Disability and Rehabilitation: Assistive Technology

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/iidt20

The ability of typically developing 2–3 year olds to infer the control mechanism for eye-gaze technology and the impact of causal language instruction

Tom Griffiths, Michael T. Clarke & John Swettenham

To cite this article: Tom Griffiths, Michael T. Clarke & John Swettenham (19 Dec 2023): The ability of typically developing 2–3 year olds to infer the control mechanism for eye-gaze technology and the impact of causal language instruction, Disability and Rehabilitation: Assistive Technology, DOI: 10.1080/17483107.2023.2293874

To link to this article: https://doi.org/10.1080/17483107.2023.2293874

© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

Published online: 19 Dec 2023.

Submit your article to this journal

Article views: 163

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=iidt20

https://doi.org/10.1080/17483107.2023.2293874

Taylor & Francis
Taylor & Francis Group

ORIGINAL RESEARCH

OPEN ACCESS   Check for updates

The ability of typically developing 2–3 year olds to infer the control mechanism for eye-gaze technology and the impact of causal language instruction

Tom Griffiths[a], Michael T. Clarke[b] and John Swettenham[c]

[a]School of Computing, University of Dundee, Dundee, UK; [b]Department of Speech, Language and Hearing Sciences, San Francisco State University, San Francisco, CA, USA; [c]Division of Psychology and Language Sciences, University College London (UCL), London, UK

ABSTRACT

Purpose: Little is known about how children learn to control eye-gaze technology, and clinicians lack information to guide decision-making. This paper examines whether typically developing 2–3 year olds can infer for themselves the causal mechanisms by which eye-gaze technology is controlled, whether a teaching intervention based on causal language improves performance and how their performance compares to the same task accessed via a touchscreen.

Methods and materials: Typically developing children's (n = 9, Mean Age 28.7 months) performance on a cause and effect game presented on eye-gaze and touchscreen devices was compared. The game was presented first with no specific instruction on how to control the devices. This was followed by a subsequent presentation with explicit instruction about how the access methods worked, using a causal language approach. A final presentation examined whether children had retained any learning.

Results: Performance in the eye-gaze condition without instruction (42.5% successful trials) was significantly below performance in the corresponding touchscreen condition (75%). However, when causal language instruction was added, performance with both access methods rose to comparable levels (90.7% eye-gaze and 94.6% touchscreen success). Performance gains were not retained post-intervention.

Conclusions: Although 2–3 years in the study could make use of eye-gaze technology with support, this study found no evidence that these children could infer the causal mechanisms of control independently or intuitively. The lack of spatial contiguity and the comparative lack of feedback from eye-gaze devices are discussed as possible contributory factors.

ARTICLE HISTORY
Received 17 May 2023
Revised 28 November 2023
Accepted 6 December 2023

KEYWORDS
Cerebral palsy; computer access; eye-gaze technology; cause and effect; teaching; assistive technology

IMPLICATIONS FOR REHABILITATION

- There are challenges in young children inferring for themselves the causal link between eye movements and control of an eye-gaze device.
- Explicit instruction may improve children's performance in a specific task, but it is debatable whether this translates to the establishment of causal mechanisms for control of the device.
- Clinicians should be cautious of making assumptions about what children are learning from activities claiming to teach cause and effect or other foundational eye-gaze control skills.

Introduction

Eye-gaze technology refers to a method of interacting with a computer using the controlled movement and fixation of the user's gaze. Eye-gaze systems typically involve a specialised camera, sensitive to infra-red and near-infra-red light, mounted beneath the screen of a computer or tablet, with one or more infra-red emitters located coaxially. By calculating the relative distance and offset between the centre of the user's pupil and the reflection of the emitter(s), image processing software determines where the user is looking on the screen, which is in turn translated into cursor movement [1]. Onscreen items can be selected by holding the gaze steadily for a pre-determined period (known as "dwell selection"). This technology can potentially offer a means of computer access for people with physical disabilities, whose ability to interact with a computer using volitional movement of limbs or other body parts may be limited. Such access can enable the use of electronic augmentative and alternative communication (AAC) systems, as well as control of standard and customised computer software, environmental control systems and other assistive technology. This, in turn, offers the potential to increase users' participation and involvement in social interaction, education, employment, play and leisure [2, 3].

Despite its obvious potential to support children with physical disabilities, and the many significant achievements that children can experience with eye-gaze, successive reviews have highlighted that use of eye-gaze technology is under-researched [1, 4]. Karlsson and colleagues [1], for example, conducted a systematic review of the literature related to the effectiveness of eye-gaze technology in supporting communication, identifying only two papers as meeting the evidence criteria for inclusion, both of which provided low levels of evidence for use of the technology from small group trials. The review's authors observe that only one study [5] involved children and that the literature generally lacks evidence relating to the impacts of cognitive or intellectual disability or assessment strategies used to guide implementation

CONTACT Tom Griffiths TGriffiths001@dundee.ac.uk School of Computing, University of Dundee, Dundee, UK
© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

of the technology. Several researchers [1, 5, 6] have highlighted that there exists little evidence to guide early interventions using eye-gaze technology and that, in practice, families, clinicians and educators can have limited insight into the learning demands of this form of access which can lead to children being given technology that they may be unable to use effectively [7]. The introduction of eye-gaze technology can also be accompanied by high expectations on all sides which, in turn, can lead to persistence with its use when it may not be the most effective access method or when greater success and satisfaction could be experienced in other activities [8]. Despite these concerns, we know little about how children first learn to control eye-gaze technology, how they acquire the understanding that their eyes are controlling the device.

