Gaze behavior when learning to link sequential action phases in a manual task

Daniel Säfström
Department of Integrative Medical Biology, Physiology Section, Umeå University, Umeå, Sweden

Roland S. Johansson
Department of Integrative Medical Biology, Physiology Section, Umeå University, Umeå, Sweden

J. Randall Flanagan
Department of Psychology and Centre for Neuroscience Studies, Queen’s University, Kingston, Ontario, Canada

Most manual tasks comprise a sequence of action phases. Skill acquisition in such tasks involves a transition from reactive control, whereby motor commands for the next phase are triggered by sensory events signaling completion of the current phase, to predictive control, whereby commands for the next phase are launched in anticipation of these events. Here we investigated gaze behavior associated with such learning. Participants moved a cursor to successively acquire visual targets, as quickly as possible, by actively keeping the cursor within the target zone (hold phase) for a required duration, before moving to the next target (transport phase). Distinct visual and auditory events marked completion of each phase and, with learning, the launching of the transport phase shifted from being reactively to predictively controlled. Initially, gaze was directed to the current target throughout the hold phase, allowing visual feedback control of the cursor position, and shifted to the next target in synchrony with the cursor. However, with learning, two distinct gaze behaviors emerged. Gaze either shifted to the next target well before the end of the hold phase, facilitating planning of the forthcoming cursor movement, or shifted to the next target after the cursor, enabling cursor exits to be monitored in central vision. These results suggest that, with learning, gaze behavior changes to support evolving task demands, and that people distribute different gaze behaviors across repetitions of the task.

Introduction

Most manual tasks involve a sequence of action phases that are implemented to achieve subgoals of the overall task. The completion of these subgoals is typically demarcated by mechanical events that give rise to discrete sensory signals, often in multiple modalities (Johansson & Flanagan, 2009). For example, in taking a drink from a glass of water, contact between the fingertips and the glass marks the completion of the reaching phase, the breaking of contact between the glass and tabletop marks the completion of the subsequent lift phase, and contact between the lips and glass marks the completion of the object transport phase. Successful task performance often requires that each action phase be successfully completed before the next phase is initiated. For example, if one attempts to lift an object before it has been securely grasped, the object may slip or be knocked over. By comparing predicted and actual sensory events associated with subgoal completion, the brain can monitor task progression, launch appropriate corrective actions when required, and update knowledge related to the task (Flanagan, Bowman, & Johansson, 2006). The prediction of sensory events associated with subgoal completion also supports predictive linking of action phases, whereby motor commands for the forthcoming phase are launched in anticipation of completion of the current phase (Bowman, Johansson, & Flanagan, 2009; Säfström & Edin, 2008). Given time delays inherent in sensorimotor feedback loops, such predictive linking enables smooth phase transitions and can substantially decrease task completion times. Learning to predictively link action phases in novel tasks is thus an important component of skill acquisition.

We have recently described changes in manual performance when learning a novel task involving a sequence of action phases (Säfström, Flanagan, & Johansson, 2013). Here we focus on the gaze behavior
recorded during the experiments that contributed to that previous report. Participants controlled the position of a cursor on a screen by applying forces to a knob instrumented with a force sensor. The cursor position relative to the screen center scaled with force (1 N: 4.8° of visual angle). Application of leftward and rightward force on the knob moved the cursor upwards and downwards on the screen, respectively, and forces directed away from and towards the body moved it rightwards and leftwards, respectively. An infrared video-based eye tracker recorded the gaze position of the right eye. (B) Examples of target transitions. There were always two targets visible on the screen, the current (highlighted) and next targets. The previous target disappeared when the cursor exited the target zone after goal attainment, which required the participants to keep the cursor in the target zone for 0.6 s (in standard trials) before aiming for the next target. At the time of goal attainment, participants received visual (target doubled in diameter for 50 ms in flashlike manner) and auditory (brief tone) feedback. Cursor and gaze positions are represented by continuous and dotted lines, respectively. (C) Gaze and cursor positions, as a function of time, corresponding to the target transitions shown in (B).

During initial performance, the transport phase was triggered, reactively, based on sensory events signaling completion of the hold phase. However, with practice, participants learned to launch the motor commands for the transport phase based on a prediction of these sensory events such that the cursor exited the target zone shortly after the end of the required hold period (Säfström et al., 2013). Moreover, given the temporal uncertainty in predictively launching the transport phase, participants balanced the trade-off between the time costs incurred by moving too early or too late in a target to acquire it, which was detrimental to task performance because of the time cost.

