This paper appears in Disability andRehabilitation: Assistive Technology,
DOI: 10.1080/17483107.2020.1729874

Effect of feedback and target size on eye gaze accuracy in an off-screen task

Isao Sakamaki[a], Kim Adams[a], Mahdi Tavakoli[b] and Sandra A. Wiebe[c]

[a]Faculty of Rehabilitation Medicine, University of Alberta, Edmonton, AB, Canada; [b]Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB, Canada; [c]Department of Psychology, University of Alberta, Edmonton, AB, Canada

ABSTRACT

Purpose: Eye gaze interfaces have been used by people with severe physical impairment to interact with various assistive technologies. If used to control robots, it would be beneficial if individuals could gaze directly at targets in the physical environment rather than have to switch their gaze between a screen with representations of robot commands and the physical environment to see the response of their selec-tion. By using a homogeneous transformation technique, eye gaze coordinates can be mapped between the reference coordinate frame of eye tracker and the coordinate frame of objects in the physical envir-onment. Feedback about where the eye tracker has determined the eye gaze is fixated is needed so users can select targets more accurately. Screen-based assistive technologies can use visual feedback, but in a physical environment, other forms of feedback need to be examined.

Materials and methods: In this study, an eye gaze system with different feedback conditions (i.e., visual, auditory, vibrotactile, and no-feedback) was tested when participants received visual feedback on a dis-play (on-screen) and when looking directly at the physical environment (off-screen). Target selection tasks in both screen conditions were performed by ten non-disabled adults, three non-disabled children, and two adults and one child with cerebral palsy.

Results: Tasks performed with gaze fixation feedback modalities were accomplished faster and with higher success than tasks performed without feedback, and similar results were observed in both screen conditions. No significant difference was observed in performance across the feedback modalities, but participants had personal preferences.

Conclusion: The homogeneous transformation technique enabled the use of a stationary eye tracker to select target objects in the physical environment, and auditory and vibrotactile feedback enabled partici-pants to be more accurate selecting targets than without it.

Introduction

The experience of physically manipulating objects in the environment has a large influence on cognitive development in children [1,2]. Cognitive development refers to the development of children in terms of thinking, resolving, learning, feeling, and knowing the environment [3]. Physical manipulation has been identified as a critical motor experience that enables children to learn skills, such as the emergence of symbols, referential communication and the understanding of relations between objects [4]. For children who have complex physical disabilities that prevent them from reaching and grasping objects, one of the biggest concerns is lacking opportunities for meaningful manipulation tasks, often in the context of play activities [5]. This lack of opportunities may negatively affect the progressive development of their learning skills and mental growth [3].

Robots have been utilized by children who have physical impairments to access play activities [6]. Robot systems can behave like extended arms, allowing children to reach what they otherwise could not reach or probe what they otherwise could not probe. However, these technologies often still require a certain degree of physical ability to access and to operate, such as using switches or joysticks. Eye gaze has been used to control assistive technology for many years [7], and recently, the cost of eye trackers has gone down, making it a feasible access method to control robots.

The most common setup for eye gaze is to have the user fixate on graphic target options on a screen (called on-screen herein). The users generally rely on feedback about where the tracker is interpreting the gaze, such as a mouse pointer for the selection. For example, individuals with severe physical impairment can generate synthesized speech for communication by selecting symbols on a screen [8]. Arai and Yajima [9] developed a feeding aid system using a robot arm controlled by an eye gaze interface. The user gazed at the desired food on the screen and then the robot picked up the food to bring it closer to the user. In Encarnação et al. [10], children controlled a LEGO Mindstorms robot (Lego A/S, Billund, Denmark) with an eye gaze-tracking system that enabled children with physical impairment to participate in academic activities. Simple robot commands were displayed on a computer screen and users controlled the robot by fixating their gaze on the command. The system showed positive impact on children with physical impairment, but Encarnação et al. [10] pointed out that it required considerable effort for children to look at the screen to select the robot command and then look at the robot to check its effect. This forced the user to keep changing their visual attention during the tasks and added a layer of complexity.

When using eye gaze to control robots, it would be better if the user did not have to look at a computer screen to select robot movements. Using head-mounted eye trackers could be one way to accomplish this [11]. For instance, eye gaze estimation using a head-mounted tracker helped to reveal how humans gather information from their environment and how they use that information in motor planning and motor execution [12]. Galante and Menezes [13] developed a head-mounted eye tracker enabling the system to estimate gaze position in the physical environment by mapping between the camera frame view and actual gaze direction using geometric calibration techniques. However, these are expensive, and some people cannot tolerate wearing them.

A stationary eye gaze interface could be used without a computer screen to select objects in a physical environment (called off-screen herein) if geometric calibration techniques are used. However, to use it, feedback to the user about where the tracker is interpreting the gaze is crucial. For robot control in the physical world, visual feedback of a mouse pointer on a screen is not possible, thus, other kinds of feedback are needed.

The feasibility of alternative feedback modalities for gaze applications on-screen have been investigated [14–17]. Majaranta et al. [15] explored a combination of auditory and visual feedback in eye typing. The authors found that the visual-auditory feedback significantly improved user text entry speed and satisfaction compared to visual feedback alone. Boyer, Portron, Bevilacqua, and Lorenceau [18] investigated whether real-time auditory feedback of eye movement improved an on-screen target tracking task in the absence of visual feedback. Although large individual differences were observed, the auditory feedback did modify the oculomotor behaviour and improved task performance. Kangas et al. [19] compared off-screen gaze interaction using gaze gestures (looking right then left to activate a command) with vibrotactile feedback and no feedback. All 12 participants performed the gaze interaction faster and preferred the vibrotactile feedback over no feedback. Auditory and/or vibrotactile feedback could be feasible ways for users to confirm their gaze interaction when using a robot in a physical environment for play.

Some research has pointed out challenges using eye gaze-based assistive technology with a clinical population. Amantis et al. [20] found that children with cerebral palsy responded more slowly and less accurately in gaze performance compared to non-disabled children. Dhas, Samuel, and Manigandan [21] found that gaze interaction applications were not suitable for children who had too many involuntary movements because the eye tracker lost accuracy in determining eye gaze direction. The main challenge for efficient gaze interaction is how to distinguish between gaze intended to gather visual information versus gaze to activate a specific command. This problem often results in unintended selections, which is called the Midas' touch problem [22]. One solution for this problem is to employ dwelling, requiring the user to fixate gaze for a prolonged period of time on the target option. A typical dwell time for eye typing using a screen-based eye gaze system is approximately 0.5–1 s [23]. Adjusting the dwell time or the diameter of the target acceptance size may allow for more tolerance with involuntary movements, and help users to be more successful in accomplishing the gaze interaction.