Contiguity and the causal mechanism

Common methods of controlling a computer, such as a mouse, joystick or mechanical switches, involve both a physical input device and a variety of clear feedback channels. These access methods often provide some auditory and/or haptic feedback, such as the auditory "click" of a mechanical switch or the physical sensation of "travel" when depressing a keyboard key, alongside several types of visual feedback – the user can see their body operating the input device and also see the resulting action or selection on a device's display. It has been well established that, for young children, relationships between objects and actions are better inferred as causal if they adhere to the principles of temporal and spatial contiguity; that is, for one event to be perceived as causing another, it should occur concurrently and at the same location [9–12]. Where spatial or temporal contiguity is absent or less obvious, it is more difficult to discern an underlying mechanism by which one event causes another [13].

Teaching eye-gaze control

It has been argued that where a causal relationship is not clear and obvious it can be learnt by young children through observation or explanation, imitation and repetition [14–16]. When supporting children learning to use a touchscreen or switch, modelling is an established principle [17]. One may, for example, demonstrate the relationship between the switch and the device by positioning the switch where the child can see it and pressing the switch oneself. This may then proceed to supporting a child through hand-over-hand explorations of the switch, which are then phased out as the child consolidates the relationship between the movement of their body, the feedback from the switch and the resulting action on the display. With eye-gaze control, however, the relationship cannot be demonstrated so easily; there is no practical way to physically affect the direction of a child's gaze and the movements of the eyes are generally too small to be useful in demonstrating how they might be used for control.

Verbal instruction is a method commonly used to support the development of causal links between events. Bonawitz and colleagues [11] looked at the impact of causal language on children's understanding and replication of causal actions. The researchers describe causal language as being simplified so as to be accessible to young children. Using core verbs and simplified, consistent grammatical forms ("The block makes the truck *go*" to be followed by "Can you make the truck *go*?") may help to embed the causal links between actions. The researchers tested typically developing toddlers' ($n=18$, Mean Age 24.4 months) ability to infer causal relationships by showing them causal events with no supporting description, where the action of a moving block contacting a toy caused the toy to activate. They found that, although the children could recognise causal relationships between the objects (as measured by observing their predictive looking – shifting gaze in anticipation of an object's action), none of the children spontaneously performed the necessary action to replicate these events in structured or free play. In a further task with a similar group ($n=20$, Mean Age 24.5 months), the researchers showed that the addition of causal language resulted in a significant increase in children spontaneously replicating the action to activate the toy: with successful replication of the event in 50% and 62% of trials in two causal language conditions, compared to only 6% in non-causal trials. Further, it was observed that causal language resulted in better performance than simply calling children's attention to the events, suggesting that this method of teaching may support the development of causal reasoning [11]. Bonawitz and colleagues propose that their results support the use of causal language as a method for teaching causal links which lack a clear and obvious link between agent and action, or where the mechanism is unclear.

The current study worked with typically developing children with no prior experience using eye-gaze. Having typically developing children as participants allowed exploration of their learning from a common baseline [18]. Working with typically developing children also allowed for comparison with touchscreen performance. Since this is the first attempt we are aware of to explore the application of causal language instruction to eye-gaze technology, children aged between two and three were selected as participants. This younger age group was chosen firstly because, if these skills can be demonstrated in younger children then it is likely that they will also be present in older children, and secondly because this age group would have language comprehension sufficient to understand the causal language instruction. The study looked at whether children could learn through feedback from the eye-gaze device, that fixating on one stimulus activates a reward while fixating on an alternative similar stimulus does not. The study addressed the following research questions: firstly, whether typically developing children of pre-school age could infer for themselves the causal mechanisms by which eye-gaze technology is controlled, secondly whether a teaching intervention based on causal language could improve their performance and finally how their performance compares to the same task accessed *via* a touchscreen.

Methods

Study design

This study used a repeated measures, two-by-two design, with children given the same task on both a touchscreen and an eye-gaze control device. In each condition, children were presented with the task first with no additional support, then with additional teaching and prompting using causal language within an intervention phase, before a final post-intervention phase with no additional support. Ethical approval for this study was granted by University College London Research Ethics Committee (Project ID 1328/009).

Participants and recruitment

Typically developing children ($n=9$; 3 male, 6 female) aged between 24 and 35 months (Mean Age = 28.7, $SD=4.8$) were recruited from a pre-school in the UK. Informed consent was obtained from parents of all children recruited. All children met the core inclusion criteria of having no diagnosed learning disability, understanding English to a level sufficient to understand the instructions (as reported by staff), and having no reported hearing or vision difficulties (other than refractive errors corrected by glasses) that would impact on their ability to engage with the research materials.