Figure 1. Experimental setup. (A) Participants controlled a cursor on a screen by applying forces to a knob instrumented with a force sensor. The cursor position relative to the screen center scaled with force (1 N: 4.8° of visual angle). Application of leftward and rightward force on the knob moved the cursor upwards and downwards on the screen, respectively, and forces directed away from and towards the body moved it rightwards and leftwards, respectively. An infrared video-based eye tracker recorded the gaze position of the right eye. (B) Examples of target transitions. There were always two targets visible on the screen, the current (highlighted) and next targets. The previous target disappeared when the cursor exited the target zone after goal attainment, which required the participants to keep the cursor in the target zone for 0.6 s (in standard trials) before aiming for the next target. At the time of goal attainment, participants received visual (target doubled in diameter for 50 ms in flashlike manner) and auditory (brief tone) feedback. Cursor and gaze positions are represented by continuous and dotted lines, respectively. (C) Gaze and cursor positions, as a function of time, corresponding to the target transitions shown in (B).
near optimal way (Säfström et al., 2013). The aim of this article is to investigate how gaze behavior supports both learning to predictively link successive action phases and skilled linking of action phases after learning.

We hypothesized that the gaze behavior would change with practice and would reflect changes in the demand for different kinds of gaze-related information with learning. Specifically, we expected that during initial performance, when the transport phase is reactively triggered by sensory events signaling completion of the hold phase and control of cursor position is still being mastered, gaze would remain at the target throughout the hold phase and shift to the next target in synchrony with the cursor. This gaze behavior would help participants keep the cursor. However, we hypothesized that some degree of independence would arise between gaze and manual control once participants learned the temporal characteristics of the task and mastery of cursor control. Specifically, we entertained two possible outcomes regarding what gaze behavior would be observed. On the one hand, gaze might shift to the next target earlier, and prior to completion of the hold phase, in order to facilitate planning of the cursor movement towards the next target, akin to gaze behavior in self-paced natural tasks (for a review, see Land, 2006) and high-speed stacking tasks (Foerster, Carbone, Koesling, & Schneider, 2011). On the other hand, gaze might remain at the target until after the cursor exited. This would allow the time of the cursor exit relative to the end of the required hold period to be monitored in central vision and could help participants maintain optimal task performance (Johansson, Westling, Bäckström, & Flanagan, 2001). It has recently been shown that gaze resources can be adaptively distributed among competing perceptual and motor components of a task in order to optimize task performance—a result that also demonstrates that eye and hand movements can be uncoupled and are not slavishly yoked (Sims, Jacobs, & Knill, 2011). One possibility is that both of the gaze behaviors we considered would be observed in different trials, consistent with the idea that gaze resources can be adaptively allocated across repetitions of the task in order to collect task-relevant visual information.

Methods

Participants

Nine healthy, right-handed participants (five men and four women, aged 21–41 years) with normal vision participated in the study after providing written informed consent prior to the experiment. They received 100 Swedish kronor per hour, with the payment doubled for the top performer. The local Ethics Committee of Umeå University approved the experiment.

Procedure and apparatus

The participants controlled the position of the cursor (filled circle of diameter 3 mm or 0.5° of visual angle) on a vertical computer screen positioned 37 cm in front of the eyes by using the right hand to apply horizontal forces to a spherical knob (Figure 1A) that was instrumented with a force sensor (FT-Nano 17, Assurance Technologies, Garner, NC). The position of the cursor relative to the center of the screen scaled linearly with force (1 N of force to 4.8° of visual angle). With zero force, the cursor was located in the middle of the screen. The mapping from hand force to cursor position (from the origin) was rotated 90° clockwise such that a leftward force moved the cursor upwards and a force away from the body moved it rightwards. The force signals that controlled the cursor were sampled at 200 Hz and low-pass filtered at 3.3 Hz to prevent cursor wobble driven by physiological tremor.

The task involved moving the cursor to attain sequentially presented targets (open circles of diameter 24.3 mm or 3.9° of visual angle) as quickly as possible (Figure 1B). Forty-four possible target locations were distributed equally across the screen under the constraint that the distance between two successive targets was 18° of visual angle and the direction from the previous target was uniformly distributed in the range between 0° and 360°.

At any given time, the current target (highlighted with a border of 0.9 mm) was displayed on the screen along with the next target (border of 0.3 mm). To attain the current target, participants had to move the cursor into the target zone and actively keep it there for a required time. After the current target had been attained, the next target became the current target, a new next target was displayed, and the previous target disappeared (Figure 1B). If the cursor exited the current target zone before the required hold phase duration, the participant had to return the cursor to the target zone and, again, attempt to keep it there for the required time.

Participants received visual and auditory feedback about completion of both the transport phase and the hold phase (Figure 1C). When the cursor entered the current target, marking completion of the transport phase (and the start of the hold phase), the thickness of the current target border increased from 0.9 mm to 1.52 mm. If the cursor remained within the target zone for the required hold duration, the target doubled in diameter for 50 ms before disappearing. The comple-
tion of the transport phase was also indicated by a click sound, and target attainment was indicated by a beep (50 ms). If the cursor exited the target zone prior to the required hold phase duration, feedback about target attainment was not provided and the current target remained on the screen.

A single trial was defined as a hold phase followed by a transport phase. The hold phase started when the cursor entered the target zone and ended when the cursor exited the target zone after the target was attained (i.e., after the required hold period). The transport phase started at the end of the previous hold phase and ended at the start of the next hold phase. The cursor was considered to have entered and exited a target zone when the center of the cursor moved across the center of the border outlining the target zone. Both visual and auditory feedback about goal completion of each phase was provided as already detailed.

