 phase. However, in tweezer trials, the shift occurred well after ball contact (0.29 ± 0.11 s),
136 mainly during the grasp phase. This finding supports the prediction that tweezers require greater
137 reliance on central vision to establish stable grasp on the ball.

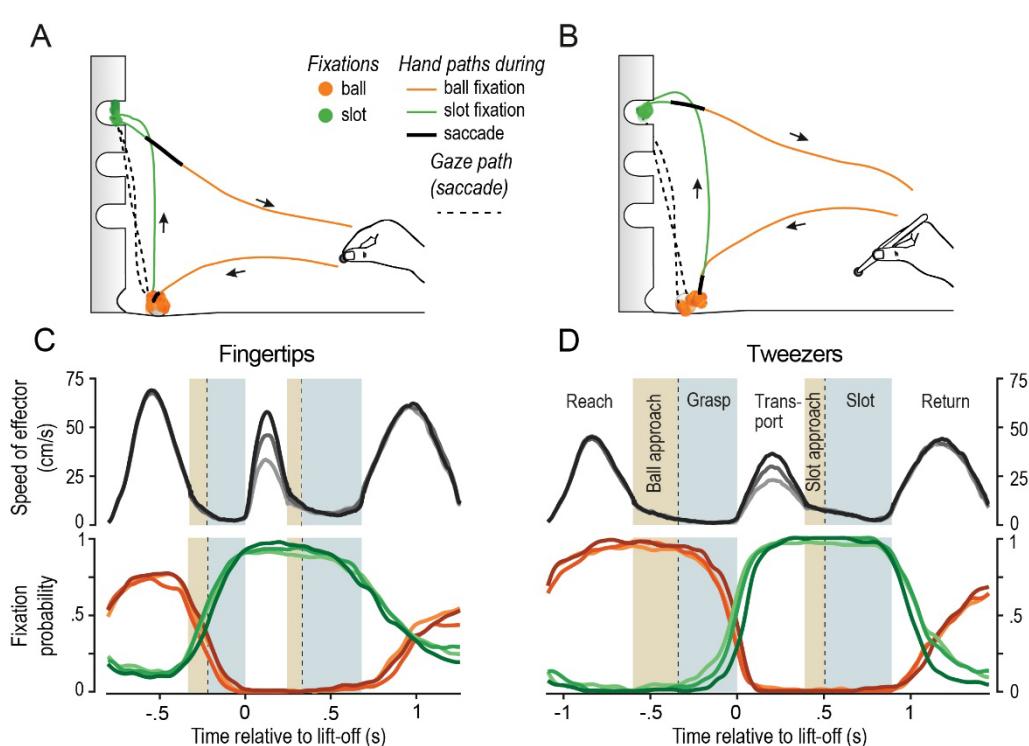


Figure 2. Gaze-hand coordination in the single-task condition. (A, B) Gaze and end-effector paths in single fingertip (A) and tweezer (B) trials. Fixations are color-coded according to landmark (ball, slot), and end-effector paths (tip of the fingers or tweezers) are color-coded based on the current gaze state: fixating the ball, fixating the slot, or executing a saccade. Note that the ball was fixated during the reach and return phases while the slot was fixated during the transport phase. (C, D) Average end-effector speed (black) and the probabilities of fixating the ball (orange) and slot (green) as a function of time relative to lift-off, for fingertip (C) and tweezer (D) trials. Separate traces are shown for each slot (top, middle, and bottom coded dark to light). Alternating white, brown, and grey-blue regions indicate movement phases labelled in D. Plots aggregate data from all trials conducted by all participants, with phase durations in each trial normalized to the median duration of the phase. Fixations were almost always directed to the landmarks (ball and slot); however, the combined probabilities of these fixations may be less than 1 due to saccades occurring between the landmarks.

Dual task conditions

138 In the dual task conditions, participants divided their gaze fixations between the text display
139 and the action-related landmarks, consistent with the fact that detecting letter changes in the
140 visual monitoring task required central vision (Fig. 3A-B). As a result, the likelihood of fixating
141 the ball or slot during the ball-drop task was reduced compared to the single task conditions,
142 regardless of end-effector used (Fig. 3C–D). Additionally, the durations of occurring action
143 landmark fixations were consistently shorter.

Propensity and patterns of action fixations depend on the end-effector

144 In fingertip trials, participants primarily fixated the display during the reach, ball approach, and
145 grasp phases (88% of trials), suggesting that grasping the ball generally did not require central
146 vision. In trials where participants did fixate the ball, these fixations occurred during the reach
147 phase, with gaze typically shifting away before the ball approach phase and almost never
148 remaining on the ball after contact. Transporting the ball and inserting it into the slot were also
149 frequently performed with gaze remaining on the display. However, participants briefly fixated
150 the slot in about half of the trials (51%), as in the example shown in Fig. 3A. The probability of
151 fixating the slot peaked midway through the transport phase, then steadily declined,
152 approaching zero by the end of the slot phase (Fig. 3C). These findings show that in fingertip
153 trials, inserting the ball into the slot often did not require central vision. However, the likelihood
154 of fixating the slot during the slot approach and slot phase was higher than fixating the ball
155 during the ball approach and grasp phase.

156 Consistent with the fixation probabilities described above, we observed two predominant gaze
157 patterns in fingertip trials: ‘display-only’, where gaze remained on the display throughout the
158 trial, and ‘slot’, where gaze shifted from the display to the slot and then back to the display (see
159 Figs. 3E-F). To examine the relationship between gaze pattern and manual performance, we
160 compared the action phase durations in display-only and slot trials. Only participants ($N=9$)
161 who demonstrated both fixation patterns were included. The duration of the slot phase was
162 shorter when participants fixated the slot ($0.30 \pm .054$ s) compared to when they did not (0.378
163 $\pm .056$ s; $t_8 = 3.76$, Holm-Bonferroni adjusted $p = 0.03$). No other phase durations were affected
164 by gaze pattern (adjusted $p > 0.51$ in all cases).

165 In tweezer trials, participants were much more likely to fixate both the ball and the slot
166 compared to fingertip trials (Fig. 3D). Participants almost always fixated the ball before contact
167 (88% of trials), and the slot before slot entry (89% of trials). The probability of fixating the ball
168 peaked towards the end of the reach phase and remained relatively high during the ball approach

169 and most of the grasp phase. Similarly, the likelihood of fixating the slot peaked towards the
 170 end of the transport phase and remained elevated during the slot approach and slot phases. These
 171 findings suggest that, in general in tweezer trials, central vision was crucial for both grasping
 172 the ball and inserting it into the slot.

