Int. J. Human–Computer Studies 216 (2026) 103896

Contents lists available at ScienceDirect

International Journal of Human - Computer Studies

journal homepage: www.elsevier.com/locate/ijhcs

Chaining subtle gaze direction cues to guide users in extended reality

Teresa Hirzle*, Jonathan Sutton, Thomas Edward Larsen, Lise A. Bruun, Kasper Hornbæk

University of Copenhagen, Department of Computer Science, Universitetsparken 1, Copenhagen, 2100, Denmark

HIGHLIGHTS

- Chaining subtle gaze direction cues effectively guides gaze to multiple targets.
- Chaining subtle gaze direction cues prolongs runs during visual foraging.
- Guiding gaze to the nearest target is the most effective chaining technique.
- Chaining subtle gaze direction cues also increases the speed of target selection.

ARTICLE INFO

Keywords:
Gaze guidance
Subtle gaze direction
Extended reality
Visual attention
Eye tracking

ABSTRACT

Directing a person's attention by guiding their gaze to an object of interest is a promising application of extended reality (XR) to support people's everyday lives. XR gaze guidance is not only useful for assisting visual search, but also for influencing a person's viewing order of objects, which affects memory and comprehension. For example, following an expert's eye movements can enhance one's proficiency in identifying relevant information. Existing gaze guidance techniques typically guide users to one target at a time. However, many real-world tasks require guidance to multiple targets, where the viewing sequence is crucial for comprehension or problem-solving. This paper addresses this gap by exploring how to guide users to multiple targets in a sequential order by applying guidance cues several times consecutively (i.e., chaining them). This paper presents a user study (N = 35) investigating a continuous gaze guidance technique that chains individual cues based on subtle flickering (i.e., subtle gaze direction). Participants perform visual foraging–a controlled search task with multiple target and distractor objects. The results suggest that subtle chaining is effective when it supports users' search strategies, but less effective in guiding users to targets that require a higher degree of cognitive processing. We discuss how the results can improve the design of gaze guidance techniques for multiple targets in XR.

1. Introduction

The way we view a scene crucially influences our comprehension. For example, the sequence in which we view objects affects our ability to remember their positions and recall them later (Johansson and Johansson, 2014; Bailey et al., 2012). Guiding gaze to relevant information by manipulating the viewing sequence can enhance problem solving (Grant and Spivey, 2003) or improve proficiency in identifying task-relevant information by following an expert's gaze (Jarodzka et al., 2013).

This paper presents a continuous gaze guidance technique based on subtle gaze direction (SGD) cues (Bailey et al., 2009). It extends prior work by introducing the chaining of SGD cues, focusing on perceptual effects such as whether sequential guidance increases the likelihood of selecting subsequent cued targets. SGD applies flicker in the periphery, where sensitivity to light changes is higher, and stops it once a saccade is initiated, preventing perception on the fovea. Thus it attracts gaze quickly, while remaining unobtrusive. To our knowledge, this is the first study of sequential SGD application in multi-target tasks, bringing the technique closer to real-world scenarios. Before detailing our technique and its contributions, we illustrate its potential through three examples in everyday contexts, where continuous gaze guidance supports attention and trains users to follow specific viewing sequences.

> Envision that as you are about to leave your house, your gaze is subtly guided toward objects in order of importance. For example, you might first confirm that the stove and lights are turned off,

* Corresponding author.
Email addresses: tehi@di.ku.dk (T. Hirzle), josu@di.ku.dk (J. Sutton), tla@di.ku.dk (T.E. Larsen), libr@di.ku.dk (L.A. Bruun), kash@di.ku.dk (K. Hornbæk).

https://doi.org/10.1016/j.ijhcs.2026.103896
Received 27 October 2025; Received in revised form 9 May 2026; Accepted 14 July 2026
Available online 16 July 2026

before noticing your keys on the shelf. Repeated guidance in this sequence can both support you in the moment and train you over time to prioritize critical items. For example, remembering to take the keys and ensuring that the stove is turned off might be more important than remembering the lights are off. Continuous gaze guidance ensures that you look at the lights last.

Similarly, when learning to drive, gaze guidance could direct attention to the most relevant vulnerable road users. For example, you might first want to notice cyclists that you are getting close to on the road before noticing pedestrians who are walking on the sidewalk. Over time, this can help learners develop safer, more expert-like scanning patterns.

Finally, when learning to play chess, gaze guidance could reflect how an expert (e.g., Magnus Carlsen[a]) evaluates the board by paying attention to the pieces in a particular order. For instance, it could guide attention first to immediate threats (e.g., checks), then to attacked pieces, and finally to longer-term strategic opportunities. This can support better decisions in the moment and facilitate the learning of effective visual strategies.

[a]Magnus Carlsen is a Norwegian chess grandmaster often renowned as the best chess player of our time.

These examples illustrate the practical advantage of continuous gaze guidance: directing attention in a particular order, helping users remember important steps (Example 1), prioritizing relevant information (Example 2), and learning to evaluate situations where multiple objects compete for attention (Example 3). This paper lays the groundwork for such applications by studying how to guide users to multiple targets through sequential SGD cues.

Originally developed for 2D displays (Bailey et al., 2009; McNamara et al., 2008), SGD has been extended to virtual reality (VR) (Grogorick et al., 2017, 2018), successfully guiding people to peripheral targets. However, its use for guiding people to multiple targets sequentially has not been studied. While technically chaining SGD cues is equivalent to applying single cues sequentially, repeated application might cause unknown perceptual effects, such as changes in responsiveness over time. Similar effects have been observed with overt guidance; for example, Sutton et al. (2024) report that some participants followed the cues instead of the targets. While we do not expect similar effects on agency, we investigate how SGD cues influence selection strategies over time, as this is crucial for their practical application in continuous gaze guidance.

This work investigates three variants of chaining SGD cues. The three variants differ in how the next modulation target is selected: (1) the closest target (*close chaining*), (2) the furthest target (*distant chaining*), or (3) a randomly chosen target (*random chaining*). We examine these variants in a controlled multi-target search task, where targets are not ordered. We expect *close chaining* to be most effective, as it aligns with typical search behavior favoring adjacent targets (Kristjánsson et al., 2020a). In contrast, *distant chaining* disrupts this pattern by directing gaze to far targets. While this should attract attention quickly due to high sensitivity to peripheral motion, we expect it to be less effective because of the disruption to typical search behavior. Lastly, *random chaining* breaks systematic search entirely, testing how easily participants can be diverted from systematic search behavior.

We conducted a within-participant study (N = 35), comparing three chaining variants to a baseline without guidance in a visual foraging task. In visual foraging, participants search for multiple targets among distractors without a predefined order (Kristjánsson et al., 2020a). It is increasingly used in cognitive psychology because it offers a more naturalistic scenario than single-target search while maintaining experimental control compared to real-world search tasks (Tünnermann et al., 2022; Wolfe, 2013). This paradigm allows us to study complex viewing behavior and low-level visual processing in a controlled setting. Real-world tasks additionally involve higher-level cognitive functions (e.g., planning and reasoning), which we abstract from in this study to isolate underlying perceptual mechanisms. We used VR to maintain control over the visual surroundings. Our primary interest is in how chaining SGD cues influences the order of target selection, analyzed using *runs* (consecutive selections of the same target type). As a secondary measure, we report performance in terms of target collection rate.

The remainder of this paper is structured as follows: We first discuss the relationship between viewing order and gaze guidance, and explain visual foraging. Next, we introduce the concept of chaining SGD cues and its three variants. Then, we present the empirical work, including study design, analysis, and results. We conclude with a discussion of the findings, their implications for the design of gaze guidance techniques, and future work.

2. Related work

2.1. The impact of viewing order on memory and learning

This section discusses how viewing patterns and viewing order impact cognitive processes, most importantly visuospatial memory and learning. Studies support that eye movements are functional for visual memory retrieval and response time. The phenomenon of “looking at nothing” shows that people tend to look at previous target locations even when they are not visible (e.g., Spivey and Geng, 2001; Martarelli and Mast, 2013; Brandt and Stark, 1997). While most studies have been conducted in 2D settings, Chiquet et al. (2021) investigated this phenomenon in VR and found a higher fixation proportion on previous target locations compared to locations that did not correspond to the target. Interestingly, 89% of the time, participants looked at locations unrelated to any objects.