An infrared video-based eye tracker (RK-726PCI pupil/corneal tracking system, ISCAN, Burlington, MA) recorded the gaze position of the right eye in the plane of the screen at 120 Hz. The eye tracker was mounted on a wooden support, and the participant’s head was stabilized by a forehead-and-chin support. Calibration was performed in a two-step procedure as previously described elsewhere (Johansson et al., 2001). For initial calibration, the participant was asked to sequentially look at five 3-mm LEDs that were mounted on a flat surface aligned with the work plane, with one LED in the center and one in each corner of the work plane. For final calibration, which was made repeatedly after every third block of trials, we used the same five points as in initial calibration as well as four additional points located at the midpoints of the four lines that defined the rectangular work plane. All experimental data were calibrated off-line using data obtained from the nearest calibration measurements. The system measured gaze in the horizontal and vertical directions with accuracies of 0.50° and 0.52° of visual angle, respectively. The eye tracker prevented vision of the participant’s hands. We considered that the gaze fixated a target when gaze position was within 3° of the center position of the target zone. We chose 3° because this is considered to be the size of the functional fovea in manipulation tasks (Johansson et al., 2001; Terao, Andersson, Flanagan, & Johansson, 2002).

**Standard trials**

Participants first performed the task with a constant required hold phase duration of 0.6 s. The task was performed over successive 72-s periods, with 18-s rests in between, until at least 1,800 targets were attained. The participants were free to take a longer break after each chunk of about 20 min of practice. Participants were told that the goal was to attain as many targets as possible, and during each rest period, the screen displayed the number of targets attained during the previous period as well as the participant’s “high score” from all previous periods.

**Reactive trials**

After completing the standard trials, the participants performed three 72-s periods in which the required hold phase duration varied randomly between 0.2, 0.4, 0.6, 0.8, and 1.0 s following a uniform distribution. Again, the goal was to attain as many targets as possible, and the same feedback about performance was provided. Because the required hold phase duration could not be predicted, participants reactively launched the transport phase based on the feedback signaling completion of the required hold phase duration. In total, individual participants performed between 149 and 163 reactive trials, including between 29 and 31 trials with a hold phase duration of 0.6 s. The latter were used to compare with standard trials, as they had the same hold phase duration. The whole experiment was performed on the same occasion and lasted in total about 2 hr (including breaks).

**Catch trials**

Three different kinds of catch trials were intermixed with the standard trials, each of which could involve one modality alone (auditory or visual feedback) or both modalities simultaneously (auditory and visual feedback): delayed feedback about goal attainment of the hold phase; time-advanced feedback about goal attainment of the hold phase; and delayed feedback about completion of the transport phase. The delayed feedback about goal attainment of the hold phase occurred 0.8, 0.85, or 0.9 s after the onset of the hold phase, corresponding to delays of 0.2, 0.25, and 0.3 s, respectively. If the cursor exited the target zone after the prescribed time (0.6 s) but before delayed feedback was provided, feedback did not appear and the participant was free to move to the next target. The time-advanced feedback about goal attainment occurred 0.2, 0.25, or 0.3 s after the start of the hold phase, corresponding to advances of 0.4, 0.35, and 0.3 s, respectively. Finally, the delayed feedback about transport phase completion occurred 0.2, 0.25, or 0.3 s after the cursor entered the target zone. For each of these kinds of catch trial, either auditory only, visual only, or both auditory and visual feedback could be manipulated (i.e., delayed or advanced). Thus, there were 27 distinct catch trial types (3 kinds × 3 modality
variants \times 3\) times). Catch trials were intermingled with the standard trials in a manner unpredictable for the participants. In addition, the sequence of catch-trial types was randomized. About 5% of all trials were catch trials.

Catch trials were included to study multimodal integration at sensorimotor control points during the task. In this article, we focus on catch trials involving delayed feedback about goal attainment of the hold phase (i.e., hold phase completion), as these trials provide useful information about the control mechanisms underlying gaze shifts from one target to the next. Moreover, we only examine catch trials in which either the auditory cue alone or both the auditory and visual cues were delayed, since both of these had similar and clear effects on gaze and cursor behavior (whereas delay of the visual cues only did not).

**Statistics**

In data analysis, we used an alpha level of 0.05. Bonferroni corrections were used to compensate for multiple comparisons. To facilitate data analysis, we partitioned the trials into seven learning stages. The stages represented approximately equal intervals on a logarithmic scale and were chosen because they could capture the changes in performance with learning (trials 1–20, 21–60, 61–140, 141–300, 301–620, 621–1,260, and >1,260; shaded gray areas in Figure 2A through C).

Only trials from Stage 7 (trial number > 1,260) were included in the analysis of gaze behavior during steady-state performance, because all significant changes in performance occurred prior to this stage.

**Results**

We have described changes in manual performance associated with learning the task in a previous report (Säfström et al., 2013). We will first provide a summary of these results and then describe the gaze behavior, and coordination of gaze and hand movements, associated with different stages of learning.