All children demonstrated cause and effect understanding with physical objects, which was affirmed through a game played at the start of the session in which the experimenter models the use of the toy, pushing a button to make it move across the table and accompanying this with a causal language observation such as "The button makes it go." The toy is then passed to the child, with the prompt "Can you make it go?", followed by encouragement to push the button and activate the toy [19].

Procedures

Eye-gaze and touchscreen setup

All children were tested in a separate, familiar room in their pre-school. Distractions were minimised and a familiar adult accompanied each child. Adults were asked not to give the child any encouragement or prompts that might impact on their performance. An eye-gaze control device (the *Mobi 2* from *Jabbla* with a *Tobii PCEye Go* camera) was positioned so that children would be seated at the distance from the screen specified in the manufacturer's usage instructions – approximately 60 cm away with the eyeline in the top third of the screen area. The *Mobi 2* runs the *Mind Express 4* software, which was used to write the experiments described in this paper. As the *Mobi 2* also has a touchscreen, the same device was used for both eye-gaze and touchscreen conditions. The touchscreen was deactivated during the eye-gaze condition.

All children successfully completed a five-point calibration of the eye-gaze system prior to undertaking the protocol. A crosshair cursor was used to provide feedback on the child's gaze position during the eye-gaze condition. A dwell selection time of 1.0 s was set for all parts of the task. The need to dwell on the item being selected reduces the possibility that the choice can occur accidentally. Progress towards a selection was indicated by a prominent red border appearing around the item being targeted and a "clock"-style indicator which appeared around the crosshair, with a selection being made when the circle had completed.

The touchscreen condition was configured to closely match the eye-gaze setup, with the same red border and clock-style indicator used to show children's progress towards a selection. In the touchscreen condition, a 1.0 s hold on an item was required to select it. This delay was introduced to replicate the delay in activation that occurs with dwell selection, meaning that control of the touchscreen would require some element of learning.

Grouping

Children were randomly assigned to complete either the touchscreen or eye-gaze condition first. The procedure for both access methods was the same, as shown in Figure 1. The full procedure for each condition lasted around 20 min, with a five minute break between the conditions.

Learning phase

Two learning phases were presented to all children. Initial instructions to the children were kept to a minimum, with children told only that they were going to play some finding games on a computer. The first of these introduced the "active" stimulus – a still image of a woman wearing a blue t-shirt with an effective area of 65×50 mm (equivalent to $20° \times 17°$ of visual angle at 60 cm distance in the eye-gaze condition). The image appeared randomly at one of nine onscreen locations (see Figure 2(a)). When children dwelled on or pressed the image for the required time, it was replaced by one of six different reward video clips which appeared at the same location and with the same dimensions as the initial picture. In all cases, the reward clips featured the woman in the original image performing an action such as dancing, waving, blowing bubbles or playing with a toy. The colocation of the stimulus and reward was intended to underline the relationship between the two.

The second learning phase introduced the "inactive" stimulus – a still image of a man in a yellow t-shirt of the same size as the active stimulus (Figure 2(b)). In contrast to the active stimulus, dwelling on or pressing the image did not lead to the video reward and it remained on screen for six seconds, before disappearing and reappearing at another location after a delay of two seconds. Both the active and inactive learning phases presented the stimuli six times each.

Baseline testing phase

Following the completion of both learning phases, children were offered a five minute break and then presented with the baseline testing phase. In this phase, both active and inactive stimuli appeared simultaneously onscreen (Figure 2(c)). Twelve iterations of this testing phase were presented to each child for each of the access methods. Children's performance on each trial was categorised as follows:

1. Successful – a successful trial is one in which the active stimulus is selected
2. Unsuccessful – selection of inactive stimulus
3. Unsuccessful – selection of another (blank) area of the screen
4. Unsuccessful – no attempt at selection made within 10 seconds
5. Unsuccessful – attempt to touch the screen (applicable only for the eye-gaze condition)

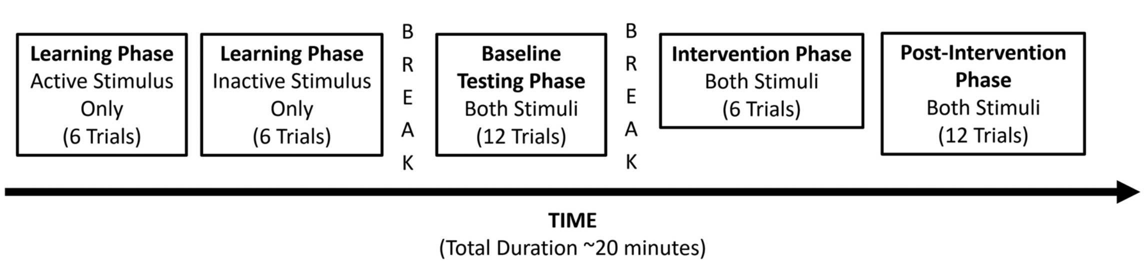

Figure 1. Procedure for both eye-gaze and touchscreen conditions.