More relevant for our work are studies suggesting that impacting viewing order can indeed help people retrieve objects more accurately, define their positions more precisely, and respond more quickly. Johansson and Johansson (2014) found that guiding eye movements to previous target locations improved accuracy and speed, while restricting gaze (i.e., participants looking at a fixation cross during memory retrieval, preventing eye movements to target locations) increased retrieval times. Bochynska and Laeng (2015) showed that following previous scanpaths improved memory retrieval, with better recognition and faster responses when patterns appeared in the original sequence. They also confirmed Johansson and Johansson’s (2014) findings that prohibiting gaze movements by fixating gaze on the screen’s center reduced accuracy and increased retrieval time, even with previously seen sequences. In these examples, geometric cues directed gaze (e.g., by drawing a shape on the blank space to guide participants’ gaze to the correct target location). Bailey et al. (2012) applied a subtle method to guide participants’ gaze already during the encoding phase (i.e., when targets are presented). The participants drew rectangles on a blank screen, indicating recalled targets. One group had their gaze guided towards targets during the encoding phase and one had their gaze guided away from targets. Results showed guiding gaze to targets reduced counting and location errors compared to not applying a gaze direction method. Guiding participants’ gazes away from targets resulted in a higher location error compared to not applying any direction.

In more applied directions, Jarodzka et al. (2013) showed that guiding gaze affects not only memory, but also learning. They overlaid expert eye movements on a video for participants classifying fish locomotion. In a second iteration of the task, participants with gaze guidance had more coherent viewing patterns and better identification of relevant parts. McNamara et al. (2012) used subtle gaze cues to guide viewers through paintings in a specific order, enhancing visual literacy.

We conclude that from controlled studies (Johansson and Johansson, 2014; Bochynska and Laeng, 2015) (e.g., following the “looking at nothing” study paradigm) to applied studies (Jarodzka et al., 2013; McNamara et al., 2012), prior work has demonstrated that influencing viewing order impacts memory and visual pattern learning. These results are promising for our approach and demonstrate that influencing a person's viewing order in their everyday life using XR could indeed have positive effects on everyday search and learning tasks. Next, we discuss gaze guidance techniques that have been developed for XR to explicitly or implicitly guide a person's gaze.

2.2. Gaze guidance techniques

Subtle gaze guidance. Subtle gaze guidance has been explored in various practical applications, such as "Gaze-Guided Narratives", where audio content is implicitly adapted to tourists' gaze points (Kwok et al., 2019). Examples include the guidance of attention through contrast and other visual techniques (Todd and Manaligod, 2018; Lu et al., 2012). Temporal approaches like flicker are known for their effectiveness in drawing attention quickly (Abrams and Christ, 2003). Different versions of flicker have been studied for their ability to guide attention subtly. Waldner et al. (2014), for example, combined two intensities—one to attract attention and one to maintain it on the stimulus. Subtle gaze direction (SGD) is a version of flicker, where the modulation is turned off when a saccade is performed towards the stimulus, such that it remains unnoticeable (Bailey et al., 2009). SGD has been applied on different settings and shown to be successful in 2D screens (Bailey et al., 2009, 2012), in real-world settings based on a projector that creates the modulation (Booth et al., 2013), and in VR (Grogorick et al., 2017).

Gaze guidance in AR and VR. Several guidance techniques have been explored, including visual cues, such as spheres, vignettes, or light changes, to attract gaze both in VR (Woodworth et al., 2023; Lankes and Ramirez Gomez, 2022) and AR (Seeliger et al., 2021). A comprehensive overview of visual attention guidance in VR is provided in a book chapter by Grogorick and Magnor (2020). This work categorizes attention guidance into overt methods (e.g., blur, highlighting, overlays) and subtle methods (e.g., spatial blur, saliency adjustment, temporal luminance modulation, binocular rivalry). Examples of subtle gaze guidance in AR and VR include binocular rivalry, a technique that creates mismatches in binocular images for preattentive visual guidance (Krekhov et al., 2020), or saliency modulation, which refers to changing properties like contrast or color to attract attention (Rivu et al., 2021; Sutton et al., 2022). Other examples include depth-of-field effects, small red dots, local magnification, and spatial blur (Lu et al., 2012; Grogorick et al., 2018). These techniques have been shown to effectively guide attention without being obtrusive. Lange et al. (2020) presented a technique for VR in which guidance is integrated into the narrative using the swarm motion of particles to direct attention. However, for content-independent guidance, subtle gaze direction remains a preferred method (Bailey et al., 2009). Flicker-based guidance in 360 video VR has been shown to increase recall rates for details in the guided area (Rothe et al., 2018).

SGD has also been explored in VR. Grogorick et al. (2017) show that SGD improves the detection speed of hidden targets. In 3D environments, target-directed saccades to stimuli in a person's periphery have been observed, demonstrating the utility of SGD in immersive settings (Grogorick et al., 2018).

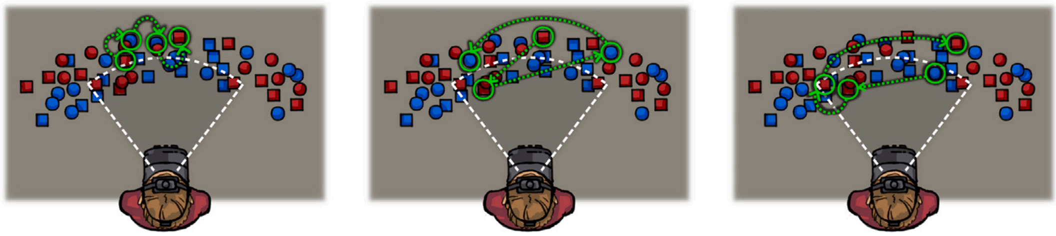
The main challenge of using subtle gaze guidance cues in AR and VR is to attract gaze without a conscious response. Techniques must balance noticeability and effectiveness. Subtle gaze direction (a temporal approach) can successfully guide the user's gaze to targets. It is fast in attracting gaze, and the interplay of showing a modulation in the periphery and turning off the modulation in the foveal part of the view has been successfully used for unobtrusive guidance for individual targets. However, the studies discussed do not systematically investigate guidance to several targets. In the cases where they guide the user's gaze to several targets, this does not happen in a systematic way, as the interest lies in guiding the gaze to an individual object rather than chaining guidance to several ones. Our work extends these studies by investigating chaining SGD cues in a controlled visual search task, systematically varying target and distractor objects.

2.3. Visual search and visual foraging

Visual search is a cognitive process that we perform constantly in our everyday lives (Wolfe, 2020). It is crucial for efficiently navigating and interacting with our surroundings. For example, we search for our keys before leaving the house, buy specific items in a grocery store, or look for familiar faces at a conference. Visual search is influenced by attention (Wolfe and Horowitz, 2017) and directed by features (Wolfe, 2020). When searching for objects, we consider them as representations of relevant features rather than specific or individual instances. Common features include color, size, or motion, but can also be more complex like luminance, shape, or shading (Wolfe, 2014). For example, when looking for a red apple among green ones, color is the most important feature, while when looking for green apples among pears, shape is more important. Visual search tasks formalize these processes. In hybrid foraging, for example, participants search "for multiple instances of several types of targets at the same time" (Wolfe et al., 2019), and in conjunctive search, a target is defined by a combination of features (e.g., size and color) (Egeth et al., 1984). Visual foraging refers to a search task with multiple target objects among multiple distractor objects (Kristjánsson et al., 2020a). It emerged as an alternative to the classical single search paradigm, where one target has to be retrieved, typically in a setting where the task is to indicate whether a target is present or not (Wolfe, 2013). Visual foraging is recognized as a more naturalistic task for investigating search behavior, allowing for the study of more complex search tasks where the observer does not know how many targets are present or when the search is completed (Wolfe, 2013). Several methodological decisions must be made when applying a foraging task. Wolfe et al. (2016) introduced the hybrid foraging paradigm, describing a search task in which observers are tasked to find multiple instances of multiple target types, combining foraging (finding multiple instances of a target) and hybrid search (finding multiple targets that are present at once). During hybrid foraging, observers tend to exclusively forage for one type of target before moving to the next type, rather than switching between target types. Going into more detail, Kristjánsson et al. (2014) introduced the concept of conjunction foraging—a version of hybrid foraging where a target is determined by two or more features (e.g., color and shape). This is in contrast to feature foraging, where only one feature (e.g., only color) determines a target. Kristjánsson et al. (2014) found similar foraging behaviors, confirming that for feature foraging observers switched between target categories (e.g., red and blue spheres are targets, green and yellow spheres are distractors), while for conjunction foraging observers tended to exclusively select targets from one category first before switching to the next (Jóhannesson et al., 2016; Kristjánsson et al., 2014). Different variations of hybrid foraging have been studied. For example, adding different values to targets (Wolfe et al., 2018), varying feature context by using more or less similar colors (Tünnermann et al., 2021), or selection modality (Ian M. Thornton and Kristjánsson, 2019; Jóhannesson et al., 2016). Kristjánsson et al.'s (2022) work is most relevant to our study, as they presented a first investigation of a foraging task in virtual reality. Therefore, we chose a similar task for our study, also using conjunction foraging with shape and color as features. In general, conjunctive visual foraging is useful for studying chaining subtle gaze guidance cues, as it is known to be cognitively demanding. Therefore, it helps us understand how chaining would work in a demanding, real-world task.