**Changes in manual performance during learning**

In the standard condition, the trial rate (number of completed targets per second) increased with practice in a manner that could be described by an exponential equation (Figure 2A). Averaged across participants, the half-life for the increase was 206 ± 15 trials (mean ± SE). Two factors contributed to the improvement in performance. First, there was a relatively rapid decrease in the transport phase duration (time between targets) as participants mastered control of the cursor (Figure 2B; half-life: 90 ± 14 trials [mean ± SE] averaged across the participants). Second, there was a much more gradual decrease in the cursor exit time relative to the start of the target hold phase (Figure 2C;
Changes in gaze behavior during learning

Figure 3A shows the distribution of gaze and cursor exit times from the target zone, relative to the onset of the hold phase, for each of the seven learning stages. For comparison, Figure 3B shows the corresponding distributions for reactive trials from the reactive control condition with a hold phase duration of 0.6 s. At the beginning of the practice period, participants typically fixated the current target throughout the hold phase, and gazed the current target, en route to the forthcoming target, at around the same time that the cursor exited the target zone heading for the forthcoming target (Figure 3A). The time distributions of gaze exits from the target zone during the first three stages (trial number \( \leq 140 \)) were unimodal, and the overall timing of these gaze exits (mean of participant medians = 0.94 s; range of participant medians = 0.87–1.08 s) suggests that the gaze shifts, like the cursor exits (mean of participant medians = 0.85 s; range = 0.73–0.92 s), were reactively triggered by sensory feedback about goal attainment related to the hold phase. This is further indicated by the similarity in the distributions of these gaze and cursor exits with those observed in reactive trials from the reactive control condition, in which participants could not predict the timing of the events signaling completion of the required hold period (Figure 3B; mean gaze exit time of participant medians = 0.80 s, range = 0.43–0.92 s; mean cursor exit time of participant medians = 0.87 s, range = 0.84–0.92 s).

During the later stages of the practice period, concomitant with the gradual decrease in the cursor exit times, gaze shifts to the forthcoming target gradually occurred earlier (Figure 3A). However, in contrast to the cursor exit times, the distribution of gaze exit times became more bimodal (see Stages 5, 6, and 7). One mode represented gaze exits that occurred in the time range of the reactively triggered cursor and gaze exits seen in reactive trials from the reactive control condition (Figure 3B). The other mode represented exits that occurred, on average, well before the cursor exits. Thus, participants learned to allocate gaze resources to the next target while still dealing with the current target. The relative frequency of these two modes of gaze behavior changed over the last few stages, and during steady-state performance (Stage 7), gaze most often shifted towards the next target well before goal attainment of the hold phase. However, later gaze shifts were still observed in a substantial proportion of trials.

Figure 3C shows, for each participant, gaze exit time relative to the onset of the hold phase as a function of trial number for all standard trials. The figure illustrates that, on average, gaze exit times decreased for all participants and that the relative frequency of early and late gaze exits, especially during later learning stages, varied across participants. Whereas some of the participants primarily exhibited early gaze shifts (P1, P2, P4, P5, and P8), with gaze shifting to the next target before the end of the required hold period (0.6 s), others exhibited both early and late gaze shifts (P3, P6, and P9), and one primarily exhibited late gaze shifts (P7). Note that in the four participants (P3, P6, P7, and P9) who exhibited substantial numbers of late gaze shifts during steady-state performance (i.e., Stage 7), the frequency of these shifts appears to be quite constant over the approximately 650 trials of Stage 7. For these four participants, the percentage of late gaze exits in the first half of Stage 7 (53%, 41%, 66%, and 45%) was similar to the percentage in the second half of Stage 7 (45%, 43%, 60%, and 34%). However, it is possible that with even more extended practice, the frequency of late gaze shifts would have gradually decreased in some or all of these participants such that mainly early gaze shifts would be observed.

Gaze behavior during steady-state performance

Figure 4A shows, for each participant, the distribution of gaze exit times (solid black histograms) and cursor exit times (gray curves) with reference to the start of the hold phase during steady-state performance (Stage 7). Although the general pattern indicates a bimodal distribution of gaze exit times, the presence and relative size of the two modes varies across participants. Tests for unimodality (Hartigan & Hartigan, 1985) revealed that the distributions for three of the participants were significantly different from unimodal (P3, P6, and P9). That is, these participants...
showed bimodal distributions. Of the other participants, five generated gaze exits that primarily occurred well before the launching of the transport phase (P1, P2, P4, P5, and P8), and one generated gaze exits that mainly occurred in the reactive time range (P7), although this participant still produced a substantial number of early gaze shifts.

Although four of the participants (P3, P6, P7, and P9) exhibited frequent late gaze shifts during steady-state performance, there were no obvious differences between these participants and the remaining participants in terms of learning to predictively control the cursor (and hence learn the duration of the hold phase). First, as shown in Figure 4A (gray curves), all
participants learned to generate predictive cursor exits by the time they reached the steady-state stage. Second, these two groups of participants exhibited similar learning rates, as revealed by exponential fits to the cursor exits for each participant. Specifically, the average half-life was 332 trials for the group of participants who seldom generated late gaze exits during steady-state performance (P1, P2, P4, P5, and P8) and 293 trials for the group of participants who frequently generated late gaze exits.