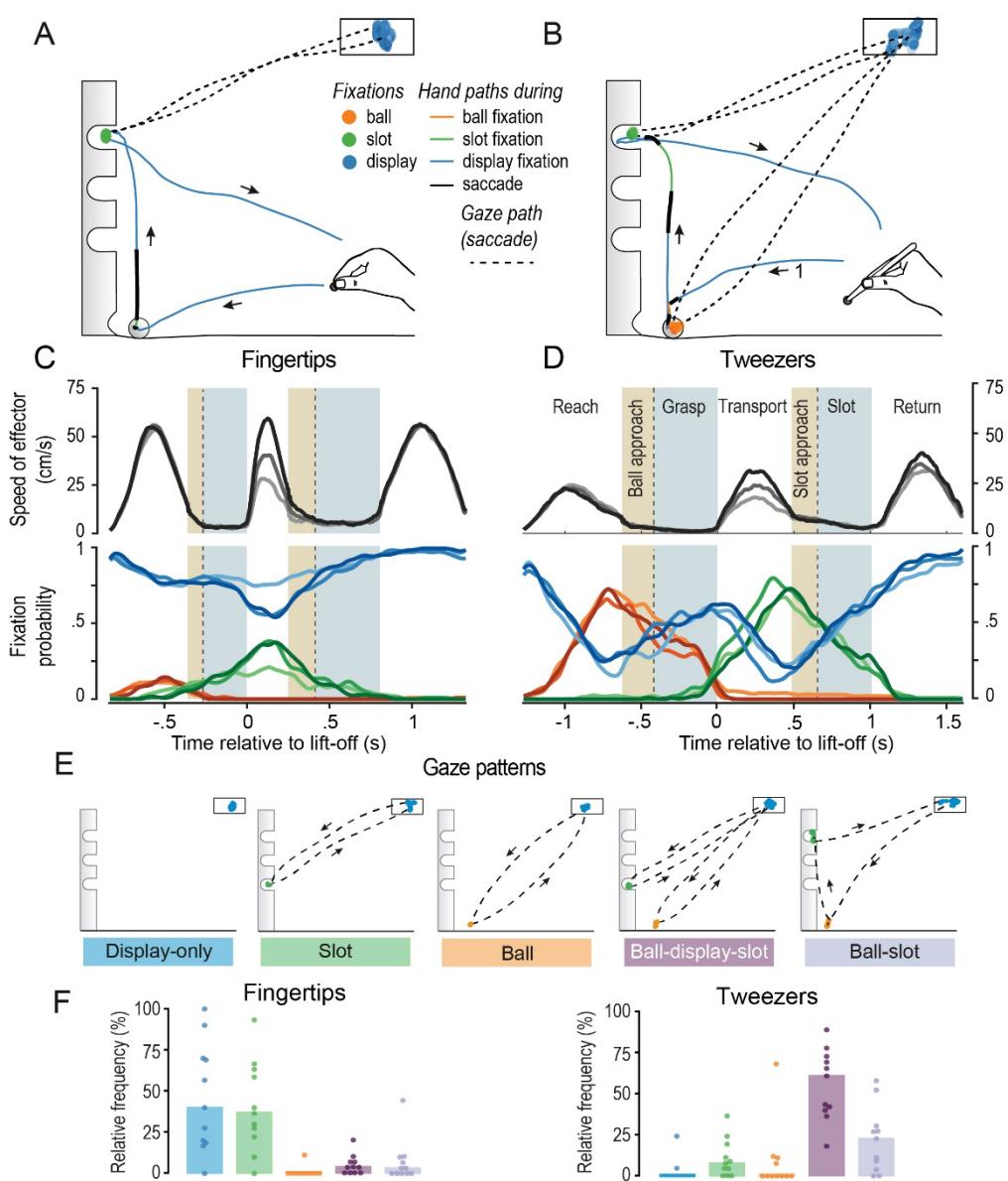


Figure 3. Gaze-hand coordination during dual-task conditions. (A, B) Gaze and end-effector paths in single fingertip (A) and tweezer (B) trials. Fixations are color-coded by landmark (ball, slot, or display), and end-effector paths (tip of fingers or tweezers) are color-coded based on the current gaze state (fixating the ball, slot or display, or executing a saccade). (C, D) Average end-effector speed (black) and the probabilities of fixating the ball (orange), slot (green), and display (blue) as a function of time relative to lift-off, for fingertip (C) and tweezer (D) trials. Separate traces are shown for each slot (top to bottom coded dark to light). Alternating white, brown, and grey-blue regions indicate movement phases labelled in D. Plots combine all trials from all participants with phase durations in each trial normalized to the median duration of the phase. (E) The five distinct gaze patterns observed during dual-task trials are illustrated with data from single trials. (F) Average percentage of each gaze pattern (color-coded as in E) observed across participants in fingertip and tweezer trials. Individual participants are represented by dots.

173 The predominant gaze pattern in tweezer trials was 'ball-display-slot', where participants shifted
174 their gaze from the display to the ball, back to the display, then to the slot and finally returned
175 to the display (see example in Fig. 3B). The second most common pattern was 'ball-slot', where
176 participants shifted their gaze from the display to the ball and then directly to the slot before
177 returning to the display (see Figs. 3E-F). To examine the relationship between gaze pattern and
178 manual performance, we compared the action phase durations between the ball-display-slot and
179 ball-slot trials. Only participants who exhibited both gaze patterns ($N = 8$) were included in this
180 analysis. We found that the transport phase was shorter when gaze shifted directly from the ball
181 to the slot ($0.32 \pm .051$ s) compared to when gaze fixated the display between the ball and slot
182 fixations (0.55 ± 0.17 s; $t_7 = 4.71$, Holms-Bonferroni adjusted $p = 0.01$). No other phase
183 durations were associated with gaze pattern (Holms-Bonferroni adjusted $p > 0.22$ in all cases).

184 As in the single-task conditions, in the dual task conditions the ball-drop task was performed
185 more slowly with the tweezers (2.77 ± 0.4 s) than with the fingertips (2.11 ± 0.23 s; $t_{10} = 5.61$,
186 $p < 0.001$; $d = 1.69$). Similar to the single task conditions, the greater time taken to perform the
187 task with tweezers was due to increased durations of the reach, grasp, and transport phases
188 (Holms-Bonferroni adjusted $p < .01$ in all three cases).

Action fixations are temporally anchored to contact events

189 To examine the trial-by-trial coordination between the timing of action fixations and specific
190 kinematic or contact events, we conducted a multiple linear regression analysis to assess how
191 well the onset and offset times of ball and slot fixations could be predicted by the timing of the
192 following six events delimiting action phases: (1) start of the reach phase, (2) start of the ball
193 approach phase, (3) time of first ball contact (start of the ball grasp phase), (4) time of ball liftoff
194 (start of the ball transport phase), (5) start of the slot approach phase and (6) time for slot entry
195 (start of the slot phase). To mitigate structural multicollinearity, predictors were centered
196 individually for each participant by subtracting their mean. Furthermore, participants were
197 included as a categorical nuisance factor to account for variance due to differences in task
198 performance speed. Separate regression analyses were conducted for fixation onsets and offsets,
199 as well as for each action landmark (ball and slot) and end-effector (fingertips and tweezers).

200 In both fingertip and tweezer trials, the initiation and termination of ball fixations were best
201 predicted by the time of first ball contact, while the initiation and termination of slot fixations
202 were best predicted by the time of slot entry. For fixation onset times, in fingertip trials, ball
203 fixation onset was predicted solely by the time of first ball contact ($t_{1,32} = 2.84$; $p = 0.008$), and

204 slot fixation onset solely by the time of slot entry ($t_{1,158} = 9.28; p < 0.001$). Similarly, in tweezer
205 trials, the ball fixation onset was primarily predicted by the time of first ball contact ($t_{1,243} =$
206 $6.26; p < 0.001$), while slot fixation onset was solely predicted by the time of slot entry ($t_{1,244} =$
207 $11.2; p < 0.001$). Comparable patterns were observed for fixation offset times. In fingertip trials,
208 ball fixation offset was solely predicted by the time of first ball contact ($t_{1,32} = 3.08; p = 0.004$),
209 while slot fixation offset was best predicted by the time of slot entry ($t_{1,158} = 8.30; p < 0.001$).
210 Similarly, in tweezer trials, ball fixation offset was best predicted by the time of first ball contact
211 ($t_{1,243} = 5.18; p < 0.001$), and slot fixation offset was solely predicted by the time of slot entry
212 ($t_{1,244} = 9.04; p < 0.001$). These findings demonstrate a robust temporal coupling between the
213 initiation and termination of action landmark fixations and their associated contact events,
214 consistent across both end-effectors.

215 Figure 4 illustrates this temporal relationship, showing that the timing of ball and slot fixation
216 periods relative to ball contact and slot entry, respectively, remained quite consistent across
217 trials. The left panels of Figs. 4A and B show the timing of ball fixations and associated action
218 phases relative to the time of first ball contact (time = 0). Each row represents an individual trial
219 and data from all participants are shown. Orange lines indicate periods of ball fixation, while
220 dots mark the onset of the reach, ball approach and transport phases in each trial. The trials are
221 arranged in order of ball approach phase duration. Similarly, the left panels of Figs. 4C and D,
222 show the timing of slot fixation and associated action phases relative to the time of slot entry.
223 Green lines represent periods of slot fixation, and dots mark the onset of the transport, slot
224 approach, and return phases for each trial. These trials are sorted by slot approach phase
225 duration. In some trials, the ball or slot were re-fixated (depicted by purple lines in Figs. 4A-
226 D). Re-fixations typically occurred when multiple attempts were required to successfully grasp
227 and lift the ball or insert it into the slot. Notably, the durations of both the ball and slot approach
228 phases could vary considerably across trials.

229 The scatter plots on the right side of each panel in Fig. 4 illustrates the position of each trial in
230 its test block (No. 1 – 30), ranked by the duration of the ball or slot approach phase, as well as
231 the position of each participant’s trials in the same ranking (Pt. Nos. 1 -11). Colored dots (orange
232 for ball and green for slot) represent trials with an action landmark fixation, while black dots
233 indicate trials without such fixations. The absence of apparent structure in these scatter plots
234 suggests that neither the decision to fixate the action landmark nor the variation in ball and slot
235 approach phase durations were influenced by trial position. Furthermore, these scatter plots
236 indicate that, overall, participants exhibited similar behaviour.