On-screen gaze interaction is well established and researched; however, off-screen gaze interaction using dwell selection, such as selecting an object in a physical play task scenario for the robot to move towards, is novel in the field of assistive technology. Auditory and vibrotactile feedback could help users to be more accurate at target selection, but appropriate settings for the device features need to be examined. In this study, we developed a gaze interaction system that maps eye gaze direction between the reference coordinate frame of a stationary eye tracker and the coordinate frames of objects in on- and off-screen environments. The effect of different feedback modalities (i.e., no-feedback, visual feedback, auditory feedback, and vibrotactile feedback) about where the eye tracker determined the eye gaze to be fixated in a target selection task, and the effect of different target sizes, were examined.

The research questions addressed in the study were as follows:

1. How do the speed and success of gaze interaction differ between on-screen and off-screen conditions?
2. Which feedback modalities and target size make the gaze-based target selection faster and more successful?
3. What is the feedback preference of the participants in the on- and off-screen gaze interaction?

Methods

Participants

Ten university students without physical impairment, three males and seven females, aged from 22 to 38 (26 ± 4.1), participated in the study (called A1–A10 herein). The system was also tested by two adults with quadriplegic cerebral palsy (a 52-year-old female and a 33-year-old female, called B1 and B2, respectively), three non-disabled children (a 10-year and 2-month-old boy, a 7-year and 10-month-old girl, and a 6-year and 4-month-old girl, called C1, C2, and C3, respectively), and a child who had right side spastic hemiplegic cerebral palsy (a 7-year and 4-month-old boy, D1). Participants B1 and B2 have mixed high and low muscle tone and involuntary movements and perform mobility by using a powered wheelchair. B1 is affected by strabismus and has difficulty focussing on objects with both eyes simultaneously, while B2 has no visual impairment. Participant D1 has no visual impairment, however, he was diagnosed with Attention Deficit Hyperactivity Disorder (ADHD) which may cause reduced gaze concentration (greater spread of vertical and horizontal eye movements) [24]. Ethical approval was received from the local Health Research Ethics Board Health Panel at the University of Alberta.

Design

There were two experimental screen conditions in the study: a gaze interaction task with a computer screen, called on-screen condition, and a gaze interaction task with the physical environment, called off-screen condition. In each screen condition, different target size and feedback modalities were examined.

Experimental setup

The system diagram of the on- and off-screen experimental set up is shown in Figures 1 and 2. A Windows-based computer and a stationary eye tracker, Tobii eye tracker 4 C (Tobii Technology,

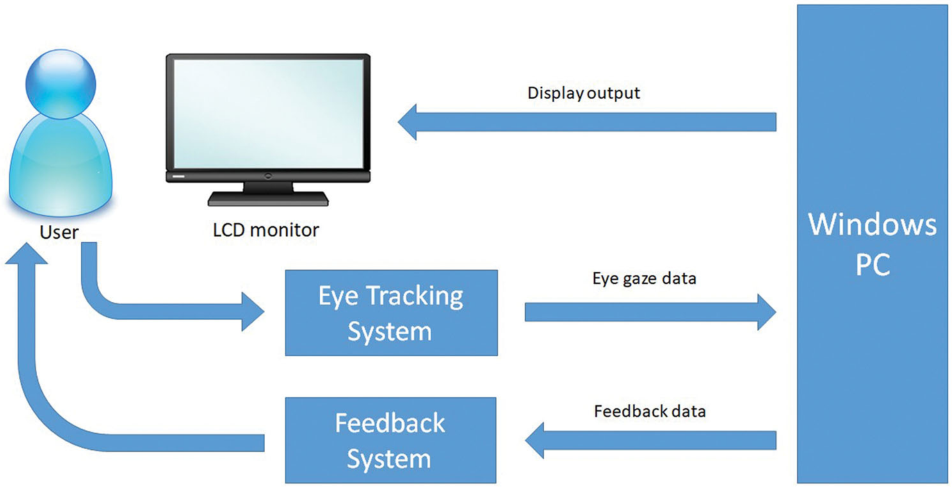

Figure 1. Schematic diagram of the system in the on-screen experiment.

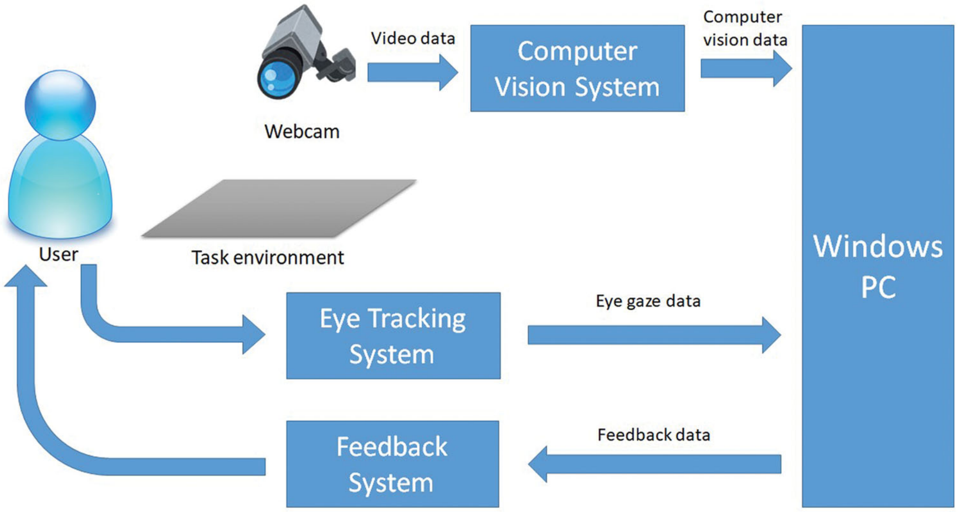

Figure 2. Schematic diagram of the system in the off-screen experiment.

Danderyd, Sweden), and external devices for feedback modalities were the basic components for both the on- and off-screen conditions. A 19-inch LCD monitor (42 cm × 24 cm) was added for the on-screen condition, and a computer vision system was added for the off-screen condition. The eye gaze acquisition was performed in MATLAB (MathWorks, Nadick, MA, USA). The feedback system, computer vision system, and interconnecting of systems were programmed in LabVIEW (National Instruments, Austin, TX, USA). Details of each component are explained below.

Eye gaze acquisition system

The eye tracker was placed in front of the task environment and connected to a Windows PC with a sampling frequency rate of 90 Hz, in order to monitor fixation of the gaze during the gaze interaction. The dwell time was set to 1.5 s in all the conditions. A longer dwell time than typical was selected in this study to make sure participants had enough time to select the target during the off-screen condition, based on pilot testing of the system. When the participant fixated their gaze on the target for 1.5 s, the system recognized it as the target that the participant desired to select. If the participant's gaze came off the target before 1.5 s and then back on the target, counting of the dwell time started over again.