3. Chaining subtle gaze direction cues

We investigate chaining gaze direction cues based on the mechanism described by Bailey et al. (2009). By *chaining* we refer to applying the guidance mechanism (i.e., subtle gaze direction or flickering) to a new



(a) *Close Chaining* (b) *Distant Chaining* (c) *Random Chaining*

Fig. 1. Illustrations of the concept of the three chaining variants.

target object after an object is selected, thus guiding users to several consecutive targets.

We examine three chaining variants that differ in how the next target to be modulated is selected (see Section 3.1). The basic steps of chaining SGD cues are:

1. An object is determined for modulation. We explored three variants to determine the next object to modulate (detailed in Section 3.1).
2. The object of interest is modulated with SGD (i.e., the object is flickering; detailed in Section 3.2). To ensure that the modulation is subtle, we determine the flicker intensity through a psychophysics staircase calibration described in Section 4.4.
3. When a person performs a saccade towards the modulated object, the modulation stops. We describe saccade detection and turning off modulation after detection in Section 3.3.

3.1. Three variants of subtle gaze direction chaining

1. *Close Chaining* modulates the target that has the shortest Euclidean distance to the one just selected (see Fig. 1a). This variant should be successful for targets that are located nearby. It is in line with typical search behavior for multiple targets, where individuals tend to select adjacent targets before switching to more distant ones (Kristjánsson et al., 2020a). As it supports typical search behavior, we expect the variant to successfully guide individuals to modulated targets. However, as a saccade towards the target direction stops the modulation, modulation time might be short. Therefore, modulation might not trigger a response. This would occur when participants perform a saccade towards a nearby target that is not the modulated one.
2. *Distant Chaining* modulates the target that has the largest Euclidean distance to the one just selected and is still in the person's field of view (see Fig. 1b). This variant presents a contrast to close chaining by interfering with typical search behavior, directing participants' gaze away from their current focus to a target in their periphery. We implemented this variant to test how strongly the flickering modulation directs participants to targets far away, disrupting their typical foraging behavior.
3. *Random Chaining* modulates a target of random Euclidean distance to the one just selected, considering that it is still in the person's field of view (see Fig. 1c). This variant introduces a random unsystematic factor that contrasts the systematicity of the other two variants. We chose this approach in addition to the two systematic ones, to explore the extent systematic redirection has on participants. The unsystematic pattern of *Random Chaining* might be less reliable and less predictable for participants.

3.2. Subtle gaze direction modulation

SGD (Bailey et al., 2009) is based on the concept that flickering (i.e., high temporal frequency changes in color or luminance) is highly effective in attracting gaze when applied in the peripheral view, where the visual system is highly sensitive to light changes. In contrast, high-frequency flicker is less visible in the foveal part of the view which has a higher resolution but is slower in processing movement information. Therefore, a flickering modulation in the periphery will alert the visual system initiating a movement towards the flickering target. As soon as the eyes move into the direction of the target (i.e., perform a saccade), the flickering modulation is stopped and thereby remains undetected by foveal vision (see Fig. 2a). The technique takes advantage of saccadic masking, which refers to the “temporary suppression of visual processing during eye movements” (Bailey et al., 2009). Let $O_{fixated}$ be the object that was last fixated and $O_{flickering}$ the currently flickering target. Let $\vec{v}$ be the vector of the current saccade, and $\vec{w}$ the vector between the last fixation and the current flickering target. Let θ be the angle between $\vec{v}$ and $\vec{w}$. θ and $\vec{v}$ are updated continuously with a frequency of 90 Hz. θ is determined by calculating the cosine distance between $\vec{v}$ and $\vec{w}$.

$$\theta = 1 - \frac{\vec{v} \cdot \vec{w}}{\|\vec{v}\|\|\vec{w}\|} \tag{1}$$

The modulation is stopped when $\theta \leq 10°$, as this indicates that a saccade is being performed in the direction of the flickering object. Thus, the flickering can be stopped without the person noticing, using saccadic masking, as illustrated in Fig. 2(b).

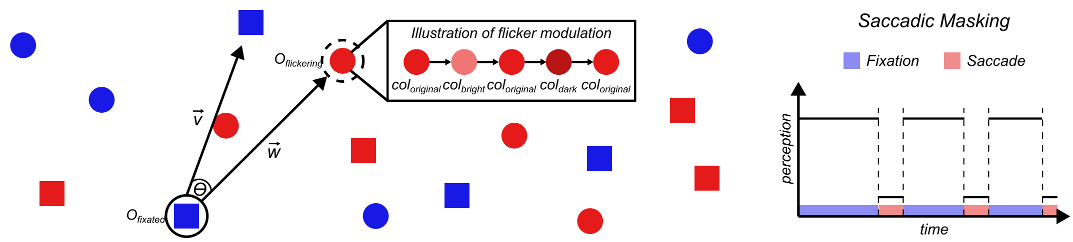
We chose a lightness-based flicker modulation that alternated the color of each object with a brighter (higher lightness value) and darker (lower lightness value) version of that color, as it was shown to be effective in the original experiment (Bailey et al., 2009) and in VR (Grogorick et al., 2017). The alternating modulations occurred at a rate of 90 Hz. We chose to modulate the entire object rather than employing a gradient or interpolation to a location like Grogorick et al. (2017), as we are interested in attracting gaze at the object level. Flickering was created by alternating the original color of a target object $col_o(t)$ with a darker or brighter version of the color. Fig. 2(a) illustrates the flicker mechanism alternating the object's color between $col_{original}$, col_{bright}, and col_{dark}. We use the *HSL* (hue, saturation, lightness) color model to explain the color modulation. We chose two colors with equal lightness values (based on the HSL color model), such that the flicker modulation would have the same effect on lightness (L) on them (red_{rgb} = $(230, 25, 25)/red_{hsl}$ = $(360, 80\%, 50\%)$, $blue_{rgb}$ = $(25, 25, 230)/blue_{hsl}$ = $(240, 80\%, 50\%)$). During flickering, the object's color alternated between a brighter $col_{bright}(t)$ and a darker $col_{dark}(t)$ color with their lightness values modulated as described below. The intensity i was determined through a perceptual threshold detection calibration, detailed in Section 4.4.

$$L(col_{bright}(t)) = L(col_{original}(t)) + i \tag{2}$$

$$L(col_{dark}(t)) = L(col_{original}(t)) - i \tag{3}$$

3.3. Saccade detection

SGD requires saccade detection to stop the modulation when a saccade is performed toward the modulated object. To achieve this,



(a) *Subtle Gaze Direction Modulation.* (b) *Saccadic Masking.*

Fig. 2. Illustrations of the subtle gaze direction modulation and saccadic masking mechanisms. Figure (b) adapted from Bailey.

we implemented an online, velocity-based saccade detection algorithm based on Engbert and Kliegl's (2003) algorithm for microsaccades. The implementation is based on von der Malsburg's (2015) implementation. We use a buffer size of 5 seconds and a lambda of 6. The algorithm needs five seconds for filling the buffer and initializing saccade detection, as the algorithm adjusts the velocity value based on previous values. Since the experimental task only started after the person had read the instructions (which took more than five seconds), this did not pose a problem. Given our task, we needed to be able to detect relatively large saccades, as the objects were distributed over the entire field of view. Thus, we used a conservative saccade delay filter of 100 milliseconds (i.e., when detection of a saccade has started, it takes 100 milliseconds to register a new saccade).