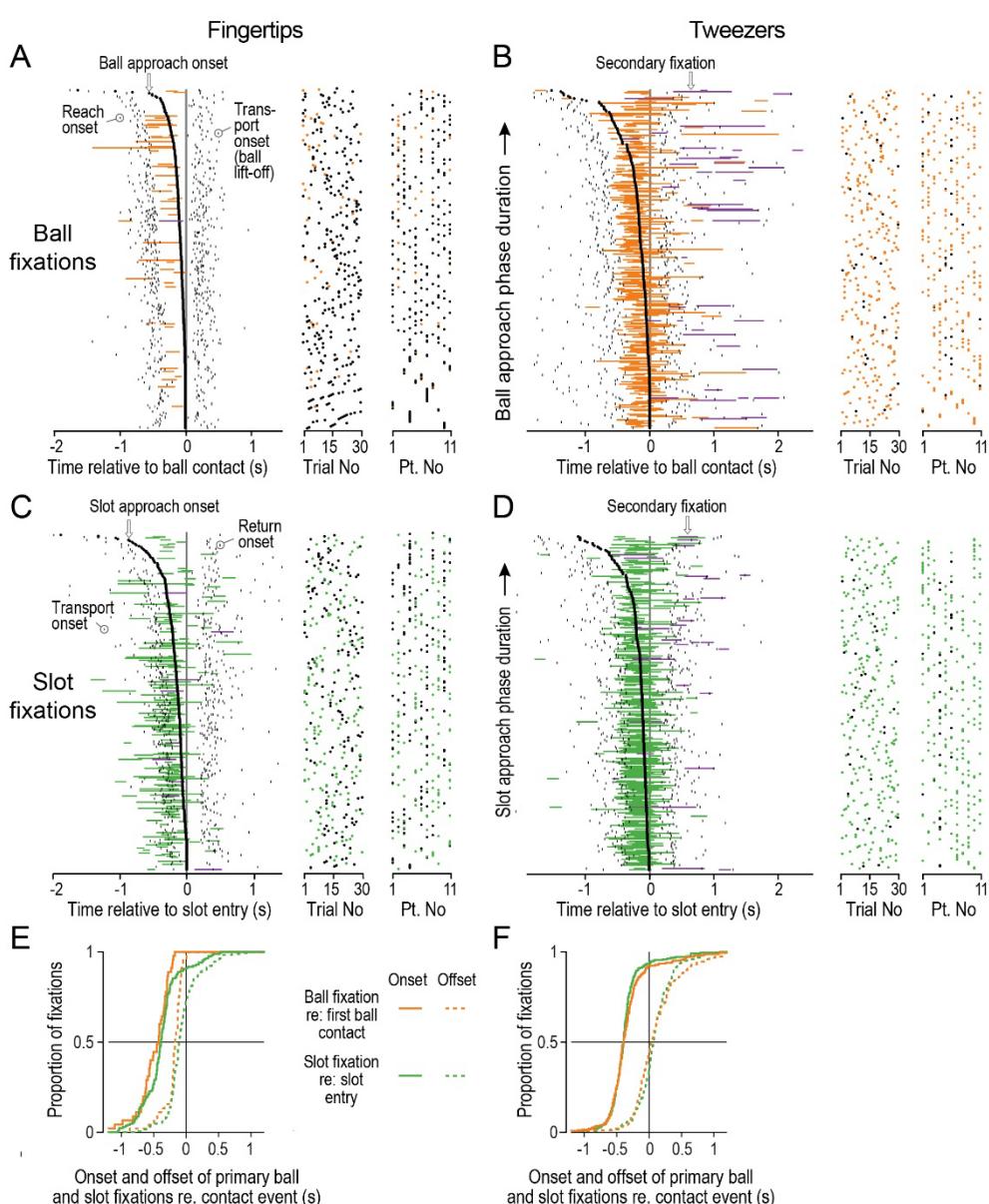


Figure 4. Timing of action fixations. (A-D) These panels show timing data for all trials stacked vertically. (A, B) The left column shows periods of initial (orange lines) and secondary (purple lines) ball fixations, aligned to the time of initial ball contact (grey vertical line), in fingertip (A) and tweezer (B) trials. For each trial, the onsets of the reach and transport phases are marked by small black dots, and the onset of the ball approach phase is marked by larger black dots. Trials are sorted by ball approach phase duration. The middle and right columns show trial number and participant number, with orange and black dots representing trials with and without a ball fixation, respectively. (C, D) Corresponding plots for initial (green) and secondary (purple) slot fixations, aligned to the time of slot contact. Small black dots mark the onsets of the transport and return phases, and the larger black dots indicate the onset of the slot approach phase. Data are sorted by the slot approach phase duration. In the middle and right columns, the green and black dots depict trials with and without a slot fixation. (E) Cumulative distributions of ball and slot fixation onsets and offsets, aligned to the initial ball contact and slot entry, respectively, in fingertip trials. (F) Corresponding distributions in tweezer trials.

237 The timing of action fixation onsets, relative to contact events, showed remarkable consistency
238 across action landmarks and end-effectors. Both ball and slot fixations typically began
239 approximately 0.4 s before ball contact and slot entry, respectively (solid line curves in Figs. 4E
240 and F). The timing of fixation offsets was also consistent across action landmarks but influenced
241 by the end effector used. In fingertip trials, gaze generally shifted away from both the ball and
242 the slot before the contact event, with an average lead time of about 0.15 s (dashed line curves
243 in Fig. 4E). In contrast, in tweezer trials, gaze typically shifted shortly after the contact event,
244 with an average lag of about 0.05 s (dashed line curves in Fig. 4F).

Variability in the function of action fixations across trials

245 The variability in the timing of ball and slot fixations, relative to their associated action phases,
246 suggests that the functions of these fixations—such as *directing* and *guiding*—may vary across
247 trials. To investigate this possibility, we analyzed the function of each individual ball and slot
248 fixation, acknowledging that an individual fixation could serve multiple functions. A fixation
249 was classified as *directing* if the ball or slot was fixated for at least 100 ms during the reach or
250 transport phase, respectively. Similarly, a fixation was classified as *guiding* if the ball or slot
251 was fixated for at least 100 ms between the start of the ball or slot approach phase and the end
252 of the grasp or slot phase, respectively (i.e., the combination of the approach and manipulation
253 phases). Additionally, fixations were considered as checking when the ball or slot was fixated
254 for any period of time after the end of the grasp or slot phase, respectively, allowing visual
255 confirmation of the action’s completion (Säfström et al., 2014). Figure 5A illustrates examples
256 of slot fixations in tweezer trials demonstrating these distinct functions.

257 In fingertip trials, ball fixations primarily served a *directing* function, whereas slot fixations
258 also frequently played a *guiding* role (Fig. 5B). In contrast, in tweezer trials, both ball and slot
259 fixations were more evenly distributed between *directing* and *guiding* (Fig. 5C). In both
260 conditions, a small proportion of fixations were involved in *checking* (Fig. 5B and C).
261 Interestingly, in fingertip trials, most of the ball and slot fixations served only one function (80%
262 overall), whereas in tweezer trials, this proportion was lower (48% overall). This distinction is
263 highlighted by the thin solid bars within each wide bar in Figs. 5B and C.

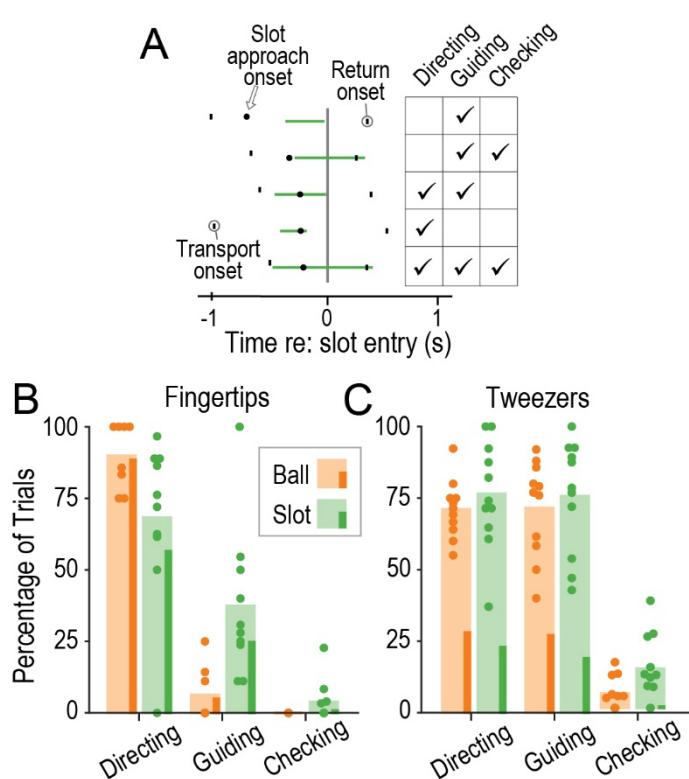


Figure 5. Classification of fixation functions. (A) Examples of slot fixations from tweezer trials that serve different functions. (B, C) Wide bars show the mean percentage of ball (orange) and slot (green) fixations engaged in *directing*, *guiding*, and *checking* across participants in fingertip (B) and tweezer (C) trials. Note that a given fixation could be engaged in more than one function. Circles represent individual participants, except for the circles at zero which represent all participants with zero. The thin bars represent percentages of single-function fixations.

Impact of letter change statistics on the allocation of gaze resources

266 In the following section, we investigate whether participants can learn and exploit the statistical
 267 properties of letter changes (LCs) to mitigate competition for gaze between the visual
 268 monitoring and manipulation tasks. Our findings reveal that participants adapt their gaze
 269 behaviour both directly, by selecting different gaze patterns, and indirectly, by modifying their
 270 manual actions in response to LC statistics (see next section).

271 In the dual-task condition, the interval between LCs was drawn from a uniform distribution
 272 spanning 1.5 to 6 s. As a consequence, there was a 1.5 s “silent period” following a LC when
 273 the next LC could not occur. Following the silent period, the likelihood of the next LC occurring
 274 (i.e., the hazard rate) gradually increased over the subsequent 5 s rate. On average, there were
 275 1.08 and 1.41 LCs per trial in fingertip and tweezer trials, respectively. Importantly, participants
 276 were very good at detecting these LCs and therefore could, in principle, exploit the LC statistics.
 277 Participants detected these LCs with an accuracy of $88.8 \pm 11.8\%$ and $87.1 \pm 9.1\%$, respectively
 278 (mean \pm standard deviation across participants).

279 To analyze whether LC detection influenced the timing action task fixations, we compared the
 280 observed frequency distributions of ball and slot fixation onsets—relative to the most recently
 281 detected LC preceding each fixation—with the distributions expected if fixation onsets occurred
 282 randomly with respect to LCs. For each end-effector, we conducted separate analyses for trials
 283 exhibiting each of the primary gaze patterns: ball-display-slot, ball-slot, and slot trials. In both

284 fingertip and tweezer trials in which participants fixated both the ball and slot (i.e., ball-slot and
285 ball-display-slot trials), the distribution of ball fixations onsets deviated from the expected
286 random distribution (Kolmogorov-Smirnov test, $p \leq 0.01$ in all four cases). Specifically, ball
287 fixation onsets during the silent period occurred more frequently than expected by chance (top
288 panels of Fig. 6A).