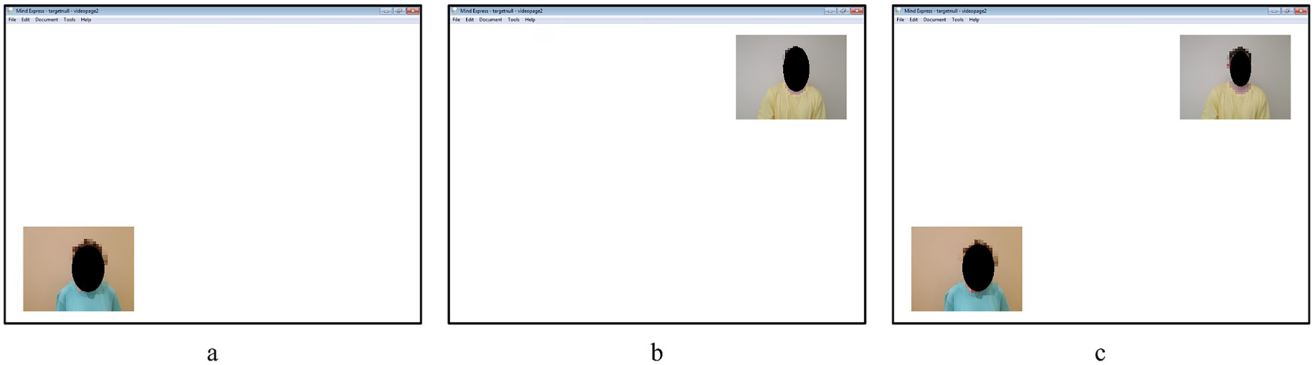

Figure 2. The three experimental phases: (a) learning phase with active stimulus, (b) learning phase with inactive stimulus, (c) simultaneous presentation of stimuli used in all other phases. Note that faces in this image are obscured for confidentiality, however faces in the real stimuli were not obscured.

If no selection was made within ten seconds, the experimenter manually advanced to the next trial and attempted to redirect the child's attention to engaging with the task. All sessions were recorded by a video camera which was focused on the screen of the device to allow for score checking.

Only minimal verbal prompting was given during the above phases, with the instructor only using prompts to direct the child's attention to the task and non-specific praise (such as "That's great!") given to keep the child engaged. This ensured that children did not receive any information or feedback about how to use the access method or complete the task. For the touchscreen condition, children were encouraged to press the screen harder if they did not achieve enough pressure to activate the active icon. They were not given information about the need to hold their finger on the screen until the dwell had completed.

Intervention and post-intervention phases

The intervention phase comprised a further six presentations featuring both active and inactive stimuli simultaneously (Figure 2(c)). During this phase, children were given specific verbal instructions using the principles of causal language and were provided with coaching and feedback on their performance. In the eye-gaze condition, children were told the following:

- That the goal was to look at the reward stimulus until it played a video (e.g. "You need to look at the woman to make the video play")
- That the non-reward stimulus would not do anything when looked at (e.g. "Nothing happens when you look at the man")
- That the cursor on the screen was being controlled by their eyes (e.g. "The red box shows where you are looking")
- That the dwell progress marker must complete in order for the video to play (e.g. "You need to look until the circle goes all the way round")

In the touchscreen condition, children were told the following:

- That the goal was to press the reward stimulus (e.g. "You need to press the woman to make the video play")
- That the non-reward stimulus would do nothing when pressed (e.g. "Nothing happens when you press the man")
- That the screen needed to be pressed until the dwell progress marker had completed for the video to play (e.g. "You need to press until the circle goes all the way round")

In both conditions, children were given feedback which explicitly referenced what they had done to trigger the video reward, for example: "Well done! You played the video by looking at the woman!", or which made explicit any errors in selection, for example: "Oh dear! Looking at the man doesn't play a video!." In order to keep the protocols for the two access methods as similar as possible, no modelling (such as pointing to the icons or simulating the pressing of the screen) was used during the touchscreen intervention. After the intervention phase, children were given a break of up to one minute, during which they were allowed to play with a toy or chat with the researcher or a familiar adult.

In the post-intervention phase of the trial the active and inactive stimuli appeared on screen simultaneously for a further twelve trials (Figure 2(c)). Support reverted to the minimal verbal prompting of earlier phases. The scoring system and the coding of unsuccessful trials for the intervention and post-intervention phases was the same as that used for the learning and baseline phases.

At the end of the experiments, children were given a sticker as a reward for participating and were thanked for taking part.

Results

Children using touchscreen first (Group A) and eye-gaze first (Group B) were matched for chronological age ($t(7)$ = .043, $p = 0.967$). A mixed ANOVA confirmed that there was no difference in performance related to the order in which the two access methods were presented ($F(1, 7)$ = .26, $p = 0.628$).

Performance on touchscreen and eye-gaze

Table 1 reports all eye-gaze and touchscreen responses across baseline, intervention and post-intervention phases. For the eye-gaze conditions we report the mean percentage of trials in which a successful selection was made (1 s fixation on the active stimulus). We also report the mean percentage of trials where the child fixated for 1 s on the inactive stimulus, fixated for 1 s on a blank area of the screen, made no fixation on the screen area, or touched the screen. The table reports the equivalent mean percentages for the touchscreen conditions.