4. Evaluation

4.1. Study goal

The main goal of this study is to assess the *Effectiveness* of chaining gaze direction in influencing viewing order. The secondary goal is to understand the impact of the techniques on *Performance*.

Effectiveness. Effectiveness refers to how likely participants are to follow the modulated targets. High effectiveness means consistently selecting the flickering target, while low effectiveness means participants are not following the flickering. Note that in both cases it is possible to avoid making any mistakes in the foraging task.

We measure effectiveness via run length, where a run is defined as consecutive selections of the same target type. Since chaining ignores target type and selects based on distance, it can either support or disrupt runs: If the modulated target matches the current target type (we call this *flicker-supported run*), following it should increase run length. If it differs (*flicker-unsupported run*), following it should decrease run length by prompting participants to switch types. Compared to average run length during baseline, we expect:

- increased run length for flicker-supported runs;
- decreased run length for flicker-unsupported runs.

Performance. Performance refers to how efficiently participants collect targets. We measure performance by the number of selected targets, reflected in the collection rate and inter-target times (ITTs).

We expect higher performance in the chaining conditions than in the baseline, assuming that guidance helps participants locate targets more quickly. This should result in an increased collection rate and reduced ITTs. In conjunctive foraging tasks, participants typically select targets of the same type (e.g., red spheres) repeatedly before switching to another type (e.g., blue cubes). If flickering is effective, this pattern should be disrupted, as modulation is applied independently of target type. Following the flicker should therefore lead to more frequent switches between target types. Compared to the baseline, we expect:

- increased overall collection rate (i.e., more targets selected);
- decreased ITTs (i.e., faster target selection);
- a reduced ratio of repeated to switched targets, reflecting more frequent target-type switches.

4.2. Participants

To determine the sample size, we conducted an a-priori power analysis using G*Power.[1] As we were interested in the difference between two means (with vs without subtle gaze direction modulation), we used specific t-tests as a comparison technique. We chose a two-tailed test, with $\alpha = 0.05$ and $1\text{-}\beta = 0.8$, and assumed a medium effect size ($d_z = 0.5$). The analysis suggested 34 participants.

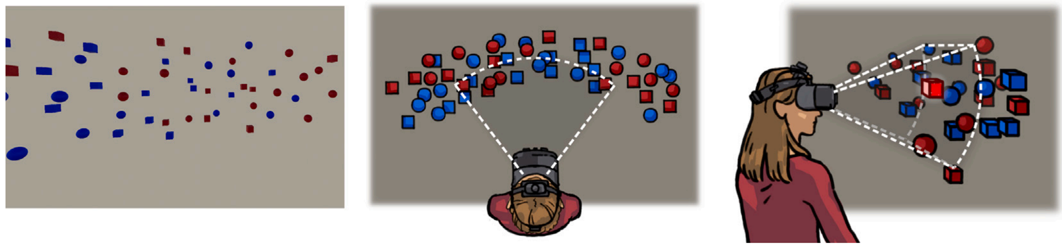
We recruited 36 volunteers (16 women and 20 men[2]) aged between 20 and 64 years with a mean age of 30.14 ($Sd = 10.94$, median age 26.5). Regarding VR experience, five participants reported that they use a VR headset daily or weekly, three reported that they use a headset once a month, two reported they use it occasionally, three reported they have used a headset regularly in the past (but not currently), 18 reported they have tried VR but never used it regularly, and finally five reported having no experience with VR at all. Regarding experience with eye tracking, one participant reported using eye tracking once a month, two reported they use it occasionally, one reported having used eye tracking regularly in the past (but not currently), eight reported they have tried eye tracking but never used it regularly, and finally 24 reported having no experience with eye tracking at all. All participants had normal or corrected-to-normal vision. All gave their informed written consent for participation and were given a present to participate in the experiment. We excluded one participant from the analysis, due to low-quality eye tracking data, resulting in a total of 35 participants for analysis. On average, participants spent 35 to 40min in VR, sometimes prolonged due to re-calibration of eye tracking or if a data logging error occurred.

4.3. Task

The task is based on Kristjánsson et al.'s (2014) conjunction foraging approach (see Fig. 3 for an illustration, note that the flickering object is exaggerated for better visibility). Similar to Kristjánsson et al. (2022), we chose color and shape as the features determining target and distractor objects. We chose blue and red for the color feature and spheres and cubes for the shape feature (see Section 3.2 for the color values). A trial consisted of a search task with 25 target and 25 distractor objects. We chose a non-exhaustive foraging task (Kristjánsson et al., 2020b), limiting each trial to 10 s. Since this increased task difficulty, we expected it to make the effects of chaining more likely. In half of the trials, the targets were blue cubes and red spheres, and the distractors were red cubes and blue spheres; this was reversed for the other half.

[1] G*Power: https://www.psychologie.hhu.de/arbeitsgruppen/allgemeine-psychologie-und-arbeitspsychologie/gpower, last accessed: July 9, 2024.

[2] Gender was collected in an open-response manner, interpreting the following answers as "woman": *woman, f, female, cis female*; the following was interpreted as "man": *man, m, male*.



(a) Participant view. (b) Object distribution. (c) Flickering illustration.

Fig. 3. Visual foraging task.

Participants were instructed to select the targets by looking at them and then pressing a button on a keyboard to confirm the selection. We chose to select the target by gaze and press and not via the controller as Kristjánsson et al. (2022) did, because we were interested in the timing and selecting a target with a controller takes more time than looking at it and pressing a button. We suspected that having more time to select a target would result in spending more time viewing the scene, which might have influenced viewing patterns, as it would decouple the selection and viewing, adding a new step. Participants were instructed to select as many targets as possible in any order, with errors counted but not penalized. They were not instructed about flickering occurring in the main study, as it was calibrated to be imperceptible. Thus, they should focus their attention entirely on the features that determined a target (shape and color). Following the standard procedure in foraging tasks, the targets disappeared after being selected (Kristjánsson et al., 2020a).

4.4. Stimuli

We used spheres and cubes as stimuli, with sphere diameters of $0.2m$ (i.e., Unity units) and cube dimensions of $0.2m \times 0.2m \times 0.2m$. The stimuli were randomly distributed within a 120° arc surrounding the user and distances between $2m$ and $5m$ as shown in Fig. 3(b). On average, target size in visual angle was 2.8°, ranging from 3.3° for the closest targets to 2.4° for the furthest. Visual angles were calculated based on the distance between each target and the participant's viewpoint (i.e., Unity camera). To ensure that stimuli did not overlap or collide with each other, we pregenerated 15 sets of stimulus positions, which were randomly presented to the participants in all conditions.

Determination of stimulus intensity. We needed to determine an intensity for the flickering that was just noticeable for the participants, such that they would follow the stimulus without consciously noticing the flickering. To do so, we used the staircase procedure to determine perceptual thresholds from psychophysics (Cornsweet, 1962). This is an established method in psychophysics, but has also been extensively used in HCI to determine detection thresholds, for example, to detect perception thresholds for visuo-haptic illusions (Feick et al., 2021) or VR redirection (Zenner et al., 2024).

In the staircase procedure, participants are exposed to stimuli at varying intensities to determine the detection threshold. We used an unweighted form, resulting in a 50% detection threshold. This method, also known as the up-and-down method, adjusts stimulus intensity based on detection in the previous trial. We employed a one up/one down method with a one alternative forced choice (1AFC) question ("Did you notice flickering?" yes/no). We used an interleaved staircase with an initial step size of $\Delta_{start} = 0.005$. After four reversals, the step size was adapted to $\Delta = 0.001$. The sequence ended after another 4 reversals occurred. The procedure was repeated eight times, with each object (red/blue cube/sphere) appearing in different locations. Each object served as the flickering object in two different locations during calibration. In half of the trials, the procedure started with an ascending sequence (initial intensity 0.001), and in the other half, with a descending sequence (initial intensity 0.1). The order was randomized to ensure equal representation. The stimulus intensity for each trial was determined by averaging the last two values, and the final intensity was calculated by averaging all eight values.