289 In tweezer trials, the choice of gaze pattern was strongly influenced by the timing of ball fixation
290 onset relative to the preceding LC. When the ball was fixated during the silent period,
291 participants were equally likely to adopt either the ball-slot or ball-display-slot pattern.
292 However, if the ball was fixated after the silent period, the ball-display-slot pattern was
293 predominantly chosen. That is, participants rarely shifted their gaze directly from the ball to the
294 slot (ball-slot pattern) unless the ball was fixated within the silent period. Conversely, when ball
295 fixation took place after the silent period, participants almost always fixated the display before
296 fixating the slot (ball-display-slot pattern).

297 As expected, in both fingertip and tweezer trials where gaze shifted directly from the ball to the
298 slot (ball-slot pattern), a peak in the distribution of slot fixation onsets followed the peak in the
299 distribution of ball fixation onsets in the silent period (bottom panels of Fig. 6A). This occurred
300 because, in the vast majority of these trials (96.4%), the last LC detected before the ball fixation
301 was also the last LC before the slot fixation. Both of these distributions differed from the
302 expected random distributions (KS test, $p < 0.02$ in both cases). Similarly, in fingertip and
303 tweezer trials with the ball-display-slot pattern, there is a peak in the distribution of slot fixation
304 onsets that is shifted to the right of the peak in the distribution of ball fixations, also representing
305 trials in which last LC detected before the ball fixation was also the last LC before the slot
306 fixation. In fingertip trials with the ball-display-slot pattern, the distribution of slot fixations
307 differed from the expected random distribution (KS test, $p < 0.01$). Note that in tweezer trials
308 with the ball-display-slot pattern, there is also an earlier peak in the distribution of slot fixations
309 occurring during the silent period, representing trials where a LC was detected while gaze
310 fixated the display prior to shifting to the slot. As a consequence of this earlier peak, the
311 distribution of slot fixations did not differ from the expected random distribution (KS test, $p =$
312 0.4). In both fingertip and tweezer trials, where only the slot was fixated (slot gaze pattern), the
313 distribution of slot fixation onsets did not differ significantly from the expected random
314 distribution (KS test, $p > 0.06$ in both cases; bottom panels of Fig. 6A).

315 Together, these results demonstrate that participants could adapt their gaze behaviour to exploit
316 the statistical properties of LCs to mitigate competition for gaze resources between the visual
317 monitoring and object manipulation tasks.

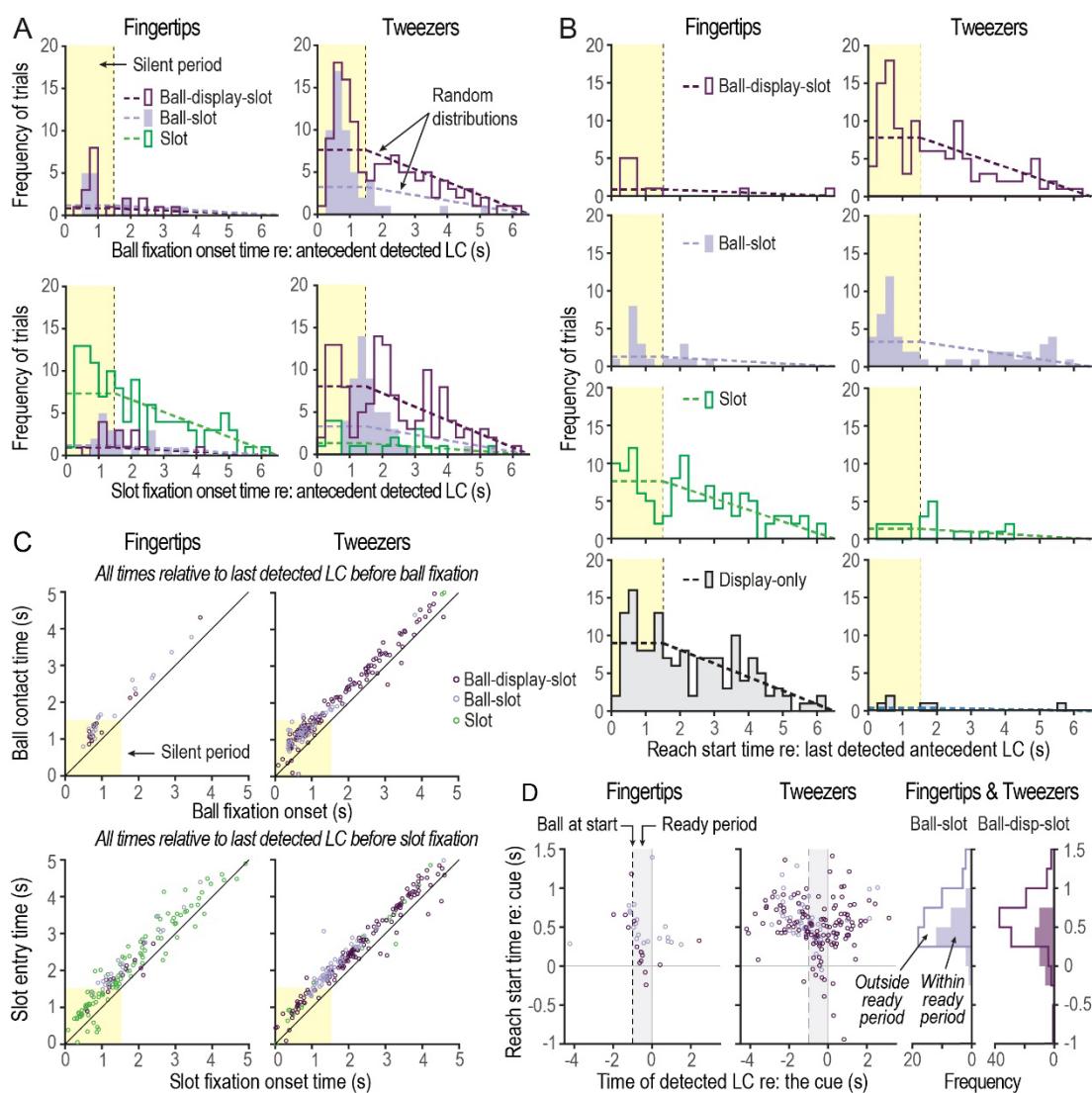


Figure 6. Impact of timing of letter changes (LCs) on gaze fixations and hand movements. (A) Frequency distributions, combining trials from all participants, of ball (top) and slot (bottom) fixation onsets—relative to the time of the last detected LC before the fixation onset—in fingertip (left) and tweezer (right) trials. Yellow regions indicate the silent period. Separate distributions are shown for trials with the main gaze patterns. Dashed lines show expected distributions assuming fixation onsets occurred randomly and independently LC timing. The expected frequency is constant within the silent period, during which the hazard rate (the probability of a LC occurring if one has not yet occurred) remains at 0, and then decreases, at a constant rate, over the next 5 s as the hazard rate increases from 0 to 1. (B) Corresponding distributions of reach start times relative to the last detected antecedent LC in fingertip (left) and tweezer (right) trials. Separate plots are shown for the four main gaze patterns. (C) The top row shows, for fingertip and tweezer trials, the relationship between ball fixation onset and ball contact times, relative to the last detected LC before fixation onset. Dots represent trials from all participants and are colour-coded by gaze pattern. The bottom row shows the corresponding relationships between slot fixation onset and slot entry times. (D) Relationship between reach start time and time of nearest detected LC, both relative to cue onset, in fingertip and tweezer trials with ball-slot and ball-display-slot patterns. (E) Frequency distributions of the reach start times in trials with ball-slot and ball-display-slot gaze patterns, aligned with the data shown in (D). These distributions combine fingertip and tweezer trials. Separate distributions are shown for trials in which the LC was within or outside the ‘ready period’ (1 s period prior to the cue).

Impact of letter change statistics on reach timing

318 We found that the LC statistics influenced manual performance in the ball-drop task,
319 particularly the timing of reach initiation (Fig. 6B). In both fingertip and tweezer trials where
320 participants fixated the ball (i.e., the ball-slot and ball-display-slot gaze patterns), the
321 distribution of reach onset times relative to the antecedent detected LC differed from the random
322 distribution (K-S test, $p < 0.05$ in all four cases). Specifically, reach onsets were biased towards
323 the silent period (top two rows of Fig. 6B), mirroring the bias observed in ball fixations. In
324 contrast, in fingertip trials where the ball was not fixated (i.e., the slot and display-only gaze
325 patterns), the distribution of reach onset times did not differ from the random distribution (K-S
326 test, $p \geq 0.4$; bottom two rows of Fig. 6B). Because of the limited number tweezer trials with
327 the slot and display-only gaze patterns, we did not test whether the distributions of reach onsets
328 in these trials differed from the random distribution. These findings suggest that initiating the
329 reach during the silent period was a strategy to reduce the risk of missing LCs.