Feedback system. In the on-screen condition, the LCD monitor showed a standard arrow-shaped mouse pointer as the visual feedback. This pointer was controlled by the participant's eye movement. There were also no-feedback, auditory feedback, and vibrotactile feedback used in both on- and off-screen conditions. For the no-feedback in the on-screen condition, there was no cursor displayed on the computer screen. A USB stereo sound adapter was used to generate the output of a 100 Hz sine wave for the auditory and vibrotactile feedback modalities. For the auditory feedback, the sine wave was outputted to earphones that the participants wore, and for the vibrotactile feedback, the wave was sent to an amplifier to drive a vibration motor (Bit Trade One, Kanagawa, Japan) on which the participants placed a fingertip during the trials. Both auditory and vibrotactile feedback were initiated when the participant's gaze was within the set target acceptance size radius, and the amplitude of the feedback increased in proportion to the time the gaze was on the target as an indication of the target being selected.

Computer vision system

The computer vision system used in the off-screen condition was a USB webcam (Dynex, Richfield, MN, USA) (see Figure 2). The location and colour of each target in the task environment were

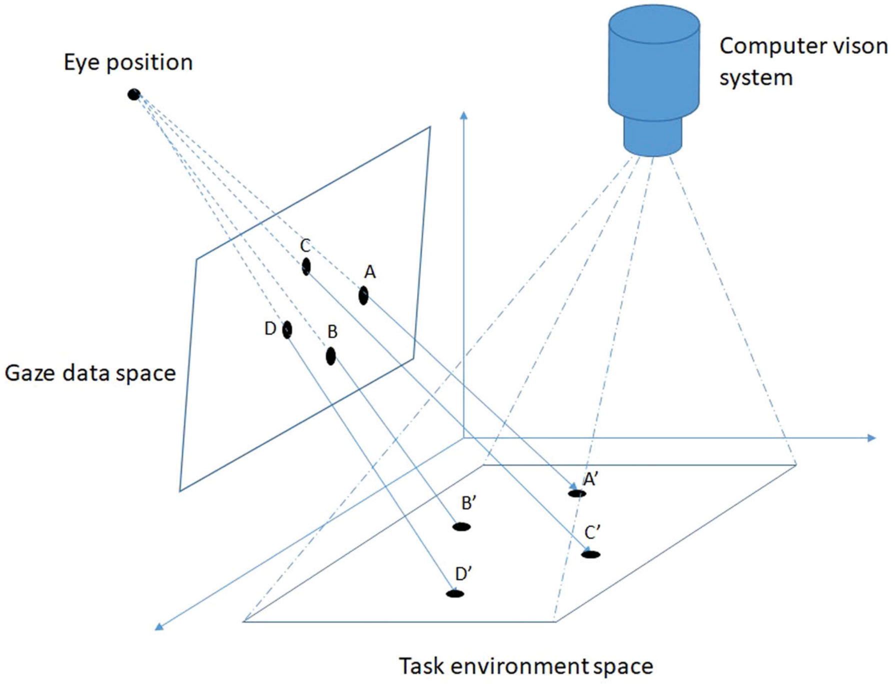

Figure 3. Two corresponding points on different space are converted by a homography.

detected by an object recognition programme coded in LabVIEW. Since the Tobii eye tracker, 4 C was designed for gaze interaction in two-dimensional on-screen space, the participant’s gaze was mapped into the two-dimensional plane of the task environment. The gaze mapping was performed by the following steps:

1. A template on which four calibration points were printed was placed in the task environment.
2. The calibration template was captured with the webcam mounted above. Then, the centre points of each calibration point were computed by the object recognition programme.
3. The participant fixated their gaze at each calibration point in turn. The gaze position detected by the eye tracker at each calibration point was collected.
4. Each gaze position was mapped to each calibration point on the task environment using a projective homogeneous transformation.

A homography is a perspective transformation of a plane, that is, a reprojection of a plane from one space into a different space as shown in Figure 3. For the homography, the relationship between two corresponding points can be written as follows [25]:

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \tag{1}$$

where $\begin{bmatrix} x & y & 1 \end{bmatrix}^{\mathrm{T}}$ represents the gaze position obtained by the eye tracker when the participant is looking at a calibration point, $\begin{bmatrix} x' & y' & 1 \end{bmatrix}^{T}$ represents a calibration point in the task environment obtained by the computer vision system, and the 3×3 matrix represents a homogeneous transformation.

Procedures

On-screen condition. There was a run of twelve trials in each of four different feedback conditions (i.e., visual feedback, no-feedback, auditory feedback, and vibrotactile feedback) in the on-screen condition (i.e., 12 trials × 4 conditions = 48 targets to select in total). The participant sat at a distance of 60 cm from the computer monitor and eye gaze tracker. At the beginning of the on-screen condition, the Tobii gaze tracking utility software was used to calibrate the participant’s eye gaze.

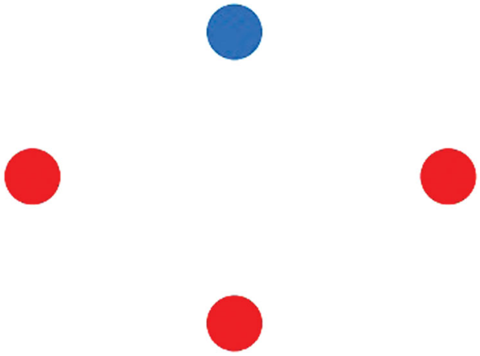

Figure 4. Four circular targets displayed on the LCD monitor for the on-screen experiment.

For the experiment, four circles with a diameter of 3 cm were displayed (see Figure 4). The target was a blue circle, and the remaining stimuli were red circles. The target location was randomized within the four circles, and the participants needed to fixate their gaze on the new blue target each time. A fixation cross was displayed at the centre of the screen during each inter-trial interval. The target acceptance size was also changed randomly in each trial. The orders of acceptance size and location were counterbalanced. The diameter of the target acceptance sizes tested for the non-disabled adult participants were as follows: 3 cm, 6 cm, and 9 cm. The diameter of the acceptance sizes used for the adult participants with physical impairments and the child participants with and without physical impairments were: 6 cm, 9 cm, and 12 cm. The sizes were larger because the 3 cm diameter was too difficult for these population groups to achieve success in selecting, according to their results in pre-experiments. With three target acceptance sizes at each of the four target