To determine the intensity of the flicker, we averaged the eight individual flicker calibration values for each participant. The total average of all values for all participants is 0.116 ($Sd = 0.042$, $\min = 0.037$, $\max = 0.197$). However, the values are slightly different for the different positions in the field of view, with 0.079 in the top left corner ($Sd = 0.038$, $\min = 0.006$, $\max = 0.159$), 0.152 in the bottom left corner ($Sd = 0.054$, $\min = 0.041$, $\max = 0.271$), 0.106 in the bottom right corner ($Sd = 0.048$, $\min = 0.02$, $\max = 0.194$), and 0.127 in the top right corner ($Sd = 0.044$, $\min = 0.046$, $\max = 0.215$).

4.5. Procedure

The study procedure is shown in Fig. 4. Experiments were conducted in a dedicated lab where the participants sat at a desk wearing the headset. The study consisted of two parts.

Firstly, participants performed a flickering calibration following a staircase procedure to identify a subtle intensity level for each participant (see screenshot in Fig. 5). Participants were instructed to look at a fixation cross in the center of their field of view. In each corner of the screen one object was presented. The calibration then successively applied flicker to each of the objects. The participants' task was to answer the question "Did you notice flickering?" with "No" or "Yes" by using the up and down arrows on a keyboard with up representing "No" and down representing "Yes". Then, the intensity of the flickering would be changed accordingly (increased if the answer was "No" and decreased if the answer was "Yes") and the procedure was repeated for the same object. After participants had switched between answering "Yes" and "No" eight times for the same object, the flickering modulation would be applied to an object in another location. The calibration ended after going through eight different flickering objects, varying in position, color, and shape.

Secondly, participants practiced the foraging task in one test trial (see Fig. 6). After the practice trial, participants completed 60 trials in total. They were informed about the target/distractor combination before each trial. They could take a break at any time during the flicker calibration and between trials in the foraging part.

4.6. Independent variable

We conducted a within-participant user study with four conditions. The study included a baseline condition where no chaining was applied, and one condition for each of the chaining variants, resulting in four conditions.

We have one independent variable *chaining type* with the four levels *none* (baseline condition), *close chaining*, *distant chaining*, and *random chaining*. Each condition had 15 trials, resulting in 4× trials in total. Trials were randomized. We abstained from a block design to avoid habituation effects.

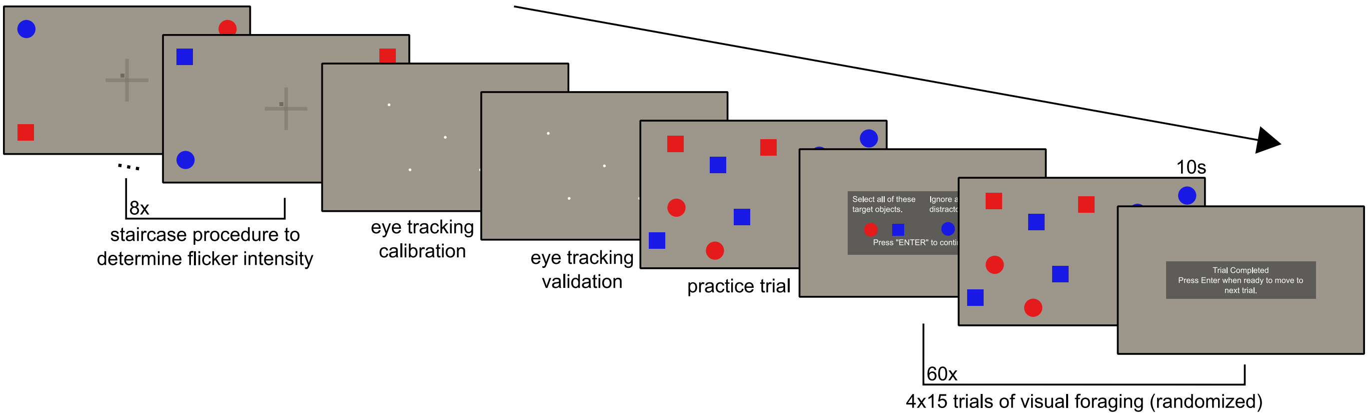


Fig. 4. Study procedure.

4.7. Dependent variables

We pick dependent measures from the foraging literature (Kristjánsson et al., 2020a) to analyze the effectiveness of the chaining techniques in influencing viewing order and their impact on overall performance. We determine the *effectiveness* of the three chaining variants, by analyzing how they interact with run behavior. We assess run behavior through run length and run number. A run is defined as the consecutive selection of the same target type (e.g., selecting two red sphere targets consecutively would describe a run of length two. Selecting a red sphere target and then a blue cube target would describe two runs, each having a length of one). Runs are thus defined by the consecutive selection of *repeated* target types. Selecting a target of the other target type is noted as *switching*. In difficult foraging tasks, such as conjunction foraging, participants typically switch less between target types (Kristjánsson et al., 2014). The effectiveness of the chaining techniques can be assessed by examining how they support or disrupt the selection of repeated and switched targets. We assess *performance* through collection rate, inter-target times (ITTs), and inter-target lengths (ITLs). Collection rate describes the number of selected targets per second, ITTs describe the time that has passed between two successive target selections, and ITLs describe the spatial distance between two successively selected targets.

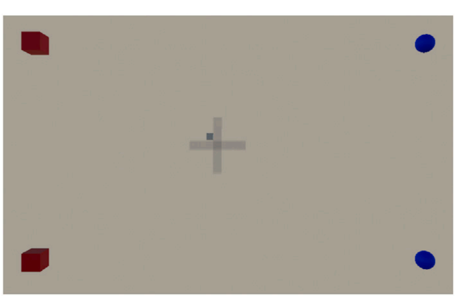


Fig. 5. Flickering calibration task.

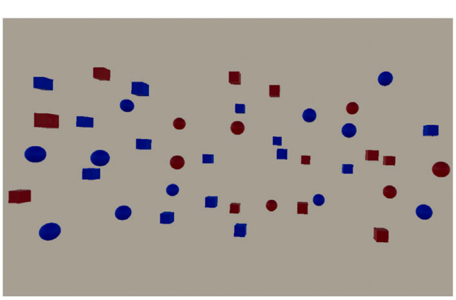


Fig. 6. Foraging task.

4.8. Apparatus

We used a Varjo XR-3 headset with a horizontal field of view of 115°, vertical field of view 90°, and refresh rate 90 Hz. The Varjo headset has a focus display (i.e., Bionic Display) with a focus area of 27° × 27° at 70 PPD uOLED and 1920 × 1920 px resolution per eye, and a peripheral display with 30 PPD LCD and 2880 × 2720 px per eye. We used a PC with an AMD Ryzen 5800X, 64GB RAM, 2TB + 4TB storage, and an Nvidia RTX 3080. The Varjo XR-3 headset has built-in eye tracking with a gaze camera resolution of 640 × 400 px per camera. The gaze tracking sample rate is 200 Hz for the Unity XR SDK. We developed the study application with Unity 3D version 2021.3.25f1 and the Varjo XR plugin is version 3.5.0.

4.9. Virtual environment

The VR environment consisted of a plain gray background (RGB: 119, 119, 119), similar to the ones used by Kristjánsson et al. (2022). Shadows were turned off and no visual depth cues were visible. The background was always the same color (see Fig. 3a).

5. Results

5.1. Data analysis

Effectiveness was analyzed using run lengths computed per trial and averaged across trials for each participant and condition. Performance was analyzed per target, with target-related measures averaged across trials for each participant and condition.