330 Importantly, we found that there was a robust temporal coupling between action landmark
331 fixations and contact events, irrespective of the gaze pattern that participant selected. Thus, for
332 all gaze patterns involving a ball fixation, and across a wide range of ball fixation onset times
333 relative to the last detected LC, we observed strong temporal coupling between ball fixation
334 onset and ball contact (top row of Fig. 6C). Similarly, across all gaze patterns involving a slot
335 fixation, we observed a strong temporal coupling between slot fixation onset and slot entry
336 (bottom row of Fig. 6C).

337 The intercept and slope of the relationship between ball fixation onset and ball contact time
338 were 0.43 s and 1.00 in fingertip trials and 0.41 s and 0.97 in tweezer trials. The intercept and
339 slope of the relationship between slot fixation onset and slot entry time were 0.368 s and 1.00
340 in fingertip trials and 0.38 s and 0.99 in tweezer trials. These intercept values are consistent
341 with our earlier observation that both ball and slot fixations typically began approximately 0.4
342 s before the corresponding contact event, irrespective of the end-effector used (Figs. 4E and F).

343 In the ball drop task, participants typically initiated their reach movement towards the ball in
344 response to the auditory cue signaling the active slot for that trial. These ‘reactive reaches’
345 generally began about 0.5 s after the cue. However, in a notable proportion of trials, participants
346 initiated their reach in anticipation of the cue, starting either before or shortly after the cue
347 (within 0.5 s). If participants generated these ‘anticipatory reaches’ in response to a LC
348 occurring shortly before the cue, and fixated the ball when doing so, it could contribute to the
349 greater-than-expected frequency of both reach onsets and ball fixation onsets during the silent
350 period in trials with the ball-slot and ball-display-slot gaze patterns.

351 To examine this further, for fingertip and tweezer trials involving ball fixation, we analyzed the
352 relationship between the timing of reach onset relative to the cue and the timing of the detected
353 LC relative to the cue. For each trial, we identified the detected LC closest in time to the
354 midpoint of the ‘ready period’, defined as the 1-second interval between the ball returning to its
355 starting position and the auditory cue. This LC could occur either before or after the midpoint
356 of the ready period. We observed that most anticipatory reaches—characterized by reach onset
357 times near or preceding the cue—occurred when the detected LC happened within the ready
358 period. This relationship is evident in the left two panels of Fig. 6D, which show reach onset
359 times relative to the cue plotted against the timing of the detected LC. These findings suggest
360 that the decision to initiate an anticipatory reach is closely linked to the detection of a LC during
361 the ready period.

362 The frequency distributions of reach onset times relative to the cue, shown in the right two
363 panels of Fig. 6D for trials with the ball-slot and ball-display-slot patterns, revealed earlier reach
364 onset times when the LC occurred within the ready period compared to when it occurred outside
365 of it. Due to the relatively small number of fingertip trials, fingertip and tweezer trials were
366 combined for these analyses. Importantly, for both gaze patterns, the distributions within and
367 outside the ready period differed (KS test, $p < 0.02$ in both cases). This suggests that the timing
368 of detected LCs had a distinct influence on reach initiation during the ball drop task.

369 Overall, these results demonstrate two strategies that participants used to exploit the statistical
370 properties of LCs to mitigate competition for gaze resources between the tasks. First, they
371 adjusted the timing of their reaching movements so that the ball could be preferentially fixated
372 during the silent period. Second, they adapted their gaze patterns on a trial-by-trial basis to
373 increase the likelihood that gaze could be allocated to the action task with minimal disruption
374 of the LC monitoring task.

Discussion

375 This study investigated three hypotheses about how individuals allocate gaze during object
376 manipulation while also monitoring the environment. These hypotheses concerned: (1) how
377 often, when, and where fixations related to the action task occur when performing the task with
378 end-effectors with different visuomotor demands; (2) whether participants learn and exploit the
379 temporal statistics of the monitoring task to guide gaze allocation; and (3) whether they adjust
380 manual action timing based on these statistics to reduce competition for gaze resources between
381 the tasks. Our findings supported all three hypotheses, as detailed below.

Gaze allocation to the ball-drop task: Effects of end-effector

As expected, when the ball-drop task was performed in isolation using either the fingertips or tweezers, gaze was directed exclusively to the ball and slot, arriving ahead of the end-effector and departing around the time of its arrival or shortly afterward. This finding aligns with prior research on eye-hand coordination in visually guided tasks (Flanagan and Johansson, 2003; Fooken et al., 2021; Hayhoe, 2017; Illamperuma and Fooken, 2024; Johansson et al., 2001; Land et al., 1999). However, when participants performed the ball-drop alongside the monitoring task, we expected gaze to be allocated to the action task when visuomotor control was most critical. Consistent with this expectation, we found that while the ball and slot were almost always fixated in tweezer trials, ball fixations—and, to a lesser extent, slot fixations—were frequently absent in fingertip trials. The increased reliance on central vision with tweezers is likely due to several factors. First, the smaller contact surfaces of the tweezer tips require greater spatial precision, particularly during ball grasping. Second, unlike the more malleable fingertips, the rigid tweezer tips do not conform to the ball's surface, reducing grasp stability. Note that in robotic manipulation, soft contact surfaces are used to enhance stability and reduce spatial precision demands (Bicchi, 2000; Bicchi and Kumar, 2002; Billard and Kragic, 2019). Third, the tweezers provide limited tactile feedback about the contact state, and impaired tactile sensibility is known to increase reliance on vision during object manipulation (Brink and Mackel, 1987; Chemnitz et al., 2013; Jenmalm and Johansson, 1997; Jerosch-Herold, 1993). These findings suggest that analyzing gaze behaviour during object manipulation tasks can offer insights into tactile impairments and the effectiveness of tools in transmitting tactile information to the user, an important consideration for applications such as teleoperation and robot assisted surgery.

Timing and role of action task fixations across end-effectors

Independent of the end-effector employed, the timing of ball and slot fixation onsets and offsets were most closely correlated with ball contact and slot entry, respectively, compared to other events in the ball-drop task. For both action task landmarks (ball and slot), and both end-effectors, gaze typically arrived around 0.4 s prior to contact, a pattern consistent with visuomotor coordination in natural action tasks (Ballard et al., 1995; Keshava et al., 2024; Land and Hayhoe, 2001). However, as hypothesized, the timing of gaze shifts differed based on the end-effector. In fingertip trials, gaze shifted away from the ball and slot well ahead of contact (0.15 s on average), whereas in tweezer trials, gaze shifted just after contact (0.05 s on average). This difference supports our hypothesis that the role of action task fixations varies depending on the end effector. Specifically, when using the fingertips, ball and slot fixations primarily

414 direct the hand (or object in hand) toward the target landmark using peripheral vision and gaze
415 position-related signals (de Brouwer et al., 2018; Prablanc et al., 1986; Wilmot et al., 2006). In
416 contrast, when using tweezers, the delayed gaze shift away from the target suggests that these
417 fixations also assist in *guiding* the end effector with central vision when close to the target. Once
418 the fingertips or tweezer tips contact the ball or slot, tactile feedback becomes available,
419 facilitating a transition from visuomotor control to haptic sensorimotor control. This transition
420 aligns with the shift from motion control to force control, which is thought to involve distinct
421 control processes (Casadio et al., 2015; Chib et al., 2009; Kolesnikov et al., 2011; Piovesan et
422 al., 2019). In the ball drop task, tactile information can not only guide force adjustments, as
423 when grasping the ball, but can also drive kinematic corrections, as when repositioning the ball
424 when inserting it into the slot. In manipulation tasks, tactile information enables rapid
425 adjustments (90-120 ms) to both force (Johansson and Flanagan, 2009) and kinematics
426 (Pruszynski et al., 2018, 2016) via automatic feedback control mechanisms.

Gaze patterns and task performance across end-effectors

427 We found that across both fingertip and tweezer trials, participants used different gaze patterns
428 when performing the ball drop task alongside the monitoring task. In fingertip trials, two main
429 gaze patterns were observed: “display only” and “slot”, distinguished by whether the participant
430 fixated the slot or kept their gaze on the display throughout the trial. In tweezer trials, two main
431 patterns were also observed: “ball-slot” and “ball-display-slot”, distinguished by whether the
432 participant decided to fixate the display between the ball and slot fixations. Importantly, the
433 choice of gaze patterns was linked to task performance. In fingertip trials, the duration of the
434 slot phase was shorter when participants fixated the slot, and in tweezer trials, the duration of
435 transport phase was shorter when participants shifted their gaze directly from the ball to the
436 slot, bypassing the display. These results suggest a trade-off between the action and visual
437 monitoring tasks; allocating gaze to the action task improves performance but increases the risk
438 of missing a letter change (LC). However, as we will discuss next, participants can mitigate this
439 trade-off by leveraging the LC statistics to guide gaze pattern selection and execution of the
440 manipulation task.