For all participants, there was a positive correlation between the times of gaze and cursor exits (Figure 4B; \( p < 0.0001 \) for each participant; \( R \) values ranged between 0.32 and 0.53, mean \( R = 0.39 \)). That is, when gaze left the current target early to fixate the next target, the cursor tended to leave the current target relatively early. However, linear regression revealed that the slopes relating gaze and cursor exits were rather shallow (mean slope across participants = 0.16, range = 0.08–0.29). Thus, changes in gaze exit times were associated with statistically significant but modest changes in cursor exit times.

Figure 5A shows gaze exit times from the target zone relative to the time of the cursor exit for all individual trials pooled across all participants (negative time indicates gaze lead). The distribution in standard trials was bimodal, with the major mode representing a substantial gaze lead. Notably, the trough in the distribution was centered on zero time difference. That is, regardless of the gaze strategy or strategies that participants employed, they appeared to avoid initiating saccades at the time of launching of the transport phase (i.e., the cursor exit). In contrast, the corresponding distribution for the reactive trials from the reactive control condition had its peak frequency around zero time difference, suggesting that both cursor exits and gaze exits were triggered reactively and synchronously.

Figure 5B shows the times at which gaze entered the next target zone relative to when the cursor exited the current target zone for all individual trials pooled across all participants. Because most gaze exits occurred well before the cursor exit in standard trials, gaze entered the next target zone before the cursor left the current target in the majority of these trials. In contrast, in the majority of reactive trials from the reactive control condition and for some of the standard trials, gaze did not enter the next target zone before the cursor left the current target. Nevertheless, in almost all standard trials and reactive trials, gaze entered the next target before the cursor. This is illustrated in Figure 5C, which shows the time at which gaze entered the next target relative to when the cursor entered the next target for all individual trials pooled across participants. That is, because the duration of saccades to the next target was shorter than the
duration of the cursor movements to the next target, gaze was on the next target when the cursor entered regardless of the gaze strategy employed.

Although premature cursor exits occurred infrequently, we asked whether, during steady-state performance (Stage 7), the likelihood of generating a late gaze shift increased in trials following a premature cursor exit. To examine this question, gaze shifts were classified as early if they occurred between 0.2 and 0.6 s after the onset of the hold phase and late if they occurred greater than 0.7 s after the onset of the hold phase. This classification reasonably captures the two distributions of gaze shifts seen during steady-state performance (see Figure 4A). We then determined the percentage of late gaze shifts for all trials following a premature cursor exit and the percentage of late gaze shifts for all trials following a correct cursor exit.

Because of the low frequency of premature cursor exits, we combined data across participants. When considering all nine participants, we found that the percentage of late gaze shifts was greater (33%) for trials following a premature cursor exit than for trials following a correct cursor exit (25%). When considering only those participants who exhibited frequent early and late gaze shifts (i.e., P3, P6, and P9), we found that the percentage of late gaze shifts was also greater (51%) for trials following a premature cursor exit than for trials following a correct cursor exit (42%).

**Gaze and cursor behavior in catch trials**

During steady-state performance, four of the nine participants (P3, P6, P7, and P9) exhibited a substantial number of late gaze shifts from the target (Figure 3C). As shown previously, these late gaze shifts tended to occur after the cursor exits, suggesting that they might be purposefully delayed in order to monitor the time of cursor exit. However, these late gaze shifts occurred in the same time range as reactive saccades. Therefore, an alternative interpretation is that these late gaze exits during steady-state performance are reactive and may arise because of a failure to predict the end of the hold phase. We think this is unlikely, because all participants learned to consistently produce predictive cursor exits during steady-state performance and, with the exception of one participant (P7), often generated predictive gaze shifts as well.
To assess the control mechanisms underlying late gaze shifts, we examined gaze and cursor behavior in catch trials in which either auditory or both auditory and visual feedback about completion of the hold phase was delayed by 0.2, 0.25, or 0.3 s (see Methods). To determine the influence of such delayed feedback on gaze (and cursor) behavior, we selected catch trials in which the gaze shift occurred later than 0.8 s after the start of the hold phase, which corresponds to the minimum time for gaze shifts in reactive trials from the reactive control condition (Figure 3B). Therefore, gaze exits that occur earlier than 0.8 s are almost certainly predictive and are unlikely to be influenced by delayed feedback. Figure 6 shows the relation between gaze exit time and cursor exit time for all selected catch trials from Stages 1–6 (gray circles and squares) and Stage 7 (black circles).