A two-way repeated measures ANOVA was conducted with two within-subject factors: testing phase (baseline or post-intervention) and access method (touchscreen or eye-gaze). There was no significant main effect of testing phase ($F(1, 8)$ = .49, $p = 0.505$). A significant main effect of access method was observed ($F(1, 8) = 8.42$, $p = 0.002$, ηp^2 = .51) with the partial eta

Table 1. Mean percentage performance on eye-control and touchscreen for each phase.

	Baseline phase	Intervention phase**	Post-intervention phase
Eye-gaze			
Fixation on Active Stimulus	42.5%	90.7%	35.8%
Fixation on Inactive Stimulus	21.7%	–	30.8%
Fixation on Blank Area	8.3%	–	6.7%
No Selection	5.8%	–	10%
Touching Screen*	21.7%	–	16.7%
Touchscreen			
Touching Active Stimulus	75%	94.6%	74.2%
Touching Inactive Stimulus	20%	–	12.5%
Touching Blank Area	2.5%	–	5%
No Selection	2.5%	–	8.3%

*Note that this error type only exists in the eye-gaze condition.
**Unsuccessful trial outcomes were not recorded for the intervention phase, since children averaged less than one error each on either access method during this part of the experiment.

squared indicating a large effect size. There was no interaction between testing phase and access method ($F(1, 8) = .36$, $p=0.563$).

No significant correlations were seen between age and performance on the touchscreen task ($r=-0.002$, $p=0.997$), age and performance on the eye-gaze task ($r = .12$, $p=0.754$), or age and overall performance ($r = .75$, $p=0.125$).

Impact of causal language intervention

The greatest success for both access methods was seen in the intervention phase, where the majority of trials were successful: 94.6% of touchscreen trials and 90.7% of eye-gaze trials resulted in selection of the active stimulus. A further ANOVA calculation for the eye-gaze condition, comparing baseline and intervention phases indicated that there was a significant increase ($F(2,16) = 30.2$, $p = < 0.001$) in mean percentage of correct responses using the eye-gaze system in the intervention ($M_{success} = 42.6\%$, $SD=27.8$) as opposed to the baseline phase ($M_{success} = 90.7\%$, $SD=16.9$), a mean increase of 48.2% (95% CI, 24.5 – 71.8%).

Discussion

This study looked at whether nine typically developing children aged 24–35 months could infer for themselves the causal mechanisms by which eye-gaze technology is controlled, whether a teaching intervention could improve their performance and whether there was a difference in performance on the same task accessed *via* a touchscreen.

The children in this study achieved an average of 42.6% successful trials on the baseline phase with eye-gaze, compared to 75% successful on the equivalent touchscreen phase. There was a significant difference in performance between the two access methods, which would suggest that the eye-gaze access method was more challenging than the equivalent activity presented on a touchscreen. All children were of a developmental age where they would be expected to understand cause and effect relationships [20], and all children had demonstrated understanding of cause and effect using physical objects prior to the start of the trials, ruling out any general lack of understanding of causal relationships potentially accounting for any difference. Similarly, age was not a factor impacting performance in either condition – it was not the case that older children in the group necessarily performed better on the eye-gaze task.

Selection of the inactive stimulus accounted for the largest proportion of errors in both the baseline and post-intervention phases of the eye-gaze and touchscreen conditions. One explanation for this phenomenon might be that children were not able to learn the difference between the stimuli during the learning phase – that the active stimulus was associated with the reward and the inactive stimulus was not. However this seems unlikely as there was a clear preference for the active stimulus (selected in 75% of trials) over the inactive stimulus (selected in 20% of trials) in the touchscreen condition of the baseline phase. This suggests that children in the study were able to learn the association between stimuli and reward. The 20% selections of the inactive stimulus in the touchscreen condition might be explained by some degree of generalisation between the stimuli. This leaves the question of why children did not select the active stimulus more frequently in the eye-gaze condition. The simplest explanation is that children did not make deliberate selections using eye-gaze: that they did not fully understand how to control the eye-gaze device and were simply looking at the stimuli since these were the only things on the screen. In any discussion of eye-gaze technology, it is important to acknowledge the potential impact of the "Midas Touch" problem [21], which posits that any access method that uses the same channel to both receive information and transmit control signals is prone to over-selectivity and accidental selection. In the case of the present study, it may be that children were simply looking at the stimuli, which could result in selections whether or not these were intended. Since the stimuli included human faces, children may have a predisposition to look at them, which has previously been discussed as an important consideration in the design and development of eye-gaze technology interfaces for children requiring augmentative and alternative communication [22]. The advantage of the methodology used in the present study is that the comparison with the touchscreen gives us some insight into what deliberate choice looks like, and to contrast this with the eye-gaze condition.