Exploiting temporal statistics of the monitoring task for gaze allocation

441 Our findings support the hypothesis that participants exploited the temporal statistics of LCs in
442 the visual monitoring task to guide gaze allocation decisions. In fingertip trials, the decision to
443 fixate the ball was strongly influenced by the LC statistics, with almost all observed ball
444 fixations occurring during the silent period when no LC could occur. Similarly, in tweezer trials,
445 the choice to fixate the display between fixating the ball and slot was also influenced by LC

446 statistics. Specifically, participants often skipped the display when the ball fixation occurred
447 within the silent period, but almost always fixated the display when the ball fixation occurred
448 outside this period. These findings align with previous research demonstrating that human gaze
449 behaviour is sensitive to probabilistic regularities in the environment (Jovancevic-Misic and
450 Hayhoe, 2009). For example, individuals adjust the timing of gaze shifts based on learned
451 temporal statistics of relevant visual events to optimize detection when events occur at different
452 intervals across two spatial locations (Hoppe and Rothkopf, 2016). In addition, during visual
453 search tasks, gaze is allocated strategically based on spatial statistics to optimize exploration
454 (Eckstein, 2017; Hoppe and Rothkopf, 2019; Najemnik and Geisler, 2005; Renninger et al.,
455 2007). Our study adds a novel perspective on how environmental statistics influence gaze
456 behaviour by demonstrating that humans can learn and exploit the temporal patterns of
457 externally determined events in the visual environment while concurrently performing an action
458 task that relies on visual guidance. This result suggests that the processes involved in using
459 visual information in sensorimotor control—including peripheral and central vision and gaze-
460 related signals—are largely independent of those used to extract and use statistical regularities
461 from the visual environment.

Timing of manual actions to mitigate gaze competition

462 Unlike visual monitoring tasks where the timing of relevant events is often externally
463 determined, in manual action tasks, the timing of events that make demands on central vision
464 can generally be controlled or adjusted by the actor. We hypothesized that participants would
465 adjust the timing of their manual actions to reduce competition for gaze resources between the
466 ball drop and visual monitoring tasks. Supporting this hypothesis, we found that participants
467 adjusted their reach onset times so that ball fixations—supporting the reaching movement and
468 ball grasping—occurred during the silent period more often than expected if reach timing was
469 independent of LC occurrences. Importantly, this indicates that participants not only learned the
470 statistical properties of LCs but exploited predictions of when and where action task fixations
471 are required during the unfolding task.

472 We found that in tweezer trials, participants often initiated their reach movement in anticipation
473 of the cue when they observed a LC during the ready period (i.e., the period from when the ball
474 returned to its start position to the auditory cue). This allowed them to both preferentially initiate
475 their reach movement, and fixate the ball, during the silent period. In contrast, in fingertip trials
476 in which the ball was rarely fixated, participants generally did not generate anticipatory reaches,
477 indicating that anticipatory reaches are linked to fixating the ball, and are not simply a
478 consequence of seeing a LC during the ready period. Indeed, in the few fingertip trials in which

479 participants did fixate the ball, they tended to generate anticipatory reaches, allowing them to
480 fix the ball “for free” during the silent period.

481 In fingertip trials, participants almost always maintained fixation at the display while reaching
482 towards and grasping the ball. Importantly, even though they were not fixating the target of the
483 action (i.e., the ball), we know from previous work that they could still use peripheral vision to
484 direct the hand towards the action target; i.e., the ball (de Brouwer et al., 2018; Neggers and
485 Bekkering, 2001, 2000). The question arises as to whether the use of peripheral vision for
486 controlling action is affected by the concurrent use of central vision to monitor external
487 environmental events. Our results suggest that these two visual functions may be independent.
488 If the processing demands associated with visual monitoring are reduced during the silent
489 period, and if visual monitoring interferes with the use of peripheral vision for action, we might
490 have expected participants to preferentially initiate their reaches during the silent period, which
491 they did not.

492 Overall, our results suggest that participants learned LC statistics to strategically adjust the
493 timing of reaching movements to mitigate potential conflicts between the monitoring and object
494 manipulation tasks, thereby supporting both tasks.

Conclusions

495 This study offers novel insights into the control and coordination of eye and hand movements
496 in real-world action tasks. First, our findings support the hypothesis that when gaze competition
497 arises, participants prioritize key control points—specifically, contact events between the hand,
498 or tool in hand, and target object in the environment—when allocating gaze during action tasks.
499 Second, our results demonstrate that participants exploit temporal regularities in the external
500 environment to enhance task performance by adapting both hand and eye movements.

Methods

Participants

501 Eleven right-handed participants (8 male; aged 22 to 33 yr) took part in the study. All
502 participants reported normal or corrected-to-normal vision and were naive to the purpose of the
503 study. The study was approved by the ethics committee of the University of Umeå and
504 participants gave written informed consent before participating in the study.

Apparatus and general procedure

505 Participants were seated at a table with the ball-drop apparatus positioned in front of them (Fig.
506 1A). It consisted of a 15 cm high vertically oriented Perspex tube (inner diameter = 14 mm;
507 wall thickness = 3 mm) mounted on a wooden platform. The tube was positioned about 2.5 cm
508 to the left of the participant's mid-sagittal plane, with its top aligned at eye level. The tube
509 contained three slots, the centers of which were positioned at heights of 5, 8, and 11 cm above
510 the platform surface.

511 Participants repeatedly performed the ball-drop task which required them to reach for and grasp
512 a small ball (12 mm diameter polished brass sphere) from a designated start position on the
513 platform, transport it towards the target slot, insert it into the tube and release it, and then return
514 the hand to a support plate located 20 cm to the right of the platform. A pre-recorded verbal
515 instruction ("bottom", "middle", or "top"), indicating which slot to use, was played 1 s after the
516 ball had rolled back to its start position. The ball's start position was 3 cm to the right of the
517 tube's vertical midline, and all movements occurred within a frontal plane at 40 cm distance
518 from the participants' eyes. The platform surface was slightly lower at the ball start position,
519 ensuring that the ball rolled back to its start position when exiting the tube. In different blocks
520 of trials, the ball was grasped either with the fingertips or with a pair of tweezers held by the
521 right hand. The tweezers, made of plastic, were 14 cm long with cylindrical tips (4 mm
522 diameter) coated for 12 mm with thin plastic tubing to increase the friction against the ball.

523 In the single task conditions, participants performed the ball-drop task only. In the dual task
524 conditions, participants completed a visual monitoring task that involved detecting letter
525 changes (LCs) on an LED text display, which required central vision. The text display was
526 positioned in the upper right quadrant of the scene (Fig. 1 C and D), 7 cm behind the frontal
527 plane in which the tube was located. The visual angle between the center of the letter area and
528 the center of the top slot and ball were 24° and 28°, respectively. The letter height and width
529 corresponded to 0.5° × 0.7° visual angle. At randomly distributed intervals (1.5—6.5 s, uniform
530 distribution), the letter M was changed to W for 300 ms before reverting to M. The participants
531 were instructed to press a button held in their left hand as soon as they detected a LC. If they
532 failed to respond within 1 second after a change, the LC was considered missed. Missed
533 detections were signaled by a brief beep and flashing hash marks on the display for 600 ms. To
534 ensure that the participants relied on central vision to detect LCs rather than peripheral visual
535 cues, the letter M randomly shifted its horizontal position by 0.6° visual angle at intervals
536 between 1 and 3 s (uniform distribution). Pilot tests confirmed that participants had to foveate
537 the display to detect LCs.

538 To encourage engagement with both tasks, participants earned 1 Swedish krona (SEK) for each
539 ball-drop and lost 3 SEK for each missed LC. Participants were informed of these reward and
540 penalty rules before testing.

541 Participants completed four conditions in the following order: single-task with fingertips,
542 single-task with tweezers, dual-task with fingertips, and dual-task with tweezers. Each
543 condition included 30 trials (10 per target slot). The target slot order varied unpredictably across
544 trials.

545 Note that each participant completed three versions of the experiment that differed in terms of
546 how the target slot was cued. In the version reported in this paper, the slot was cued with an
547 auditory cue. In the visual cue version, the slots were indicated by illuminating LEDs located
548 on the tubes, and in the ‘memory’ version, the participants were instructed to follow a specific
549 target slot sequence that they had to remember. The order in which these three versions were
550 completed was counterbalanced across participants. Importantly, we observed no clear
551 differences in manual or gaze behaviour across these experiments and therefore decided to focus
552 on one of them.

Data collection and analysis

553 Gaze position was recorded at 120 samples/s using an infrared video-based eye-tracking system
554 (RK-726PCI pupil/corneal tracking system, ISCAN Inc., Burlington, MA). Head stability was
555 maintained with an adjustable chin support and a forehead support to which the head was
556 secured with Velcro straps. The standard deviations of gaze position measurement errors were
557 0.50° (horizontal) and 0.52° (vertical) of visual angle, corresponding to 0.35 cm and 0.36 cm in
558 the vertical plane—parallel to the participant’s coronal plane—in which the ball was moved
559 during the task.

560 The position of the right index finger and the tweezers tips was tracked at 60 samples/s using
561 miniature electromagnetic position-angle sensors with six degrees of freedom (RX1-D
562 miniature receiver; FASTRAK, Polhemus, Colchester, VT). The fingertip sensor, attached to
563 the nail, recorded the preferred contact point with the ball. The mapping between the position
564 and orientation of the sensor and the preferred contact points was established during calibration
565 trials in which the participants grasped the stationary ball at a known position. The position of
566 the tweezer sensor, attached to the proximal end of the tweezers, was offset to the midpoint
567 between the tweezer tips. When the ball was at its start position, it rested on a rectangular ‘force
568 plate’ that formed part of the platform surface. A six-axis force-torque transducer (Nano F/T
569 transducer, ATI Industrial Automation, Apex, NC; sampling rate 400 Hz), mounted under the

570 plate, detected when the impacted on the platform after being dropped through the tube, when
571 the ball was first contacted by the fingertips or tweezers, and when the ball lifted off the
572 platform. An optical reflex detector (SG-2BC, Kodensi, Japan), mounted in a small hole
573 beneath the ball's start position in the platform, provided a digital signal that indicated when
574 the ball was at its start position. Additional reflex detectors, positioned at the lower edge of each
575 slot, registered when the ball had fallen ~5 mm below the slot's lower edge. To estimate the
576 position of the grasped ball during the transport phase, we used the sensors that recorded the
577 fingertip position and the tips of the tweezers.