Two distinct clusters of catch trials from Stages 1–6 can be observed, which can be clearly distinguished based on the time of the catch. First consider the catch trials that resulted in a substantial delay in the cursor exit as well as the gaze exit (filled gray circles). In these trials, most of which occurred in the earliest stages, it seems clear that the cursor and gaze shifts were generated in response to the delayed feedback about hold phase completion. In the reactive condition, cursor and gaze exits occurred, on average, approximately 0.3 s after hold phase completion when the unpredictable hold phase duration was 0.6 s (Figure 3B). The dashed vertical and horizontal black lines in Figure 6 are located at 0.3 s after the earliest delayed feedback in catch trials (which occurred at 0.8 s) and the dashed vertical and horizontal gray lines are located 0.3 s after the latest delayed feedback in catch trials (which occurred at 0.9 s). It is perhaps not surprising that most of the delayed cursor and gaze exits in these catch trials occurred more than 0.3 s after delayed feedback was provided, since this feedback occurred well after the expected time. Note that the catch trials fall along the unity line (dashed-dotted line in Figure 6), indicating that gaze and cursor exits were initiated in synchrony. A linear regression applied to these trials yielded a significant positive relation \( r = 0.72; p < 0.001 \) with a slope of 1.087 and an intercept of \(-0.101 \) s.

Consider next the catch trials from Stages 1–6 that did not result in a substantially delayed cursor exit (open gray circles in Figure 6), indicating that the cursor exit was not reactively triggered. Although many of the gaze exits in these trials occurred in the reactive range (i.e., after 1 s), there is evidence to suggest that they are not reactively triggered but rather are intentionally delayed so that the cursor exit can be monitored in central vision. Specifically, a linear regression applied to these trials yielded a significant positive relation \( r = 0.77; p < 0.001 \) with a slope of 1.008 and an intercept of 0.238 s. Thus, these gaze exits were temporally coupled with the cursor exits but delayed by about a quarter of a second on average.

Finally, consider the catch trials from Stage 7 (filled black circles in Figure 6). None of the catch trials resulted in a delayed cursor exit, and only three (out of 18) resulted in a gaze exit in the range expected for reactively triggered gaze shifts (i.e., \( >1 \) s). With the exception of these three trials, the gaze and cursor exit times in these catch trials were similar to those seen in corresponding noncatch (i.e., standard) trials. The solid vertical and horizontal lines represent the mean cursor and gaze exit times for all noncatch trials from Stage 7 in which the gaze shift occurred later than 0.8 s after the
start of the hold phase (the same selection criterion as the catch trials). These results strongly suggest that the late gaze shifts observed during steady-state performance are not equivalent to the reactive gaze exits observed in early learning stages.

All of the late gaze shifts observed when the cursor exit was not reactively triggered (i.e., both the filled black and open gray circles in Figure 6) can be well fit by a single regression line (not shown in Figure 6) with a slope of 1.020 and an intercept of 0.220 s (r = 0.76; p < 0.001). This indicates that, on average, the gaze exit occurred about 220 ms after the cursor exit, which is considerably less than the average reaction time for reactivly triggered gaze shifts (approximately 300 ms). Thus, whereas it is possible that some of these late gaze shifts are “triggered” by the cursor exit, it seems likely that most are planned in advance.

Discussion

For spatial motor control, it has previously been demonstrated that gaze behavior changes dramatically across learning stages during practice of a novel visuomotor task (Sailer, Flanagan, & Johansson, 2005). However, little is known about how gaze behavior changes during learning of novel tasks that also involve significant temporal control, such as tasks that comprise sequentially linked action phases. The aim of the current study was to investigate how gaze behavior supports both learning to predictively link successive action phases and skilled linking of action phases after learning. To that end, we used a laboratory task that involved both spatial and temporal control and that captures several important features of natural manipulation tasks. Specifically, action phases are demarcated by discrete multimodal sensory events that signal completion of the goal, or subgoal, of the phase; each phase needs to be completed before the next phase can be performed; and optimal performance in the task requires predictive linking of action phases, whereby the motor commands for the forthcoming phase are launched in anticipation of expected sensory events associated with completion of the current phase (Johansson & Flanagan, 2009; Säfström et al., 2013). Our results show that although participants directed their gaze to each successive target throughout the experiment, the timing of gaze shifts between targets changed dramatically with learning.

Gaze behavior when initially learning the task

During initial learning of the task, as well as in reactive trials from the reactive control condition, gaze shifts to forthcoming targets appeared to be reactively triggered by sensory events signaling completion of the required hold period. These gaze shifts were initiated in synchrony with cursor movements to the forthcoming targets, which were also reactively triggered. Thus, gaze typically fixed the current target and the cursor during the entire hold phase. By maintaining gaze in the target zone, participants could use central vision to help keep the cursor within the target zone during the hold phase, which required actively applying precisely controlled forces on the handle used to control the position of the cursor. Moreover, because gaze did not exit the target zone until well after the end of the required hold period, visual events signaling completion of this period (i.e., the brief doubling of the size of the current target) could be monitored in central vision. Such monitoring may have been useful in learning the duration of the required hold period, a prerequisite for learning to predictively launch the transport phase. Approximately synchronous initiation of eye and hand motor commands has been shown to occur in goal-directed movements generated in response to the appearance of visual targets (Biguer, Jeannerod, & Prablanc, 1982; Gribble, Everling, Ford, & Mattar, 2002; Sailer et al., 2005). Because the inertia of the arm is much larger than that of the eye, such synchrony results in eye movement onset preceding arm movement onset in reaching and pointing tasks (Prablanc, Echallier, Jeannerod, & Komilis, 1979). However, in the current task, inertial differences between hand and gaze actions were reduced because cursor movements were generated by applying isometric forces to a grasped handle. Thus, we would expect synchronous gaze and hand motor commands to result in near synchronous gaze and cursor movement onset.