We analyzed the data using generalized linear mixed models (GLMMs),[3] with participants as random intercepts to account for repeated measures. Since we used the same random intercepts for the models, we only report the terms we included as fixed effects. Before fitting the models, we analyzed the data distribution.[4] We chose a link function based on the best-fitting distribution, typically log-normal, leading us to use the gamma family with log or inverse link functions. All models were fitted using maximum likelihood estimation with Laplace approximation. Models were tested against simpler versions, and better-fitting simpler models were chosen. Full model specifications are given in the appendix in Appendix A. For better readability, we do not report

[3] We used the `glmer` R package https://www.rdocumentation.org/packages/lme4/versions/1.1-35.5/topics/glmer.

[4] We compared distribution fits using Akaike's Information Criterion and Bayesian Information Criterion.

all details for individual models in the results section, but refer the reader to the appendix for details. Pairwise comparisons were based on estimated marginal means for each level of independent variables, using the Tukey adjustment.

For the analysis of collection rate, we conducted paired t-tests after confirming with the Shapiro-Wilk test that the data were normally distributed. For the cases where the data followed non-parametric distributions, we performed Wilcoxon-signed rank tests.

5.2. Eye tracking quality

We determined eye tracking quality by first calculating the accuracy for the x- and y-direction separately as the absolute distance of each validation point to the mean of the gaze points recorded during validation $M_{Acc_{x/y}}$ (300 gaze points per validation point, 5 validation points). We then determined the visual angle $\theta_{VA_{x/y}}$ as follows, using the depth value of $1.5m$ as the distance value:

$$\theta_{VA_{x/y}} = 2 * arctan\left(\frac{M_{Acc_{x/y}}}{2 * 1.5}\right) * \frac{180}{\pi} \tag{4}$$

We calculated each value for the x- and y-direction, resulting in $\theta_{VA_x} = 0.91°$ ($Sd = 0.70$) and $\theta_{V_y} = 0.79°$ ($Sd = 0.36$), with an overall value of $\theta_{VA} = 0.85°$. We report the precision using the Root Mean Square Error (RMSE). We achieved a value of $RMSE_x = 0.36$ ($Sd = 0.39$) and $RMSE_y = 2.13$ ($Sd = 0.32$), resulting in an overall $RMSE$ of 1.25. We excluded one participant from the study because of very low eye tracking quality ($\theta_{VA} = 30.8°$). Therefore, we report the results of 35 participants.

5.3. Summary of main findings

The findings are grouped into two parts, focusing on (1) the *Effectiveness* of the chaining variants analyzed through run behavior and (2) the chaining variants' influence on *Performance* analyzed through target selection behavior. The impatient reader may read the summary below and then skip directly to the discussion.

Effectiveness – run behavior

1. The *close chaining* variant was most effective, producing significantly longer flicker-supported runs (runs that contain at least one flickering target) than flicker-unsupported runs (runs that do not contain flickering targets) (see Fig. 7).
2. With chaining techniques, participants selected targets faster and at shorter distances, as demonstrated through lower ITTs and ITLs for flicker-supported runs compared to flicker-unsupported runs (see Fig. 8).
3. Effectiveness varied by technique. Participants collected the most flickering targets within one run with *close chaining* (1.6), followed by *random chaining* (1.18), and *distant chaining* (1.01).
4. Chaining was most effective when targets were found mid-run, yielding the highest run length. Interestingly, runs starting or ending with a flickering target had the same length.
5. Chaining effectiveness is also demonstrated by increasingly cognitively demanding selections. Flickering increased the likelihood to switching of the other target type when already following the flicker.

Performance – target selection

1. Collection rate was constant across conditions (11.9 targets/trial; 1.19 targets/second).
2. Most flickering targets were selected in *close chaining* (44%), followed by *random* (15%), and *distant* (9%) (see Fig. 9a). Note that we calibrated to a 50% detection threshold and would expect rates close to the close one without considering the impacts of the task.
3. Flickering increased switching, as indicated by a smaller ratio of repeated to switched target selections during *close chaining*, along with faster and shorter-distance selections (see Fig. 10). Participants were also significantly faster and selected targets of shorter distance from each other when selecting flickering targets with *close chaining* compared to the baseline condition ((see Figs. 10 and 11)).
4. Participants consistently selected more repeated targets (7.26) than switched targets (3.07), which aligns with typical foraging behavior.

5.4. Effectiveness

We evaluated the effectiveness of the three chaining variants by analyzing the run behavior, focusing the analysis on *run length*, *run number*, *inter-target times (ITT)* and *inter-target lengths (ITL)* as outcome variables. Table 1 provides summary statistics for all four conditions.

Starting with the analysis, we marked runs as "flicker-supported runs" when at least one target was flickering and as "flicker-unsupported runs" when none of the targets in the run were flickering. We focused this analysis on the *close chaining* technique, since this technique produced the most flicker-supported runs. For the statistical analysis, we created a new variable named "run category". As we focused the analysis on the *close* and *baseline* conditions, we have three values for "run category": *none* (baseline), *flicker-supported close chaining*, and

Table 1
Summary statistics for runs: grouped by condition.

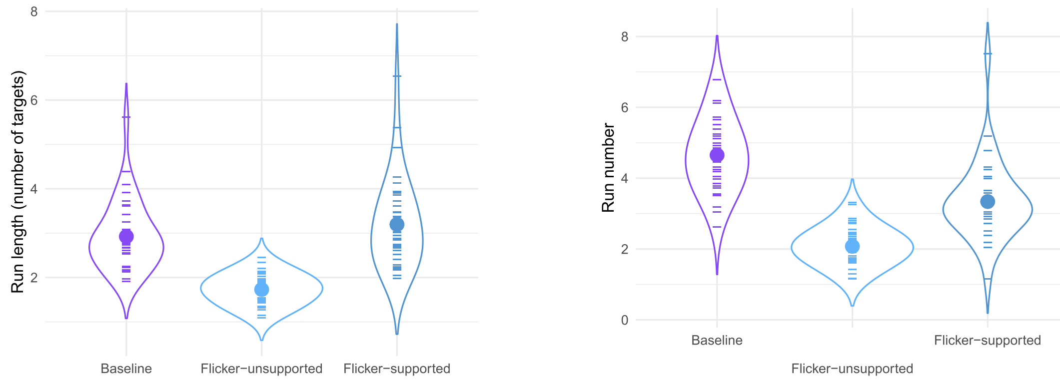
Chaining Variant	N	Run Length	Run Number	ITT (ms)	ITL (m)
None (baseline)	35	$M = 2.92$ ($Sd = 0.77$)	$M = 4.65$ ($Sd = 1.03$)	$M = 766$ ($Sd = 219$)	$M = 0.90$ ($Sd = 0.08$)
Close chaining	35	$M = 2.79$ ($Sd = 0.90$)	$M = 4.91$ ($Sd = 1.39$)	$M = 787$ ($Sd = 234$)	$M = 0.90$ ($Sd = 0.08$)
Distant chaining	35	$M = 2.93$ ($Sd = 0.82$)	$M = 4.65$ ($Sd = 1.11$)	$M = 776$ ($Sd = 215$)	$M = 0.91$ ($Sd = 0.10$)
Random chaining	35	$M = 2.94$ ($Sd = 0.83$)	$M = 4.62$ ($Sd = 1.22$)	$M = 805$ ($Sd = 256$)	$M = 0.92$ ($Sd = 0.09$)

Note. ITT stand for inter-target times, ITL stands for inter-target lengths.

Table 2
Summary statistics for runs: grouped by run category.

Run Category	N	Run Length	Run Number	ITT (ms)	ITL (m)
None (baseline)	35	$M = 2.92$ ($Sd = 0.77$)	$M = 4.65$ ($Sd = 1.03$)	$M = 766$ ($Sd = 219$)	$M = 0.89$ ($Sd = 0.08$)
Flicker-unsupported	35	$M = 1.73$ ($Sd = 0.35$)	$M = 2.08$ ($Sd = 0.53$)	$M = 881$ ($Sd = 296$)	$M = 0.98$ ($Sd = 0.17$)
Flicker-supported	35	$M = 3.20$ ($Sd = 0.99$)	$M = 3.34$ ($Sd = 1.10$)	$M = 731$ ($Sd = 168$)	$M = 0.85$ ($Sd = 0.07$)

Note. ITT stand for inter-target times, ITL stands for inter-target lengths. Flicker-supported and flicker-unsupported runs were analyzed for *close chaining* only.