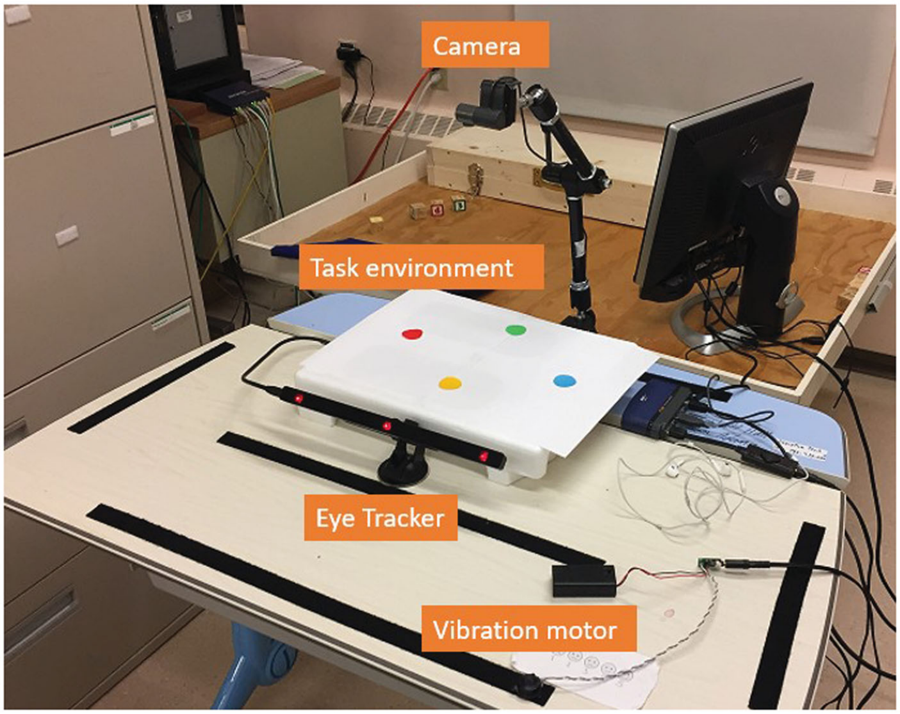

Figure 5. Four circular targets in the task environment for the off-screen experiment.

locations, a participant experienced each location and target size combination once for each feedback condition.

Off-screen condition. There was a run of 12 trials in each of the three different feedback conditions (i.e., no-feedback, auditory feedback, and vibrotactile feedback) in the off-screen condition (i.e., 12 trials × 3 conditions = 36 targets to select in total). The target acceptance sizes, which were the same dimensions as in the on-screen condition (i.e., 3, 6, and 9 cm for non-disabled adult participants and 6, 9 and 12 cm for the adult participants with physical impairments and the child participants with and without physical impairments), were changed randomly during each trial.

The dimensions of the task environment were set to resemble the 19 inch LCD monitor used in the on-screen condition. The environment where the targets were placed was located 60 cm away from the participants, and the eye tracker was placed in front of the targets as shown in Figure 5. At the beginning of the off-screen condition, the calibration procedure using the computer vision system was performed as described above. For the experiment, an image of four printed circular objects with different colours (i.e., red, green, yellow, and blue) with a diameter of 3 cm was placed in the task environment. The four different colours were used in this screen condition because the targets were fixed, and not able to change colour like they could in the on-screen condition. The participants were given verbal instructions from the computer on which coloured objects they needed to fixate their gaze during the trial. The order of the target acceptance size and the location were counterbalanced. There was a cross at the centre of the task environment for the participants to return their gaze between the target selections. A participant experienced each target size and location combination once in each feedback condition (3 sizes × 4 locations = 12 targets).

At the end of the session, the participants answered a questionnaire where they were asked to rank the feedback modalities according to their preference. The participants were also asked if they had any comments.

Measurements and analysis

The dependent measures were as follows:

- *Target selection time:* To compare which condition was faster, the time from the task cue until the target was selected by the gaze was measured in milliseconds.
- *Timeout error rate:* To compare success at selecting targets, the timeout error rate was calculated, i.e., when the participant could not select the target within 10 seconds. The timeout error rate was the proportion of trials when the task timed out.

To examine the target selection time of the 10 non-disabled adult participants (A1–A10), the Shapiro–Wilk normality test was performed first to check if the data were normally distributed. If the data were normally distributed, a two-way repeated-measures analysis of variance (ANOVA) was applied with the following factors: factor 1 was the target acceptance size (3 levels: 3, 6, and 9 cm); factor 2 was feedback modality (with 4 levels for the on-screen condition: visual feedback, no-feedback, auditory feedback, and vibrotactile feedback, and 3 levels for the off-screen condition: no-feedback, auditory feedback, and vibrotactile feedback). A probability of $p<.05$ was considered significant. For the adult participants with physical impairments (B1 and B2), non-disabled child participants (C1, C2, and C3), and the child participant with physical impairments (D1), individual task performance was evaluated based on visual inspection of the data and descriptive statistics because of the small, heterogeneous sample.

To compare on- and off-screen conditions overall, an overall average target selection time in each screen condition was calculated per participant group by averaging the target selection times of all the feedback modalities and the acceptance sizes. Also, the average timeout error rates in each experimental condition were calculated and compared for each target size and feedback for the participant groups.

Results

Target selection time

Non-disabled adult participants

Results of the target selection time for the non-disabled adult participants in the on- and off-screen experiments are shown in Figure 6(a,b), respectively. Target acceptance size had a significant effect for both the on- and off-screen conditions (*F* [2, 18] = 44.77, *p*=.001 for on-screen and *F* [2, 18] = 30.84, *p*=.001 for off-screen). Feedback modality was also significant (*F* [3, 27] = 4.40, *p*=.012 for on-screen and *F* [2, 18] = 6.588, *p*=.019 for off-screen). The main effects were qualified by interactions between target acceptance size and feedback modality only for the on-screen condition, but the interaction did not reach significance for the off-screen condition (*F* [6, 54] = 5.186, *p*=.008 for on-screen and *F* [4, 36] = 2.80, *p*=.091 for off-screen).

According to the *post hoc* Tukey test for the paired comparison, the acceptance size of 3 cm differed significantly from other acceptance sizes for both the on- and off-screen conditions. In both screen conditions, the no-feedback was significantly different from other feedback modalities. Lastly, the ANOVA showed a difference in the screen conditions where the target selection time in the off-screen condition was significantly longer than in the on-screen condition (*F* [1, 9] = 62.541, *p*=.0001). Note that the visual feedback modality was excluded for the comparison because it was only presented for the task in the on-screen condition.

Adult participants with physical impairments

The results of the on- and off-screen conditions for the two adults with physical impairments, B1 and B2, are shown in Figure 7(a,b). In the on-screen condition, there was a clear performance difference between feedback modalities in the 6 cm target acceptance size for B1. The visual feedback and the no-feedback selection time was longer than those for vibrotactile and the auditory feedback. However, not much difference was observed between feedback modalities in all the target acceptance sizes for B2. For the off-screen condition, the data for B1 shows that the target selection time increased as the target acceptance size got smaller. Also, the target selection time in the no-feedback condition with the 6 cm target acceptance size appears to be longer than other feedback modalities. There seems to be no trend in the data for B2 in terms of the performance in each of the feedback modalities in both screen conditions. The overall average target selection time for the off-screen condition was 43.4% longer than the on-screen condition for B1 and 24.3% longer for B2.