Gaze behavior when predictively linking action phases after learning

Along with the development of predictive linking of action phases that occurred during learning, the timing of gaze shifts between targets changed considerably. Two different modes of gaze behavior were observed, both within and across participants.

The first mode of gaze behavior involved gaze shifts to the next target well in advance of the end of the hold phase. By fixating the next target in advance of the initiation of the cursor movement, participants would be able to use proprioceptive and/or motor signals (i.e., efference copy) related to gaze position to help plan the movement of the cursor to the next target well in advance of movement initiation (Paillard, 1996; Prablanc et al., 1979). These gaze-position-related signals would also be available throughout the cursor movement to help guide the cursor to the target (Crawford,
Medendorp, & Marotta, 2004; Goodale, Pelisson, & Prablanc, 1986; Prablanc & Martin, 1992). Therefore, shifting gaze to the target before the cursor starts to move would enable optimal use of visual feedback throughout the cursor movement. The early gaze shifts we observed during late learning and steady-state performance are reminiscent of “look ahead” saccades that occur in many familiar manual tasks involving object manipulation (Hayhoe & Ballard, 2005; Land & Hayhoe, 2001; Mennie, Hayhoe, & Sullivan, 2007). For example, in tasks such as making tea or sandwiches, gaze fixations often move to the next object to be manipulated as early as 1 s before the current action phase is complete (Hayhoe, Shrivastava, Mruczek, & Pelz, 2003; Land, Mennie, & Rusted, 1999). For all participants, we found that the sooner gaze shifted to the next target, the earlier the cursor exited. A possible explanation for this finding is that the time required to initiate the planned movement depends on how well the planned movement is prepared. This suggestion is in line with studies showing that preparatory processes in premotor cortex during delay periods (between instruction stimuli and go signals) reduce reaction times in initiating planned movements (Churchland, Santhanam, & Shenoy, 2006; Churchland & Shenoy, 2007). Given this explanation, our finding suggests that directing gaze to the forthcoming target ahead of movement initiation facilitates the specification of the motor commands used for cursor transport towards the next target.

The second mode of gaze behavior observed during late stages of learning and steady-state performance involved maintaining gaze at the current target until after the cursor exit. This gaze behavior is reminiscent of “checking” fixations that occur in many natural tasks, for example in making tea to confirm that the kettle is full (Land et al., 1999). Our analysis of catch trials in which feedback about hold phase completion was delayed indicates that these late gaze shifts are distinct from the reactively triggered gaze shifts observed during earlier stages of learning. Although these late gaze shifts occurred roughly at the same time, relative to the onset of the hold phase, as reactive gaze shifts, our results suggest that they are intentionally delayed so that the time of the cursor exit relative to the end of the required hold period can be monitored in central vision. We have previously suggested that such gaze behavior contributes to the development and maintenance of sensorimotor correlation matrices that support predictive motor control in manipulation (Flanagan et al., 2006; Johansson et al., 2001) and thus could help participants in our task to maintain optimal performance. Consistent with the idea that late gaze shifts are involved in maintaining the calibration of cursor exit time, we found that, during steady-state performance, the likelihood of observing a late gaze shift increased in trials following a premature cursor exit.

The two modes of steady-state gaze behavior we observed have been documented in sequential target-directed action tasks. Look-ahead fixations have been described for a variety of tasks ranging from tea making (Land et al., 1999) to high-speed cup stacking (Foerster et al., 2011) and sequential target reaching (Wilmut, Wann, & Brown, 2006). Gaze shifts that occur at around the same time as or just after hand movement onset have also been reported in a range of tasks. For example, in a previous study (Bowman et al., 2009) involving a task in which participants successively contacted a series of virtual target objects, we found that, on average, gaze shifted to the next target 50–100 ms after contact (see also Epelboim et al., 1995; Sailer et al., 2005). Likewise, in our work on eye-hand coordination in naturalistic object manipulation tasks, we have reported that gaze often shifts from the current object to the next object at around the time the hand (or object in hand) departs from the current object towards the next object (Flanagan & Johansson, 2003; Johansson et al., 2001).