(a) Run length. (b) Run number.

Fig. 7. Run length and run number for flicker-supported and flicker-unsupported runs compared to the baseline.

flicker-unsupported close chaining. Table 2 shows summary statistics for the runs grouped by *run category*.

Run length. We analyzed the effect of *run category* on run length using a GLMM. Fig. 7 shows the mean run length and run number grouped by *run category*.

The GLMM suggested that run length is different for all three run categories, with pairwise comparisons confirming that run length is longest for flicker-supported runs in the close condition (3.2 targets), followed by runs in the baseline condition (2.92 targets), and flicker-unsupported runs in the close condition (1.73 targets). The full model specifications and statistical results are shown in Appendix A.1.1.

Run number. The model showed significant differences between baseline, flicker-supported, and flicker-unsupported runs (see Fig. 7b), with the highest run number in baseline (4.65), followed by flicker-supported (3.34) and flicker-unsupported runs (2.08). Full model specifications and statistical details are given in Appendix A.1.2.

ITTs. Participants were significantly faster in selecting targets in flicker-supported runs ($M_{ITT} = 731ms$) than in flicker-unsupported runs ($M_{ITT} = 881ms$). They were also significantly slower in selecting targets in flicker-unsupported runs than in selecting targets in the baseline condition ($M_{ITT} = 766ms$). Fig. 8(a) visualizes these differences.

ITLs. The distance between two successive target selections was significantly shorter during flicker-supported runs ($M_{ITL} = 0.853m$), compared to runs in the baseline condition ($M_{ITL} = 0.894$), and flicker-unsupported runs ($M_{ITL} = 0.983$) (see Fig. 8b). See Appendix A.1.3 ITT and Appendix A.1.4 ITL for model specifications.

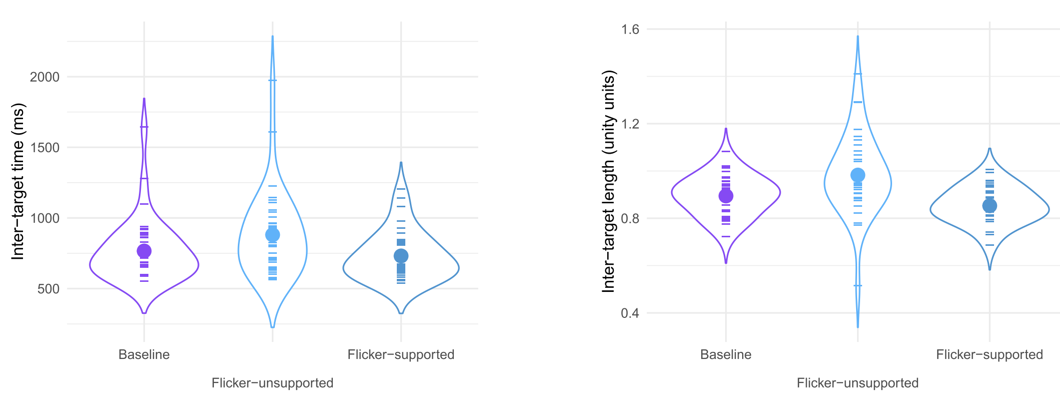
5.4.1. Analysis of flicker-supported runs

To understand the effectiveness of the different conditions on flicker-supported runs in more detail, we next examined flicker-supported runs more closely.

Number of flickering targets in flicker-supported runs. Participants collected the most flickering targets per run during *close chaining* (1.6 targets), followed by *random chaining* (1.18 targets), and *distant chaining* (1.01 targets). Summary statistics are given in Table 3. A GLMM and pairwise comparisons confirmed that the most flickering targets per run were found in the *close chaining* variant. See Appendix A.2.1 for full model specifications and statistical results.

Position of flickering target in a flicker-supported run. For flicker-supported runs, we analyzed *where* in the run the flickering occurred and grouped runs into “first position”, “middle position”, and “last position” runs. Table 4 shows summary statistics.

We fitted a GLMM with *flicker position* (non-flicker, start, middle, end) as a fixed effect and run count per participant as the outcome. Pairwise comparisons showed significant differences for all levels. Compared to non-flicker runs, participants had more runs at the end



(a) Inter-target times. (b) Inter-target lengths.

Fig. 8. ITTs and ITLs for flicker-supported and flicker-unsupported runs compared to baseline.

Table 3
Summary statistics for flicker-supported and flicker-unsupported runs.

Chaining Variant	Flicker	N	Run Length	Run Number	ITT (ms)	ITL (m)
Close Chaining	No	35	$M = 1.72$ ($Sd = 0.35$)	$M = 2.09$ ($Sd = 0.53$)	$M = 0.88$ ($Sd = 0.29$)	$M = 0.98$ ($Sd = 0.16$)
Close Chaining	Yes	35	$M = 3.18$ ($Sd = 0.99$)	$M = 3.35$ ($Sd = 1.08$)	$M = 0.73$ ($Sd = 0.17$)	$M = 0.86$ ($Sd = 0.07$)
Distant Chaining	No	35	$M = 2.87$ ($Sd = 0.82$)	$M = 4.64$ ($Sd = 1.13$)	$M = 0.77$ ($Sd = 0.21$)	$M = 0.90$ ($Sd = 0.09$)
Distant Chaining	Yes	17	$M = 3.78$ ($Sd = 2.51$)	$M = 1.04$ ($Sd = 0.11$)	$M = 0.81$ ($Sd = 0.22$)	$M = 1.42$ ($Sd = 0.37$)
Random Chaining	No	35	$M = 2.71$ ($Sd = 0.66$)	$M = 3.71$ ($Sd = 1.03$)	$M = 0.81$ ($Sd = 0.27$)	$M = 0.90$ ($Sd = 0.09$)
Random Chaining	Yes	35	$M = 3.41$ ($Sd = 1.00$)	$M = 1.54$ ($Sd = 0.42$)	$M = 0.72$ ($Sd = 0.17$)	$M = 0.97$ ($Sd = 0.13$)

Note. Flicker-supported and -unsupported runs are indicated by the column “Flicker”. ITT stands for inter-target time, and ITL for inter-target length.

Table 4
Summary statistics for position of flickering target.

Flicker Position	N	Run Number	Run Length	No of Flickering Targets	ITT (ms)
Non-Flickering	35	$M = 2.08$ ($Sd = 0.53$)	$M = 1.73$ ($Sd = 0.35$)	–	$M = 881$ ($Sd = 296$)
Start Position	35	$M = 2.40$ ($Sd = 1.08$)	$M = 2.47$ ($Sd = 0.64$)	$M = 1.72$ ($Sd = 0.29$)	$M = 745$ ($Sd = 169$)
Middle Position	35	$M = 1.28$ ($Sd = 0.192$)	$M = 4.59$ ($Sd = 1.10$)	$M = 1.81$ ($Sd = 0.39$)	$M = 666$ ($Sd = 134$)
End Position	35	$M = 1.25$ ($Sd = 0.219$)	$M = 2.71$ ($Sd = 0.48$)	$M = 1.00$ ($Sd = 0.00$)	$M = 707$ ($Sd = 304$)

Note. Position indicates where in the run a flickering target was selected the first time.

(×1.37) and middle (×1.34), and fewer runs at the start (×0.94). Results are based on geometric means rather than arithmetic means to reduce the influence of extreme values. Full model specifications are in Appendix A.2.2.

We compared run length by flicker position within a run. Runs were longest when the first flickering target appeared in the middle (+2.64 targets), compared to smaller increases at the start (+1.41) and end (+1.58) relative to non-flicker runs. The middle condition was also significantly longer than the start and end conditions. These effects were confirmed by a GLMM with *flicker position* as a fixed effect, indicating that mid-run flicker allows participants to collect significantly more targets. Full model details are in Appendix A.2.3.

Comparing ITTs across run categories, participants were fastest in flickering runs of the middle category (183ms faster than non-flicker runs), which were also significantly faster than the start and end categories. Full GLMM details are provided in Appendix A.2.4.