Non-disabled child participants

Participant C1 seemed to have a trend that the longer selection time was with the auditory feedback in the on-screen conditions as shown in Figure 8(a). However, it does not appear that the auditory feedback was clearly longer than other feedback modalities in the off-screen condition (see Figure 8(b)). The on-screen condition for C2 appeared not to have much difference in the target selection time among the feedback modalities. On the contrary, no-feedback was greater than the other feedback modalities for all the target acceptance sizes in the off-screen condition. Also, the target selection time appeared to increase as the acceptance size got smaller. From the data for C3, the feedback that had the shortest selection time was the visual feedback in the on-

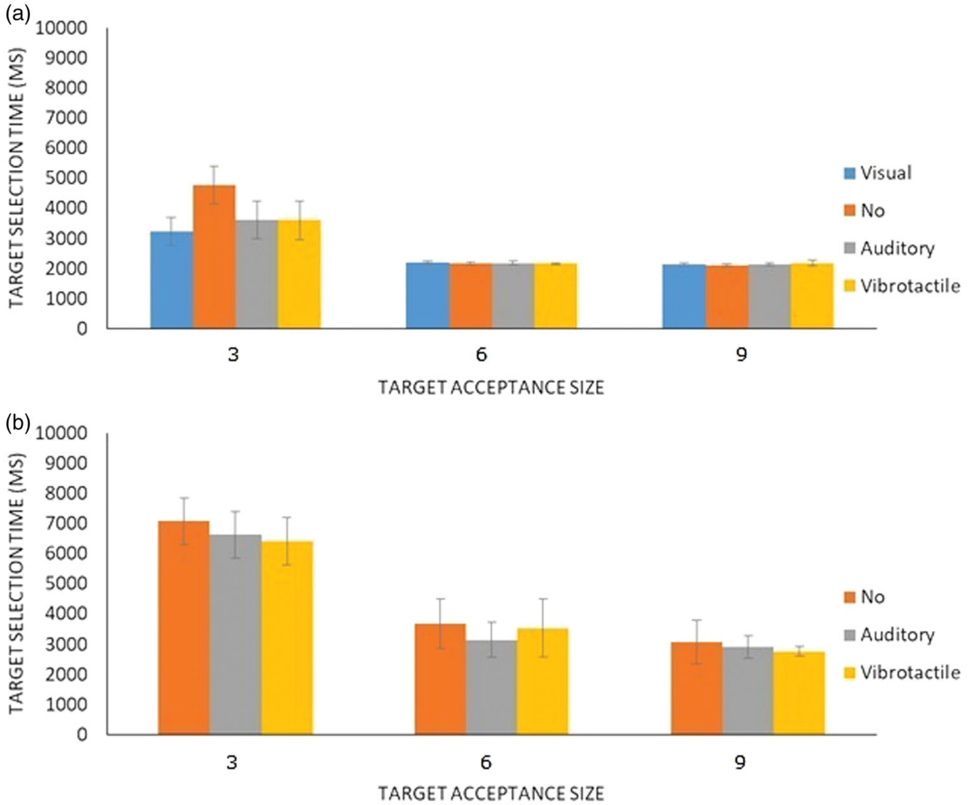

Figure 6. Target selection time with the different target acceptance size and feedback modalities for the 10 non-disabled adult participants for (a) the on-screen experiment and (b) the off-screen experiment.

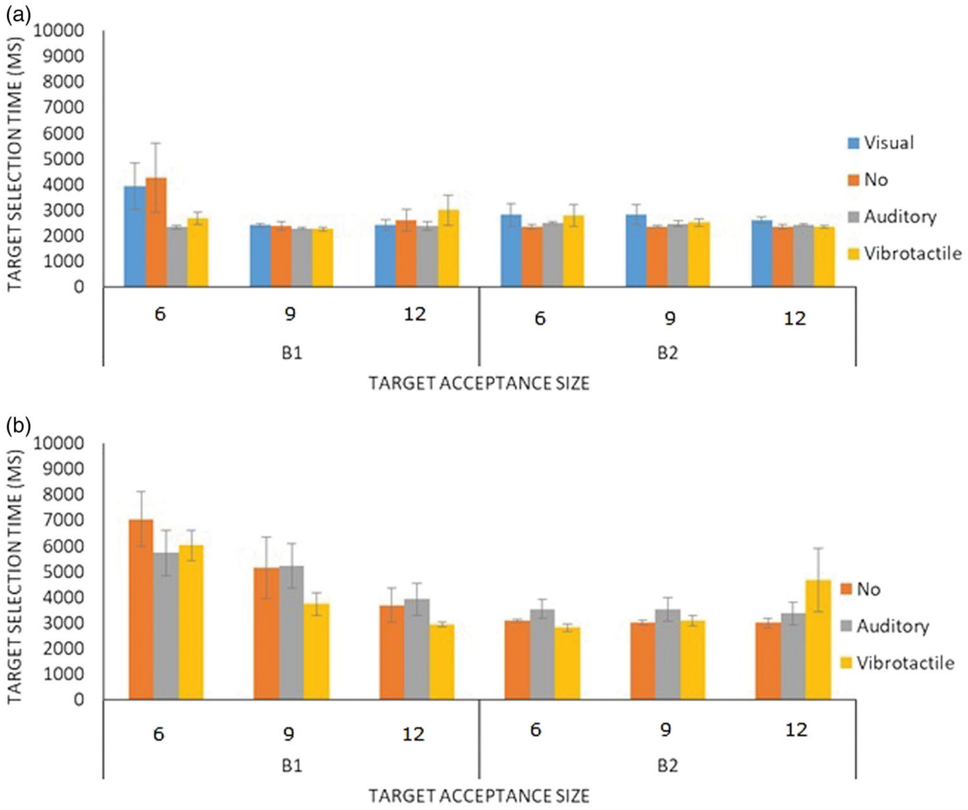

Figure 7. Target selection time with the different target acceptance sizes and feedback modalities for the two adult participants with physical impairments for (a) the on-screen experiment and (b) for the off-screen experiment.