578 All data were sampled using the SC/ZOOM software (Department of Medical and Translational
579 Biology, Umeå University). Signals from the various sensors were time synchronized and stored
580 at 200 samples/s using linear interpolation between consecutive samples.

581 Gaze fixations were identified based on previously established criteria (Johansson et al., 2001).
582 Fixations were classified into three categories according to their location, each defined as a
583 centroid with a 2.5 cm radius: the ball at its start position, the target slot, and the text display.
584 Note that fixation data were pooled across all slot positions (bottom, middle, and top). To
585 qualify as a fixation, gaze had to remain within a designated zone for at least 100 ms. Notably,
586 practically all fixations occurred within these zones (~95 %). Consistent with previous findings
587 on object manipulation, participants almost never fixated their hand, or tweezers in hand, during
588 hand movement.

589 To characterize the movement sequence in the ball-drop task we defined seven action phases
590 based on the speed of the fingertip and tweezers (Fig. 1B)). Speed was computed as the vector
591 sum of the first-time derivative of filtered horizontal and vertical position signals (2nd order
592 Butterworth low-pass filter, 10 Hz cut-off frequency). (2nd order Butterworth low-pass filter,
593 10 Hz cut-off frequency). The start of the reach phase began when the end-effector speed
594 exceeded 2 cm/s. The start of the subsequent ball approach phase was defined as the first
595 minimum (or inflection point) in the speed profile of the end-effector following the initial reach
596 movement. The ball grasp phase spanned from the first contact with the ball to when it was
597 lifted off the platform. The following transport phase began at lift off and ended at the notch in
598 the speed profile marking the transition to the slot approach phase. The slot entry phase started
599 when the ball was 1 cm to the right of the position where it was released inside the tube and
600 ended when the ball was dropped, as detected by the optical sensor in the tube. Finally, the slot
601 entry phase was followed by the return phase and the offset of the return phase were defined at
602 the times at which the end-effector speed exceeded 2 cm/s.

603 We used paired t-tests with Holm-Bonferroni correction, Kolmogorov-Smirnov (KS) tests, and
604 multiple linear regression analyses in our statistical analyses as detailed in the Results. We
605 defined a statistically significant outcome if $p < 0.05$. All statistical analyses were conducted in
606 R (R Core Team, 2022; www.r-project.org). Trials in which participants dropped the ball after
607 initiating the transport phase were excluded. A total of 64 trials (4.7 %) were removed, with no
608 participant losing more than 7 trials.

Citation diversity statement

609 Recent work in several fields of science has identified a bias in citation practices such that
610 papers from women and other minority scholars are under-cited relative to the number of such
611 papers in the field (Bertolero et al., 2020; Caplar et al., 2017; Chatterjee & Werner, 2021; Dion
612 et al., 2018; Dworkin et al., 2020; Fulvio et al., 2021; Maliniak et al., 2013; Mitchell et al.,
613 2013; Wang et al., 2021). Here we sought to proactively consider choosing references that
614 reflect the diversity of the field in thought, form of contribution, gender, race, ethnicity, and
615 other factors. First, we obtained the predicted gender of the first and last author of each reference
616 by using databases that store the probability of a first name being carried by a woman (Dworkin
617 et al., 2020; Zhou et al., 2020). By this measure (and excluding self-citations to the first and last
618 authors of our current paper), our references contain 22.73% woman(first)/woman(last),
619 11.36% man/woman, 16.71% woman/man, and 49.19% man/man. This method is limited in
620 that a) names, pronouns, and social media profiles used to construct the databases may not, in
621 every case, be indicative of gender identity and b) it cannot account for intersex, non-binary, or
622 transgender people. Second, we obtained the predicted racial/ethnic category of the first and
623 last author of each reference by databases that store the probability of a first and last name being
624 carried by an author of colour (Ambekar et al., 2009; Chintalapati et al., 2023). By this measure
625 (and excluding self-citations), our references contain 5.3% author of colour (first)/author of
626 colour(last), 12.04% white author/author of colour, 16.71% author of colour/white author, and
627 65.95% white author/white author. This method is limited in that a) names and Florida Voter
628 Data to make the predictions may not be indicative of racial/ethnic identity, and b) it cannot
629 account for Indigenous and mixed-race authors, or those who may face differential biases due
630 to the ambiguous racialization or ethnicization of their names. We look forward to future work
631 that could help us to better understand how to support equitable practices in science.

References

- 632 Ambekar A, Ward C, Mohammed J, Male S, Skiena S. 2009. Name-ethnicity classification
633 from open sourcesProceedings of the 15th ACM SIGKDD International Conference on
634 Knowledge Discovery and Data Mining. Presented at the KDD09: The 15th ACM
635 SIGKDD International Conference on Knowledge Discovery and Data Mining. Paris
636 France: ACM. pp. 49–58. doi:10.1145/1557019.1557032
- 637 Ballard DH, Hayhoe MM, Li F, Whitehead SD. 1992. Hand-eye coordination during
638 sequential tasks. *Philos Trans R Soc Lond B Biol Sci* **337**:331–339.
639 doi:10.1098/rstb.1992.0111
- 640 Ballard DH, Hayhoe MM, Pelz JB. 1995. Memory Representations in Natural Tasks. *J Cogn*
641 *Neurosci* **7**:66–80. doi:10.1162/jocn.1995.7.1.66
- 642 Bertolero MA, Dworkin JD, David SU, Lloreda CL, Srivastava P, Stiso J, Zhou D, Dzirasa K,
643 Fair DA, Kaczkurkin AN, Marlin BJ, Shohamy D, Uddin LQ, Zurn P, Bassett DS. 2020.
644 Racial and ethnic imbalance in neuroscience reference lists and intersections with
645 gender. doi:10.1101/2020.10.12.336230
- 646 Bicchi A. 2000. Hands for dexterous manipulation and robust grasping: a difficult road toward
647 simplicity. *IEEE Trans Robot Autom* **16**:652–662. doi:10.1109/70.897777
- 648 Bicchi A, Kumar V. 2000. Robotic grasping and contact: a reviewProceedings 2000 ICRA.
649 Millennium Conference. IEEE International Conference on Robotics and Automation.
650 Symposia Proceedings (Cat. No.00CH37065). Presented at the 2000 ICRA. IEEE.
651 International Conference on Robotics and Automation. San Francisco, CA, USA: IEEE.
652 pp. 348–353. doi:10.1109/ROBOT.2000.844081
- 653 Billard A, Kragic D. 2019. Trends and challenges in robot manipulation. *Science*
654 **364**:eaat8414. doi:10.1126/science.aat8414
- 655 Brink EE, Mackel R. 1987. Sensorimotor performance of the hand during peripheral nerve
656 regeneration. *J Neurol Sci* **77**:249–266. doi:10.1016/0022-510X(87)90127-4
- 657 Caplar N, Tacchella S, Birrer S. 2017. Quantitative evaluation of gender bias in astronomical
658 publications from citation counts. *Nat Astron* **1**:0141. doi:10.1038/s41550-017-0141
- 659 Casadio M, Pressman A, Mussa-Ivaldi FA. 2015. Learning to push and learning to move: the
660 adaptive control of contact forces. *Front Comput Neurosci* **9**.
661 doi:10.3389/fncom.2015.00118
- 662 Chatterjee P, Werner RM. 2021. Gender Disparity in Citations in High-Impact Journal
663 Articles. *JAMA Netw Open* **4**:e2114509. doi:10.1001/jamanetworkopen.2021.14509
- 664 Chemnitz A, Dahlin LB, Carlsson IK. 2013. Consequences and adaptation in daily life –
665 patients' experiences three decades after a nerve injury sustained in adolescence. *BMC*
666 *Musculoskelet Disord* **14**:252. doi:10.1186/1471-2474-14-252
- 667 Chib VS, Krutky MA, Lynch KM, Mussa-Ivaldi FA. 2009. The Separate Neural Control of
668 Hand Movements and Contact Forces. *J Neurosci* **29**:3939–3947.
669 doi:10.1523/JNEUROSCI.5856-08.2009