Addressing the first research question, our study found no evidence that children could necessarily infer for themselves the causal mechanism by which an eye-gaze device functions. This finding emphasises that it would be risky to assume children will simply be able to acquire these skills without significant ongoing support. The current literature lacks empirical evidence on how difficult these skills are for children to acquire and what evidence does exist suggests that such skills can take a long time to develop. One interesting study [23] looked at children with severe physical impairments ($n=10$), none of whom had any previous experience of using eye-gaze technology. These children were all issued with eye-gaze devices and their parents and support team were given a dedicated two-day introduction to the technology and the software they would be using. Thereafter the technology was used daily, with regular input from a multi-disciplinary team for 9–10 months and no other interventions being carried out during this time. Longitudinal follow-up of these children indicated that they all showed improvements on an activity involving a single target: with children improving in their speed of targeting after 5 months and in their accuracy after 15–20 months. Notably, this study involved a task which was chosen for its low cognitive and language demands. The challenges for children learning to use eye-gaze for more complex tasks including AAC or language learning will likely be greater.

One explanation for the difficulty of children's inferring this causal mechanism for themselves may be the difference in contiguity between the two access methods. A touchscreen provides clear spatial and temporal contiguity – the finger press on the screen is at the location where the effect occurs. Eye-gaze technology may present a unique challenge for young children attempting to infer a causal mechanism. Although there is temporal contiguity and distal spatial contiguity between cause (one's own eye movement) and effect (events such as cursor movement on screen), this relationship may be particularly difficult to discern. The feedback of seeing your finger press a screen icon or hand move a mouse may be more salient than only having the proprioceptive feedback of moving one's own eyes.

Examples of children's inferring a causal mechanism in the absence of spatial contiguity exist in the wider literature. Kushnir and Gopnik [24] used a play-based activity in which a selection of objects activated a musical toy when either placed on that toy (contiguous) or held over it (non-contiguous) to demonstrate that 3–4 year old children are more likely to make the correct inferences about causal mechanisms when the cause had spatial contiguity with the effect. However, their study also demonstrated that children could use both probabilistic and deterministic reasoning to infer non-contiguous causal relationships and apply this to new interactions with novel objects – which did not appear to be the case with eye-gaze interactions in the present study. The difference in the amount and type of feedback provided by eye-gaze technology may be a factor here. It has been highlighted [25] that the lack of any proprioceptive, auditory or haptic feedback from eye-gaze systems may make the mechanism by which eye movements control the system harder to identify for new users. Knowledge and understanding of a mechanism underpins how events with only distal spatial contiguity (or no spatial contiguity at all) can be treated as causal, for example the turning on of a light using a light switch, using a switch-activated toy or controlling an onscreen character with a joystick. Eye-gaze technology lacks both spatial contiguity and an obviously inferred causal mechanism, especially in the context of a comparative paucity of feedback provided to the user.

The lack of success in the eye-gaze condition was not due to children's repeatedly touching the screen. The percentage of trials in the eye-gaze condition that resulted in children touching the screen was relatively low – an average of 2.7 trials (21.7%) during the baseline phase and 2.0 trials (16.7%) in the post-intervention phase. This suggests that children learned quickly that the touchscreen was deactivated and did not persist with attempting to complete the task using this access method.

The second research question explored whether a teaching intervention based on causal language could improve the performance of children in the study. The addition of causal language had the immediate effect of improving performance with both access methods during the intervention phase, compared to the baseline and post-intervention phases when only minimal prompting and non-specific encouragement were given. These findings suggest that causal language instruction may be a practical way to scaffold eye-gaze access. The results also add further weight to the idea that the causal relationship that underpins use of eye-gaze technology is not intuitive for younger children, since their performance only approached that of touch screen access, a more familiar and more intuitive access method, with the additional support of proactive teaching.

The use of causal language instruction to support children's learning merits some further discussion. Whilst this scaffolding was in place, children's performance did increase to a level similar to their performance with a touchscreen. However, there may be questions about what this approach is actually teaching children. The approach uses consistent verb forms ("looking at the woman makes the video play, can you make the video play?") but children do not need to understand all of this instruction to be successful, or explicitly understand that the action of looking produces the outcome. In fact, simply following the instruction to "look at the woman" would lead to a successful outcome. In this way, the intervention may be focusing children on the target (the woman) but not necessarily explicitly teaching the child that there is a causal link between eye movements and outcomes on the screen. This suggests that the higher levels of performance during intervention are not necessarily due to an understanding of a causal link between eye gaze and outcomes on the screen. Interestingly, a similar pattern of performance increase and regression in the touchscreen condition was also seen, suggesting that in both conditions key prompts are being followed but causal mechanisms may not have been learned.

The final research question concerned how performance differed on the same task controlled by eye-gaze and by touchscreen. Children performed better on the touchscreen version of the task across all three phases, which suggests that the weaker performance on the task when using eye-gaze technology may be attributable to difficulty inferring the mechanism of gaze control. The modification of the touchscreen to include a delayed activation, whilst not a perfect analogue for the novelty of eye-gaze control, makes it less likely that children's better performance with a touchscreen is solely attributable to pre-existing familiarity, although it would be logical to assume that familiarity played some part. Whilst children's amount of touchscreen use in everyday life was not enquired about in this study, results published elsewhere in the literature suggest that 92.05% of children aged 26–36 months used a touchscreen daily for an average of 43.95 min [26]. Nationally available statistics in the UK suggest that the majority of 3–4 year olds are familiar with using touchscreen tablet computers to access online content [27]. By contrast, eye-gaze control was entirely novel to all children.