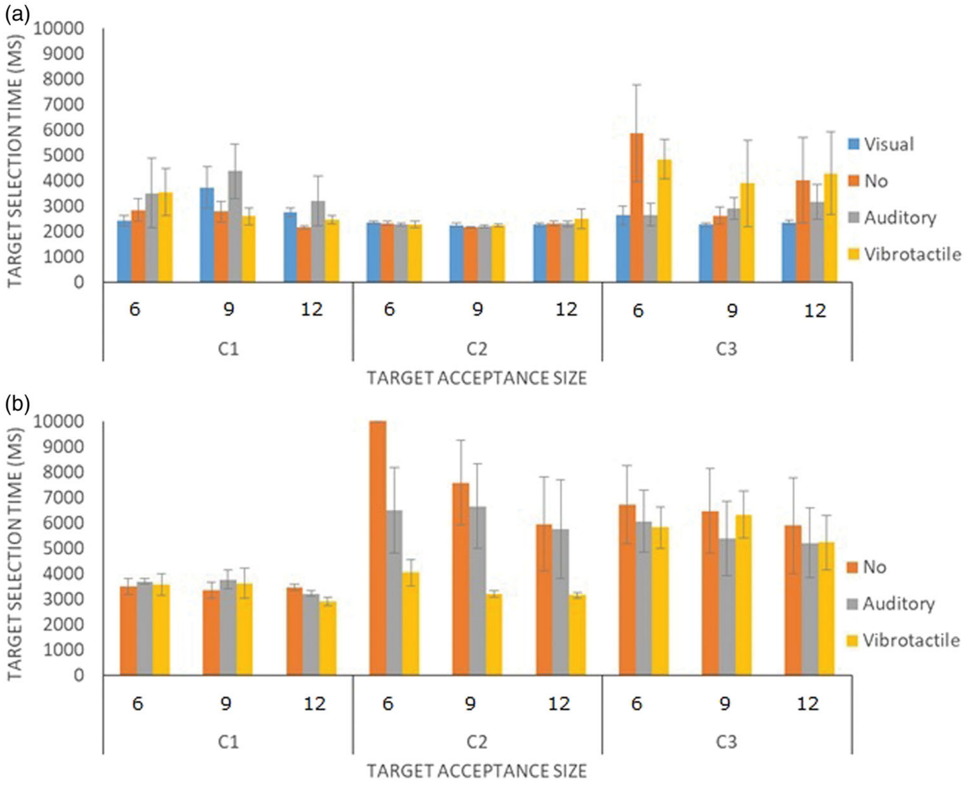

Figure 8. The target selection time with the different target acceptance sizes and feedback modalities for the three non-disabled child participants for (a) the on-screen experiment and (b) the off-screen experiment.

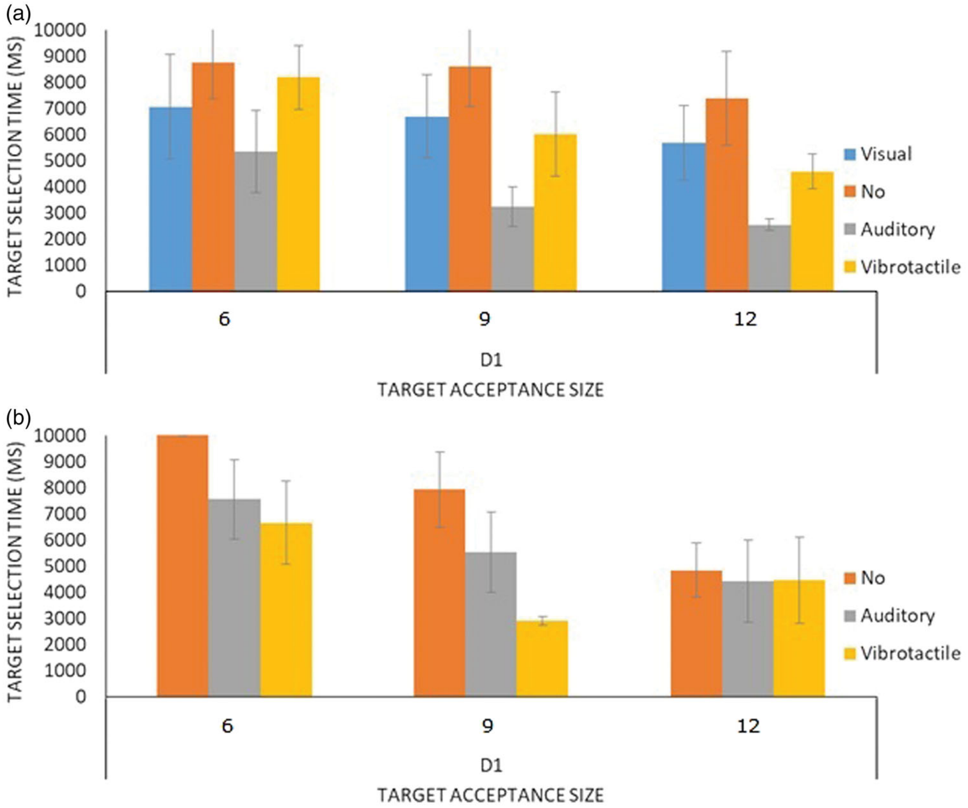

Figure 9. The target selection time with the different target acceptance sizes and feedback modalities for the child participant with physical impairments for (a) the on-screen experiment and (b) the off-screen experiment.

screen condition. The no-feedback modality was clearly longer with the target acceptance size of 6 cm in both the on- and off-screen conditions. In general, the target selection time in the off-screen condition took longer than in the on-screen condition. The average target selection time increased by 12.1% for C1, 61.4% for C2, and 40.9% for C3 from the on-screen to the off-screen conditions.

Child participant with physical impairments

Figure 9(a,b) indicates the target selection time for the on- and off-screen conditions of the child participant who had a physical impairment, D1. The figures show that no-feedback had the longest target selection time in all the target size and all the screen conditions. The auditory feedback had the shortest selection times for the on-screen condition, but vibrotactile feedback had the shortest times for the off-screen condition. The target selection time increases as the target acceptance size gets smaller in both screen conditions. In terms of the difference in the screen conditions, the average target selection time in the off-screen condition was 0.6% longer than in the on-screen condition. However, the average target selection time for the smallest target acceptance size was 9.1% longer.

Timeout error rate

The timeout error rate for each participant group in the screen conditions are shown in Table 1. The timeout error only occurred with the target acceptance size of 3 cm in both screen conditions for the non-disabled adult participants. For the adult participants with physical impairments, a timeout error was rarely seen in any of the conditions. The non-disabled child participants performed well in the on-screen condition. However, from 20 to 46% timeout error rates were observed in the off-screen condition. The highest timeout error rate among all the participants occurred for the child participant with physical impairment, D1; the error rate for D1, especially with the 6 cm target acceptance size, was remarkably higher than the timeout error of other groups. The table also indicates that the no-feedback modality had the highest timeout error rates in both screen conditions. The auditory feedback had the lowest time out rate in the on-screen condition, and the vibrotactile feedback had the lowest time out rate in the off-screen condition, overall. In general, the time out error rate in the off-screen condition was higher than the on-screen condition for all the participant groups.