- 670 Chintalapati R, Laohaprapanon S, Sood G. 2023. Predicting Race and Ethnicity From the
671 Sequence of Characters in a Name.
- 672 de Brouwer AJ, Gallivan JP, Flanagan JR. 2018. Visuomotor feedback gains are modulated by
673 gaze position. *J Neurophysiol* **120**:2522–2531. doi:10.1152/jn.00182.2018
- 674 Dion ML, Sumner JL, Mitchell SM. 2018. Gendered Citation Patterns across Political Science
675 and Social Science Methodology Fields. *Polit Anal* **26**:312–327.
676 doi:10.1017/pan.2018.12
- 677 Dworkin JD, Linn KA, Teich EG, Zurn P, Shinohara RT, Bassett DS. 2020. The extent and
678 drivers of gender imbalance in neuroscience reference lists. *Nat Neurosci* **23**:918–926.
679 doi:10.1038/s41593-020-0658-y
- 680 Eckstein MP. 2017. Probabilistic Computations for Attention, Eye Movements, and Search.
681 *Annu Rev Vis Sci* **3**:319–342. doi:10.1146/annurev-vision-102016-061220
- 682 Flanagan JR, Bowman MC, Johansson RS. 2006. Control strategies in object manipulation
683 tasks. *Curr Opin Neurobiol* **16**:650–659. doi:10.1016/j.conb.2006.10.005
- 684 Flanagan JR, Johansson RS. 2003. Action plans used in action observation. *Nature* **424**:769–
685 771. doi:10.1038/nature01861
- 686 Fooken J, Baltaretu BR, Barany DA, Diaz G, Semrau JA, Singh T, Crawford JD. 2023.
687 Perceptual-Cognitive Integration for Goal-Directed Action in Naturalistic Environments.
688 *J Neurosci* **43**:7511–7522. doi:10.1523/JNEUROSCI.1373-23.2023
- 689 Fooken J, Kreyenmeier P, Sperling M. 2021. The role of eye movements in manual
690 interception: A mini-review. *Vision Res* **183**:81–90. doi:10.1016/j.visres.2021.02.007
- 691 Fulvio JM, Akinnola I, Postle BR. 2021. Gender (Im)balance in Citation Practices in
692 Cognitive Neuroscience. *J Cogn Neurosci* **33**:3–7. doi:10.1162/jocn_a_01643
- 693 Goodale MA, Pelisson D, Prablanc C. 1986. Large adjustments in visually guided reaching do
694 not depend on vision of the hand or perception of target displacement. *Nature* **320**:748–
695 750. doi:10.1038/320748a0
- 696 Hayhoe MM. 2017. Vision and Action. *Annu Rev Vis Sci* **3**:389–413. doi:10.1146/annurev-
697 vision-102016-061437
- 698 Hoppe D, Rothkopf CA. 2019. Multi-step planning of eye movements in visual search. *Sci
699 Rep* **9**:144. doi:10.1038/s41598-018-37536-0
- 700 Hoppe D, Rothkopf CA. 2016. Learning rational temporal eye movement strategies. *Proc Natl
701 Acad Sci* **113**:8332–8337. doi:10.1073/pnas.1601305113
- 702 Illamperuma NH, Fooken J. 2024. Towards a functional understanding of gaze in goal-
703 directed action. *J Neurophysiol* **132**:767–769. doi:10.1152/jn.00342.2024
- 704 Jenmalm P, Johansson RS. 1997. Visual and Somatosensory Information about Object Shape
705 Control Manipulative Fingertip Forces. *J Neurosci* **17**:4486–4499.
706 doi:10.1523/JNEUROSCI.17-11-04486.1997

- 707 Jerosch-Herold C. 1993. Measuring Outcome in Median Nerve Injuries. *J Hand Surg* **18**:624–
708 628. doi:10.1016/0266-7681(93)90019-C
- 709 Johansson RS, Flanagan JR. 2009. Coding and use of tactile signals from the fingertips in
710 object manipulation tasks. *Nat Rev Neurosci* **10**:345–359. doi:10.1038/nrn2621
- 711 Johansson RS, Westling G, Bäckström A, Flanagan JR. 2001. Eye–Hand Coordination in
712 Object Manipulation. *J Neurosci* **21**:6917–6932. doi:10.1523/JNEUROSCI.21-17-
713 06917.2001
- 714 Jovancevic-Misic J, Hayhoe MM. 2009. Adaptive Gaze Control in Natural Environments. *J*
715 *Neurosci* **29**:6234–6238. doi:10.1523/JNEUROSCI.5570-08.2009
- 716 Keshava A, Nezami FN, Neumann H, Izdebski K, Schüler T, König P. 2024. Just-in-time:
717 Gaze guidance in natural behavior. *PLOS Comput Biol* **20**:e1012529.
718 doi:10.1371/journal.pcbi.1012529
- 719 Kolesnikov M, Piovesan D, Lynch KM, Mussa-Ivaldi FA. 2011. On force regulation strategies
720 in predictable environments2011 Annual International Conference of the IEEE
721 Engineering in Medicine and Biology Society. Presented at the 2011 33rd Annual
722 International Conference of the IEEE Engineering in Medicine and Biology Society.
723 Boston, MA: IEEE. pp. 4076–4081. doi:10.1109/IEMBS.2011.6091013
- 724 Kowler E. 2011. Eye movements: The past 25years. *Vision Res* **51**:1457–1483.
725 doi:10.1016/j.visres.2010.12.014
- 726 Land M, Mennie N, Rusted J. 1999. The Roles of Vision and Eye Movements in the Control
727 of Activities of Daily Living. *Perception* **28**:1311–1328. doi:10.1068/p2935
- 728 Land MF. 2006. Eye movements and the control of actions in everyday life. *Prog Retin Eye*
729 *Res* **25**:296–324. doi:10.1016/j.preteyeres.2006.01.002
- 730 Land MF, Furneaux S. 1997. The knowledge base of the oculomotor system. *Philos Trans R*
731 *Soc Lond B Biol Sci* **352**:1231–1239. doi:10.1098/rstb.1997.0105
- 732 Land MF, Hayhoe M. 2001. In what ways do eye movements contribute to everyday
733 activities? *Vision Res* **41**:3559–3565. doi:10.1016/S0042-6989(01)00102-X
- 734 Maliniak D, Powers R, Walter BF. 2013. The Gender Citation Gap in International Relations.
735 *Int Organ* **67**:889–922. doi:10.1017/S0020818313000209
- 736 Mitchell SM, Lange S, Brus H. 2013. Gendered Citation Patterns in International Relations
737 *Journals. Int Stud Perspect* **14**:485–492. doi:10.1111/insp.12026
- 738 Najemnik J, Geisler WS. 2005. Optimal eye movement strategies in visual search. *Nature*
739 **434**:387–391. doi:<https://doi.org/10.1038/nature03390>
- 740 Neggers SFW, Bekkering H. 2001. Gaze Anchoring to a Pointing Target Is Present During the
741 Entire Pointing Movement and Is Driven by a Non-Visual Signal. *J Neurophysiol*
742 **86**:961–970. doi:10.1152/jn.2001.86.2.961
- 743 Neggers SFW, Bekkering H. 2000. Ocular Gaze is Anchored to the Target of an Ongoing
744 Pointing Movement. *J Neurophysiol* **83**:639–651. doi:10.1152/jn.2000.83.2.639

- 745 Piovesan D, Kolesnikov M, Lynch K, Mussa-Ivaldi FA. 2019. The Concurrent Control of
746 Motion and Contact Force in the Presence of Predictable Disturbances. *J Mech Robot*
747 **11**:060903. doi:10.1115/1.4044599
- 748 Prablanc C, Pélinson D, Goodale MA. 1986. Visual control of reaching movements without
749 vision of the limb: I. Role of retinal feedback of target position in guiding the hand. *Exp*
750 *Brain Res* **62**. doi:10.1007/BF00238848
- 751 Pruszynski JA, Flanagan JR, Johansson RS. 2018. Fast and accurate edge orientation
752 processing during object manipulation. *eLife* **7**:e31200. doi:10.7554/eLife.31200
- 753 Pruszynski JA, Johansson RS, Flanagan JR. 2016. A Rapid Tactile-Motor Reflex
754 Automatically Guides Reaching toward Handheld Objects. *Curr Biol* **26**:788–792.
755 doi:10.1016/j.cub.2016.01.027
- 756 Renninger LW, Vergheze P, Coughlan J. 2007. Where to look next? Eye movements reduce
757 local uncertainty. *J Vis* **7**:6. doi:10.1167/7.3.6
- 758 Säfström D, Johansson RS, Flanagan JR. 2014. Gaze behavior when learning to link
759 sequential action phases in a manual task. *J Vis* **14**:3–3. doi:10.1167/14.4.3
- 760 Saunders JA, Knill DC. 2004. Visual Feedback Control of Hand Movements. *J Neurosci*
761 **24**:3223–3234. doi:10.1523/JNEUROSCI.4319-03.2004
- 762 Saunders JA, Knill DC. 2003. Humans use continuous visual feedback from the hand to
763 control fast reaching movements. *Exp Brain Res* **152**:341–352. doi:10.1007/s00221-003-
764 1525-2
- 765 Wang X, Dworkin JD, Zhou D, Stiso J, Falk EB, Bassett DS, Zurn P, Lydon-Staley DM. 2021.
766 Gendered citation practices in the field of communication. *Ann Int Commun Assoc*
767 **45**:134–153. doi:10.1080/23808985.2021.1960180
- 768 Wilmut K, Wann JP, Brown JH. 2006. How active gaze informs the hand in sequential
769 pointing movements. *Exp Brain Res* **175**:654–666. doi:10.1007/s00221-006-0580-x
- 770 Yarbus AL. 1967. Eye Movements and Vision. Boston, MA: Springer US. doi:10.1007/978-1-
771 4899-5379-7
- 772 Zhou D, Bertolero MA, Stiso J, Cornblath EJ, Teich EG, Blevins AS, Virtualmario Camp,
773 Dworkin J, Bassett DS. 2020. Gender diversity statement and code notebook.