Although the relative frequency of these two modes of gaze behavior varied across participants, several participants regularly employed both modes during steady-state performance. Importantly, due to the timing of the task, only one mode of gaze behavior could be employed for a given target or trial. This suggests that when repeatedly performing manual tasks, people may use multiple gaze behaviors that serve different functions, and then integrate the information gained from these behaviors across repetitions. This strategy would circumvent the fundamental limitation that gaze can only be directed to one location at a time. Such a strategy would presumably be useful in many continuous or repetitive tasks in which there is competition among events and locations for attention and central vision, and learning may involve fine-tuning of gaze allocation. Indeed, Sims and colleagues (2011) have shown that people adaptively allocate gaze among competing perceptual and motor components of a task so as to optimize task performance. Of course, given that gaze behavior is fundamentally task specific (for a review, see Land, 2006), a varying number of modes with different prominence may emerge in different tasks. The extent to which early (look-ahead) versus late (checking) gaze shifts are observed in a given task likely depends on a number of features of the task, including spatial and temporal accuracy requirements, various performance demands, and the skill of the performer. For example, whereas Neggers and Bekkering (2000, 2001) found that gaze is anchored to the reach target when reaching to relatively small targets, this anchoring is not observed in reaching to larger targets, requiring less
spatial accuracy, such that participants are able to redirect their gaze to another target while reaching to the first (Terrier et al., 2011). In their study of a high-speed cup-stacking task, Foerster and colleagues (2011) found that participants rarely monitored the completion of task subgoals, presumably because there was a premium on looking ahead in order to perform the task as quickly as possible.

With learning in our task, the initiation of eye and hand movements became separated such that, regardless of the mode of gaze behavior employed, participants tended to avoid making saccades at the time of the cursor exit. We suggest that this separation allowed participants to monitor events related to the cursor exit in central vision. When gaze exited after the cursor, participants directly viewed the cursor exit from the target as well as the disappearance of the target that occurred at the same time. When gaze shifted to the next target well in advance of the cursor exit, participants viewed the change in the visual appearance of the next target, which occurred at the time of the cursor exit. In contrast, during initial performance in standard trials, as well as in the reactive condition, the synchronous initiation of eye and hand movements implies that participants did not reliably monitor events associated with the cursor exit in central vision. This difference in gaze behavior can be understood when one considers the information required by the sensorimotor system at different stages of learning. During steady-state performance, the cursor exited, on average, very soon after the end of the required hold phase. Optimal performance in the task requires carefully controlling the timing of cursor exits so that they occur at short latency, on average, but not so short as to result in an excess of premature exits due to temporal uncertainty (Säfström et al., 2013). By monitoring the timing of these cursor exits relative to the end of the required hold period, the brain may improve the timing of these exits across trials and thus optimize task performance. In contrast, when cursor exits are reactively controlled during early learning, the sensorimotor system is not primarily concerned with the precise timing of cursor exits. That is, the system is not attempting to maintain a small time clearance between the end of the required hold phase and the cursor exit. Instead, the system is concerned with learning to estimate the duration of the required hold phase, so that subsequent predictive control can be effective. Thus, gaze need only capture the sensory events associated with the end of the required hold period (in addition, of course, to the start of the hold period).

Recently, Shalom and Sigman (2013) examined a sequential task in which participants were required to type letters on a keyboard that corresponded to letters on a computer screen, and found that participants avoided making simultaneous manual and eye movements. This finding was interpreted in terms of a refractory period between the initiation of manual and eye movements reflecting a processing bottleneck which makes it difficult to perform both movements simultaneously. In contrast, we suggest that, in our task, the asynchrony of gaze movements and hand movements is planned and allows extraction of task-relevant information that is useful for sensorimotor control. We also note that during early stages of learning (and in the reactive control condition), participants generated synchronous gaze and cursor movements.

### Conclusions

We have shown that learning to predictively link action phases, which is a key component of skill acquisition in manual tasks, is associated with marked changes in gaze behavior. We suggest that these learning-related changes reflect the functional requirements regarding gaze behavior imposed by the task at different stages of learning. During early learning, during which the action phases were reactively linked, gaze and hand movements were initiated synchronously. Gaze supported fine cursor control associated with the hold phase as well as learning about the duration of the hold phase. However, as predictive linking of action phases developed with practice, gaze and hand movements became asynchronous, with two distinct modes of gaze behavior observed both within and across participants. On the one hand, gaze supported the planning of the cursor movement towards the next target. On the other hand, gaze monitored the cursor exit relative to the end of the required hold phase in central vision so as to maintain optimal task performance. We suggest that, when repeatedly performing manual tasks, people may use multiple gaze behaviors that serve different functions, and then integrate the information gained from these behaviors across repetitions. This strategy would entail that gaze resources can be adaptively allocated across repetitions of the task and would circumvent the fundamental limitation that gaze can only be directed to one location at a time.

**Keywords:** gaze behavior, object manipulation, sensorimotor control, motor learning, multisensory

### Acknowledgments

This work was supported by the Swedish Research Council (Project 08667), the Strategic Research Program in Neuroscience at the Karolinska Institute,
References


Prablanc, C., Echallier, J. E., Jeannerod, M., & Komilis, E. (1979). Optimal response of eye and
hand motor systems in pointing at a visual target: II. Static and dynamic visual cues in the control of hand movement. *Biological Cybernetics, 35,* 183–187.