Whilst the typically developing children in this study could perform to a high level with the eye-gaze device whilst instruction was ongoing, we found no evidence that they could do this either intuitively or independently. Applying these findings to the clinical population of children most likely to be considered candidates for eye-gaze technology, it appears that there may be some advantages to using approaches such as causal language instruction, although it remains likely that learning will take time and repetition. Further, this research highlights risks in making assumptions about what children are learning from eye-gaze activities, particularly those claiming to teach cause and effect or other foundation skills needed for control of a device. Clinicians should be wary of making assumptions about what children are learning from such activities, particularly in regard to developing an understanding of the causal mechanism by which eye-gaze is controlled. Eye-gaze itself may not, by extension, be a helpful method of teaching cause and effect for children in which the skill is not already established.

Importantly, these results do not suggest that children should be denied access to eye-control technology. Rather, the findings underline the importance of planning interventions carefully, recognising the complexity of acquiring this novel method of control and managing expectations. This is in line with recently published guidelines for the assessment and introduction of eye-gaze technology [6, 28] where stakeholders including eye-gaze users, their support teams and other professionals involved in assessment, underlined the importance of understanding the expectations related to this complex technology.

Limitations

The number of children included in this work is small and it may be interesting to replicate the present study with a larger group. In addition, the methodology only includes one session of causal language intervention. There is scope for a longer-term study which looks at the impact of more teaching sessions on the consolidation of these skills over a longer period. The study does not contrast causal language instruction with other methods of instruction or teaching, which may be helpful to explore in future work. It is also conceivable that the instruction given to some children to push the touchscreen harder if it did not activate could be construed as causal, although this was considered by the researchers to be acceptable, since the prompt does not follow the format of causal language instructions used elsewhere in the study and was only used in the event of the screen not activating despite a clear, intentional press.

Summary

This paper offers new insight into the potential challenges for younger children in inferring for themselves the causal link between eye movements and control of an eye-gaze device. The results may suggest that this link is not straightforward for children to learn without significant support: children were able to improve their performance with explicit instruction, but their regression to baseline performance levels when this was removed calls into question whether they had fully acquired this causal link.

Disclosure statement

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

Funding

The author(s) reported there is no funding associated with the work featured in this article.

Notes on contributors

Tom Griffiths is a Lecturer in Assistive Technology in the Department of Computing at the University of Dundee and a Clinical Scientist. His areas of professional practice are augmentative and alternative communication (AAC), assistive technology and computer access. His research areas include access to computers for children with physical disability, with a particular research interest in eye-gaze and eye tracking technologies.

Michael T. Clarke is a Speech Language Pathologist (SLP) and an Assistant Professor in the Department of Speech, Language and Hearing Sciences in the Graduate College of Education. His area of professional expertise is in augmentative and alternative communication (AAC). His research in the area of AAC is motivated by questions arising from clinical practice. He has a particular research interest in in-person interaction involving children who are provided with AAC systems and tools.

John Swettenham is a Reader in Developmental Psychopathology at University College London. His research focuses on the development of cognition and communication in children with developmental disorders such as autism and motor disorders such as cerebral palsy. He has a particular interest in the role of attention and perception in the development of joint attention skills and key milestones in development such as theory of mind. His work on cerebral palsy includes studies examining how eye gaze is used for communication.

ORCID

Tom Griffiths http://orcid.org/0000-0002-1542-8128
Michael T. Clarke http://orcid.org/0000-0002-8933-8934
John Swettenham http://orcid.org/0000-0001-8844-6346

References

[1] Karlsson P, Allsop A, Dee-Price B-J, et al. Eye-gaze control technology for children, adolescents and adults with cerebral palsy with significant physical disability: findings from a systematic review. Dev Neurorehabil. 2018;21(8):497–505. doi:10.1080/17518423.2017.1362057.

[2] Griffiths T, Price K. A proposed framework for decision-making for assistive communication technology support: many perspectives, but one common goal. J Assist Technol. 2011;5(4):242–248. doi:10.1108/17549451111190641.

[3] Chantry J, Dunford C. How do computer assistive technologies enhance participation in childhood occupations for children with multiple and complex disabilities? A review of the current literature. Br J Occup Ther. 2010;73(8):351–365. doi:10.4276/030802210X12813483277107.

[4] Vessoyan K, Smart E, Steckle G, et al. A scoping review of eye tracking technology for communication: current progress and next steps. Curr Dev Disord Rep. 2023;10(1):20–39. doi:10.1007/s40474-023-00271-x.

[5] Borgestig M, Sandqvist J, Ahlsten G, et al. Gaze-based assistive technology in daily activities in children with severe physical impairments–an intervention study. Dev Neurorehabil. 2017;20(3):129–141. doi:10.3109/17518423.2015.1132281.