Participant preferences

Table 2 shows the preference among the feedback modalities for all the participants. For the non-disabled adult participants, visual feedback was the preferred modality for the on-screen condition, and auditory feedback was the preferred modality for the off-screen condition. The no-feedback condition stands out as the least preferred feedback modality for all the screen conditions. The preferred feedback modality for the other participants was distributed quite evenly for the on-screen condition. However, in the off-screen condition, the most preferred modality was the auditory feedback. No-feedback was the least preferred feedback modality of all the participants for both conditions, except C1 and C3 for the on-screen.

The participants who preferred the visual feedback commented that the visual feedback was most intuitive because they could easily see where their gaze was being tracked by the system. However, C1 and C3 who chose the visual feedback as the least preferred feedback pointed out that the visual feedback was

Table 1. Time-out error rate (%) for the different target acceptance size and feedback modality in each population group for on-screen and off-screen experiment.

	Size	Visual feedback	No-feedback	Auditory feedback	Vibrotactile feedback
On-screen experiment					
Non-disabled adult participants	3	3.3	12.1	8.5	10.2
	6	0.0	0.0	0.0	0.0
	9	0.0	0.0	0.0	0.0
Adult participants with physical impairments	6	11.1	0.0	0.0	0.0
	9	0.0	0.0	0.0	0.0
	12	0.0	0.0	0.0	0.0
Non-disabled child participants	6	0.0	14.3	0.0	0.0
	9	0.0	0.0	0.0	6.7
	12	0.0	7.2	0.0	7.1
Child participant with physical impairments	6	60.0	80.0	20.0	40.0
	9	20.0	80.0	0.0	20.0
	12	20.0	60.0	0.0	0.0

	Size	No-feedback	Auditory feedback	Vibrotactile feedback
Off-screen experiment				
Non-disabled adult participants	3	37.9	35.0	37.9
	6	0.0	0.0	0.0
	9	0.0	0.0	0.0
Adult participants with physical impairments	6	11.1	0.0	0.0
	9	0.0	0.0	0.0
	12	0.0	0.0	0.0
Non-disabled child participants	6	46.7	20.0	0.0
	9	38.5	20.0	0.0
	12	26.7	20.0	0.0
Child participant with physical impairments	6	100.0	40.0	40.0
	9	60.0	20.0	0.0
	12	0.0	20.0	0.0

$n = 10$ for the non-disabled adults, $n = 2$ for the adults with physical impairment, $n = 3$ for the non-disabled children, and $n = 1$ for the child with physical impairment.

Table 2. The preferences of the participants for the feedback modalities in the on- and off-screen conditions.

	Most preferred feedback modality		Least preferred feedback modality	
	On-screen	Off-screen	On-screen	Off-scree
Non-disabled adult participants	Visual (4/10)	Auditory (6/10)	No-feedback (8/10)	No-feedback (9/10)
B1	Visual	Auditory	No-feedback	No-feedback
B2	Visual	Auditory	No-feedback	No-feedback
C1	Auditory	Vibrotactile	Visual	No-feedback
C2	Auditory	Auditory	No-feedback	No-feedback
C3	Vibrotactile	Auditory	Visual	No-feedback
D1	Vibrotactile	Vibrotactile	No-feedback	No-feedback

$n = 10$ for the non-disabled adult participants.

distracting when the location of pointer did not exactly match with the actual location of their gaze. The participants who ranked either the auditory feedback or the vibrotactile feedback as the most preferred modality liked how they knew how long to fixate their gaze on the target based on the intensity of the feedback provided. Another participant commented that the auditory and vibrotactile feedback modalities were intuitive but might take more time to get used to. Also, some participants commented that they preferred the auditory over the vibrotactile feedback because they liked the ramp-up sound that was given during the gaze fixation because it was more noticeable than the ramp-up vibration. One participant commented that an advantage of those feedback modalities was less eyestrain compared with the visual feedback.

Discussion

Overall, participants were slower and less successful selecting targets using gaze fixation in off-screen interactions with the physical environment than in on-screen interactions. The smaller the target acceptance size, the slower the participant to complete the target selection task in both the on- and off-screen conditions. The tasks performed with feedback modalities were accomplished faster and more accurately than tasks performed without feedback, and similar results were observed in both screen conditions. However, the choice for which feedback modality to use going forward might be the user's preference. Some of the participants with physical impairments had difficulty performing the eye gaze in the tasks because they could not keep their head position still during the gaze interaction, but providing feedback and increasing the target acceptance size appeared to help them to improve the speed of the gaze interaction task.

The longer target selection time and higher timeout rate in the off-screen condition is likely because the targets in the off-screen condition were placed on the surface in a horizontal plane. A small difference of gaze movement in a vertical angle affected the accuracy of the gaze interaction with the task environment, especially for gazing at the target that was far from the participant. However, the timeout rarely happened with the larger target acceptance sizes in the on-screen as well as the off-screen condition if any feedback was provided.

Statistically significant differences in the target selection time were found only in the 3 cm target acceptance size in both on- and off-screen conditions for the non-disabled adults.

Interestingly, even though the target selection time with the smallest target acceptance size was quite different from the other two larger target acceptance sizes, performance with these two larger sizes was nearly the same. This is probably because the target selection time only increased when the degree of task difficulty exceeded what the user could handle.

The *post hoc* test revealed that the selection time in the no-feedback condition was significantly longer than the other feedback modalities. Thus, any feedback provided to the participants helped them to perform the most difficult target selection task. No significant difference between the audio, the vibrotactile, and the visual feedback was found in either the on or off-screen conditions. Therefore, the three feedback modalities were similar in their effectiveness. Regarding the timeout error with respect to each feedback modality, no-feedback had the highest error rate in both screen conditions, while the auditory feedback was lowest in the on-screen condition, and vibrotactile feedback was lowest in the off-screen condition. Other researchers have also found that it was not possible to quantitatively identify a clear “optimal” feedback, for example, Rantala et al. [17] found that feedback improved user’s performance in gaze interaction on-screen, but all the modalities generally performed equally.

Looking at the qualitative experience of the participants, according to the questionnaire, the visual feedback was the most preferred feedback in the on-screen condition, auditory feedback was the most preferred feedback in the off-screen condition, and the no-feedback condition was the least preferred among all the feedback modalities. Even though the visual feedback was most often ranked as the best for the on-screen condition in adults, two children participants ranked it as the least preferred feedback. The problem with the disparity between the location of the gaze-based mouse pointer and the location they were gazing probably lowered their preference, whereas the disparity didn’t seem to affect the preference for the adults. This type of disparity is one of the common issues for on-screen gaze applications [26]. The comments about preferring the auditory over the vibrotactile feedback because auditory had the distinct ramp-up sound could be addressed for the vibrotactile feedback. The vibration amplitude is adjustable, so if the vibration amplitude was larger, more participants might have ranked it higher. Even though it was not most preferred, six out of sixteen participants did prefer it in the off-screen condition, and it could still have potential. As found in the review of Burke et al. [14] visual-auditory feedback was most effective when a single task is being performed under normal workload conditions, which was the case in this study, but visual-vibrotactile feedback was more effective for multiple tasks requiring high workload conditions, which would be the case when selecting targets for robot interaction in future studies. Auditory feedback or vibrotactile feedback could be usable for off-screen gaze interaction, with the choice depending on individual preference.