[6] Karlsson P, Griffiths T, Clarke MT, et al. Stakeholder consensus for decision making in eye-gaze control technology for children, adolescents and adults with cerebral palsy service provision: findings from a Delphi study. BMC Neurol. 2021;21(1):63. doi:10.1186/s12883-021-02077-z.

[7] Light J, McNaughton D. Putting people first: re-thinking the role of technology in augmentative and alternative communication intervention. Augment Altern Commun. 2013;29(4):299–309. doi:10.3109/07434618.2013.848935.

[8] Gosnell J, Costello J, Shane H. Using a clinical approach to answer "what communication apps should We use? Perspect Augment Altern Commun. 2011;20(3):87–96. doi:10.1044/aac20.3.87.

[9] Michotte A. The perception of causality. New York: Basic Books; 1963.

[10] Saxe R, Carey S. The perception of causality in infancy. Acta Psychol (Amst). 2006;123(1–2):144–165. doi:10.1016/j.actpsy.2006.05.005.

[11] Bonawitz EB, Ferranti D, Saxe R, et al. Just do it? Investigating the gap between prediction and action in toddlers' causal inferences. Cognition. 2010;115(1):104–117. doi:10.1016/j.cognition.2009.12.001.

[12] Gopnik A, Glymour C, Sobel DM, et al. A theory of causal learning in children: causal maps and Bayes nets. Psychol Rev. 2004;111(1):3–32. doi:10.1037/0033-295X.111.1.3.

[13] Schlottmann A. Perception versus knowledge of cause and effect in children: when seeing is believing. Curr Dir Psychol Sci. 2001;10(4):111–115. doi:10.1111/1467-8721.00128.

[14] Goswami UC. Children's cognitive development and learning (report no.: 3). York: Cambridge Primary Review Trust; 2015.

[15] Sobel DM, Sommerville JA. Rationales in children's causal learning from others' actions. Cogn Dev. 2009;24(1):70–79. doi:10.1016/j.cogdev.2008.08.003.

[16] Gopnik A, Sobel DM, Schulz LE, et al. Causal learning mechanisms in very young children: two-, three-, and four-year-olds infer causal relations from patterns of variation and covariation. Dev Psychol. 2001;37(5):620–629. doi:10.1037/0012-1649.37.5.620.

[17] Sennott SC, Light JC, McNaughton D. AAC modeling intervention research review. Res Pract Pers Sev Disabil. 2016;41(2):101–115. doi:10.1177/1540796916638822.

[18] Higginbotham DJ. Use of nondisabled subjects in AAC research: confessions of a research infidel. Augment Altern Commun. 1995;11(1):2–5. doi:10.1080/07434619512331277079.

[19] Sharma A, O'Sullivan T, Baird G. Clinical evaluation of child development from birth to five years. Psychiatry. 2008;7(6):235–241. doi:10.1016/j.mppsy.2008.04.004.

[20] Cohen LB, Oakes LM. How infants perceive a simple causal event. Dev Psychol. 1993;29(3):421–433. doi:10.1037/0012-1649.29.3.421.

[21] Jacob RJK. Eye movement-based human-computer interaction techniques: toward non-command interfaces. In: Harston HR, Hix D, editors. Advances in human-computer interaction. Norwood (NJ): Abex Publishing Co.; 1993. p. 151–190. Available from: http://www.cs.tufts.edu/~jacob/papers/hartson.pdf.

[22] Wilkinson KM, Light J. Preliminary investigation of visual attention to human figures in photographs: potential considerations for the design of aided AAC visual scene displays. J Speech Lang Hear Res. 2011;54(6):1644–1657. doi:10.1044/1092-4388(2011/10-0098).

[23] Borgestig M, Sandqvist J, Parsons R, et al. Eye gaze performance for children with severe physical impairments using gaze-based assistive technology—a longitudinal study. Assist Technol. 2016;28(2):93–102. doi:10.1080/10400435.2015.1092182.

[24] Kushnir T, Gopnik A. Conditional probability versus spatial contiguity in causal learning: preschoolers use new contingency evidence to overcome prior spatial assumptions. Dev Psychol. 2007;43(1):186–196. doi:10.1037/0012-1649.43.1.186.

[25] Sowers DJ. The effects of intermediary feedback on eye tracking as an access method for AAC [Doctoral thesis]. Pennsylvania (PA): The Pennsylvania State University; 2022 [cited 2023 Jul 12]. Available from: https://etda.libraries.psu.edu/catalog/20385djs6974.

[26] Bedford R, Saez De Urabain IR, Cheung CHM, et al. Toddlers' fine motor milestone achievement is associated with early touchscreen scrolling. Front Psychol. 2016;7:1108. [cited 2023 Aug 10]. Available from: https://doi.org/10.3389/fpsyg.2016.01108.

[27] Ofcom. Children and parents: media use and attitudes report 2018. Warrington: Office of Communications (Ofcom); 2018. p. 1–17.

[28] Eyes on Communication Research Group. Eye-gaze control technology for people with cerebral palsy - clinical guidelines [Internet]. Cerebral Palsy Alliance, The University of Sydney; 2021 [cited 2023 Apr 28]. Available from: https://redcap.sydney.edu.au/surveys/?s=EDC7P4E3TP.