Next steps will involve using toys as target objects in the play environment. Using the homogeneous transformation technique, a robot system could determine which toy a child is interested in interacting with his or her eye gaze. The child will have feedback that the system knows he is looking at something because of the feedback. If using auditory feedback, the actual target chosen can be spoken aloud. If using vibration feedback, each target could correspond to a different vibration frequency. After the correct target is selected, robot movement towards the toy can be controlled by another input method or by autonomous programming. This low-cost eye gaze technique could also be used in other situations where it is beneficial not to have to switch attention between a computer screen with options and the environment, for example, power wheelchair direction selection or remote-controlled toys.

Conclusions

This study showed that the gaze interaction in an off-screen condition could be performed with a stationary eye tracker using the homogeneous transformation technique. The participants required more time to interact and select the target object in the physical off-screen environment than the target in the on-screen condition. The participant’s performance in the target selection tasks varied depending on the age and the impairment with selection time generally being slower for younger children and physically impaired participants. However, they performed the target selection tasks in both conditions comparatively accurately and quickly if the size of the target was not too small for the participants to sustain their gaze upon. The results also indicated that providing feedback to inform where gaze is fixated could make the gaze interaction performance faster and more accurate in both screen conditions. However, none of the feedback modalities emerged as performing better than the others. With future development, this eye gaze system and feedback modalities will be integrated with an assistive robot platform and used for play activities in the real physical world. This will contribute towards the goal of enabling children with physical impairment opportunities to perform object manipulation in the physical environment.

Acknowledgement

We would like to acknowledge Archit Shah for his assistance on this project.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This research was supported by a Collaborative Health Research Project (CHRP), a joint initiative of the National Sciences and Engineering Research Council (NSERC) and Canadian Institutes of Health Research (CIHR), and the Glenrose-Rehabilitation-Hospital-Foundation.

References

[1] Gibson EJ. Exploratory behavior in the development of perceiving, acting, and the acquiring of knowledge. Annu Rev Psychol. 1988;39(1):1–42.

[2] Musselwhite CR. Adaptive play for special needs children: strategies to enhance communication and learning. San Diego (CA): College-Hill Press; 1986.

[3] Bjorklund DF. Children’s thinking: cognitive development and individual differences: Australia. 4th ed. Belmont (CA): Thomson/Wadsworth; 2005.

[4] Piaget J. Play, dreams, and imitation in childhood. New York: Norton; 1962.

[5] Taylor NF, Imms C, Dodd K. Physiotherapy and occupational therapy for people with cerebral palsy: a problem-

based approach to assessment and management. London: Mac Keith Press; 2010.

[6] Adams K, Rios A, Alvarez L. Robotic systems for augmentative manipulation to promote cognitive development, play, and education. Boca Raton (FL): CRC Press; 2017.

[7] Cook AM, Polgar JM. Cook & Hussey's assistive technologies: principles and practice. 3rd ed. Philadelphia (PA): Elsevier; 2008.

[8] Biswas P, Langdon P. A new input system for disabled users involving eye gaze tracker and scanning interface. J Assist Technol. 2011;5(2):58–66.

[9] Arai K, Yajima K. Robot arm utilized having meal support system based on computer input by human eyes only. IJHCI. 2011;2(1):120–128.

[10] Encarnação P, Leite T, Nunes C, et al. Using assistive robots to promote inclusive education. Disab Rehab Assist Technol. 2017;12(4):352–372.

[11] Ryan WJ, Duchowski AT, Birchfield ST. Limbus/pupil switching for wearable eye tracking under variable lighting conditions. Paper presented at the Proceedings of the 2008 symposium on Eye tracking research applications no. 38; Savannah, Georgia; 2008.

[12] Land MF, Hayhoe M. In what ways do eye movements contribute to everyday activities?. Vision Res. 2001; 41(25–26):3559–3565.

[13] Galante A, Menezes P. A gaze-based interaction system for people with cerebral palsy. Procedia Technol. 2012;5: 895–902.

[14] Burke JL, Prewett MS, Gray AA, et al. Comparing the effects of visual-auditory and visual-tactile feedback on user performance: a meta-analysis. Proceedings of the 8th international conference on Multimodal interfaces; 2006. p. 108–117.

[15] Majaranta P, MacKenzie S, Aula A, et al. Effects of feedback and dwell time on eye typing speed and accuracy. Univ Access Inf Soc. 2006;5(2):199–208.

[16] Pomper U, Chait M. The impact of visual gaze direction on auditory object tracking. Sci Rep. 2017;7(1):1–16.

[17] Rantala J, Majaranta P, Kangas J, et al. Gaze interaction with vibrotactile feedback: review and design guidelines. Human Comput Inter. 2017;35(1):1–39.

[18] Boyer EO, Portron A, Bevilacqua F, et al. Continuous auditory feedback of eye movements: an exploratory study toward improving oculomotor control. Front Neurosci. 2017;11:197–197.

[19] Kangas J, Akkil D, Rantala J, et al. Gaze gestures and haptic feedback in mobile devices. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; 2014. p. 435–438.

[20] Amantis R, Corradi F, Maria Molteni A, et al. Eye-tracking assistive technology: Is this effective for the developmental age? Evaluation of eye-tracking systems for children and adolescents with cerebral palsy. Asst Technol Res Ser. 2011; 29:489–496.

[21] Dhas BN, Samuel PS, Manigandan C. Use of computer access technology as an alternative to writing for a preschool child with athetoid cerebral palsy—a case report. Occupat Ther Health Care. 2014;28(3):318–332.

[22] Møllenbach E, Hansen JP, Lillholm M. Eye movements in gaze interaction. J Eye Mov Res. 2013;6(2):1–15.

[23] Bednarik R, Gowases T, Tukiainen M. Gaze interaction enhances problem solving: Effects of dwell-time based, gaze-augmented, and mouse interaction on problem-solving strategies and user experience. J Eye Mov Res. 2009;3: 1–10.

[24] Munoz DP, Armstrong IT, Hampton KA, et al. Altered control of visual fixation and saccadic eye movements in attention-deficit hyperactivity disorder. J Neurophysiol. 2003; 90(1):503–514.

[25] Szeliski R. Computer vision. [electronic resource]: algorithms and applications. New York; London: Springer; 2011.

[26] Zhang Y, Hornof AJ. Mode-of-disparities error correction of eye-tracking data. Behav Res. 2011;43(3):834–842.