Number of flickering targets ignored during a run. Finally, we analyzed ignored flickering targets. In flicker-supported runs, 45.8% of ignored targets matched the current type and would have prolonged the run; 54.2% were of a different type and would have canceled the run. In flicker-unsupported runs, 38.3% matched and 61.7% differed. This suggests participants were more likely to switch target type and cancel a run when following the flicker. The difference was significant ($t = 2.82$, $df = 34$, $p = 0.0079$, Cohen’s $d = 0.48$).

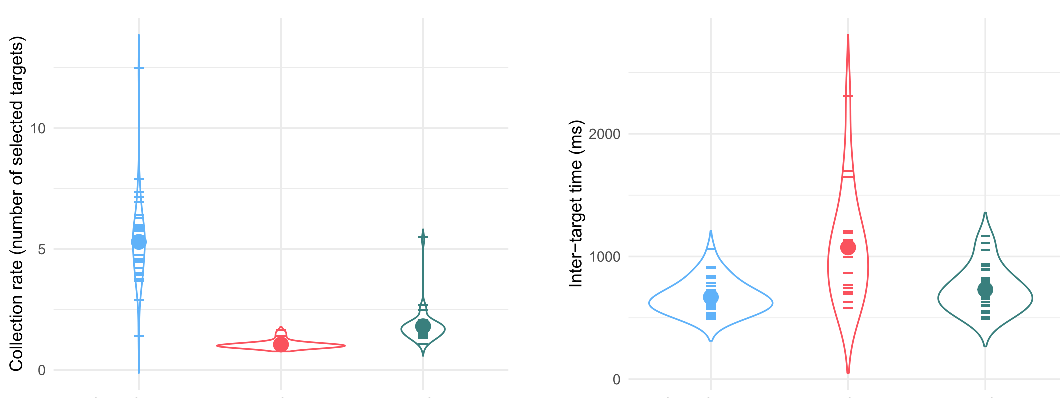
5.5. Performance

5.5.1. General collection rate

The number of collected targets was consistent across conditions ($M = 11.9$ per trial, 1.19 per second; see Table 5). Paired t-tests showed no significant differences between baseline and any chaining condition (Close: $t(34) = 0.09$, $p = 0.93$; Distant: $t(34) = 0.26$, $p = 0.80$; Random: $t(34) = 0.14$, $p = 0.89$), indicating no effect of chaining variant on collection rate. ITTs were similarly unaffected, as confirmed by Wilcoxon signed rank tests (Close: $W = 247$, $p = 0.27$; Distant: $W = 321$, $p = 0.93$; Random: $W = 223$, $p = 0.14$).

Influence of flicker modulation on collection rate. During *close chaining*, 44% of selected targets were modulated, compared to 9% in *distant chaining* and 15% in *random chaining* (see Fig. 9a). A GLMM confirmed these differences (details in Appendix A.3.1). Refer to columns N_F, $CollectionRate_F$, ITT_F in Table 5 for summary statistics.

ITTs also differed significantly (see Fig. 9b). Participants were fastest in *close chaining* ($M = 669ms$, $Sd = 133$), followed by *random* ($M = 730ms$, $Sd = 181$), and *distant chaining* ($M = 1074ms$, $Sd = 451$). See Appendix A.3.2 for model details.



(a) Collection Rate (b) Inter-target times

Fig. 9. Average number of flickering targets selected and ITTs grouped by chaining technique.

Table 5
Summary statistics for collection rate and ITTs.

Chaining Variant	N_G	Collection $Rate_G$	ITT_G (ms)	N_G	Collection $Rate_F$	ITT_F (ms)
None (Baseline)	35	$M = 11.9$ ($Sd = 2.50$)	$M = 753$ $Sd = 216$	–	–	–
Close Chaining	35	$M = 11.9$ ($Sd = 2.63$)	$M = 767$ $Sd = 221$	35	$M = 5.29$ ($Sd = 1.79$) (44%)	$M = 669$ $Sd = 133$
Distant Chaining	35	$M = 11.9$ ($Sd = 2.64$)	$M = 755$ $Sd = 217$	17	$M = 1.05$ ($Sd = 0.15$) (9%)	$M = 1074$ $Sd = 451$
Random Chaining	35	$M = 11.9$ ($Sd = 2.59$)	$M = 778$ $Sd = 250$	35	$M = 1.80$ ($Sd = 0.72$) (15%)	$M = 730$ $Sd = 181$

Note. G indicates *general*, *F* indicates *flickering targets*, ITT stands for inter-target times.

Table 6
Summary statistics for collection rate: grouped by flickering/non-flickerng targets and target type.

Chaining-Flicker	Target Type	N	Collection Rate	ITT (ms)	ITL (m)
None (baseline)	Repeated	35	$M = 7.32$ ($Sd = 2.08$)	$M = 706$ ($Sd = 196$)	$M = 0.97$ ($Sd = 0.11$)
None (baseline)	Switched	35	$M = 2.97$ ($Sd = 0.90$)	$M = 816$ ($Sd = 265$)	$M = 1.00$ ($Sd = 0.10$)
Close chaining non-flickering	Repeated	35	$M = 4.05$ ($Sd = 1.54$)	$M = 718$ ($Sd = 187$)	$M = 1.16$ ($Sd = 0.11$)
Close chaining non-flickering	Switched	35	$M = 1.94$ ($Sd = 0.49$)	$M = 993$ ($Sd = 385$)	$M = 1.27$ ($Sd = 0.14$)
Close chaining flickering	Repeated	35	$M = 3.34$ ($Sd = 0.89$)	$M = 623$ ($Sd = 129$)	$M = 0.71$ ($Sd = 0.04$)
Close chaining flickering	Switched	34	$M = 2.09$ ($Sd = 0.95$)	$M = 747$ ($Sd = 184$)	$M = 0.76$ ($Sd = 0.07$)

Note. ITT stands for inter-target time, ITL stand for inter-target length.

Influence of target type on collection rate. When looking at *what type* of target participants collected, we find, as anticipated, that they collected significantly more targets of *repeated* target type ($M = 7.26$, $Sd = 2.08$) than of *switched* target type ($M = 3.07$, $Sd = 1.03$), confirmed by a Wilcoxon signed rank test ($W = 9724$, $p = 1.70e-24$, $r = 0.866$).

5.5.2. Interaction effects of flicker modulation and foraging strategies on collection rate

Previous results describe *which* targets were selected (e.g., flickering vs. non-flickering, repeated vs. switched). Next, we examine how chaining variants interacted with participants' *foraging strategies*.

Interaction effect of flickering and target type on collection rate. We analyzed how selecting *flickering* and *non-flickering* targets interacts with *target type* (repeated/switched). As *close chaining* was most effective and provided the most data, we focus on comparing flickering vs. non-flickering targets in this condition against the baseline. We modeled this using a combined factor (*chaining-flicker*) with three levels: *baseline*, *close-chaining-flickering*, and *close-chaining-nonflickering*. See Table 6 for summary statistics.

We fitted a GLMM with *chaining-flicker* and *target type* as fixed effects. As anticipated, the model confirmed fewer switches than repeated targets and fewer flickering and non-flickering targets compared to baseline (see Appendix A.4.1) for full model specifications and statistical results.

The most interesting finding is an interaction effect, suggesting that the ratio between the number of repeated and switched targets was significantly lower in the *close condition* than in baseline (see Fig. 10), with significant differences between *baseline* and *close chaining - flickering targets* and between *close chaining - flickering targets* and *close chaining - non-flickering targets*.

For ITTs (see Fig. 11a), participants were significantly faster when selecting flickering targets in *close chaining* than non-flickering targets in the baseline, but slower for non-flickering targets in *close chaining*. Switched targets again showed longer ITTs. Full model details are in Appendix A.4.2.

For ITLs (see Fig. 11b), flickering targets in *close chaining* were selected at shorter distances than in the baseline, while non-flickering targets were selected at longer distances. Full model specifications are in Appendix A.4.2.

Interaction effect of chaining variant and target type on collection rate. We analyzed the effects of *chaining variant* (close, distant, random) and *target type* (repeated, switched) on collection rate (see Table 7). A GLMM confirmed fewer switched than repeated targets across all variants, but no main effect of *chaining variant* (see Appendix A.4).

6. Discussion

The results show that SGD chaining affects both run length and target selection, with implications for the design of future guidance techniques.

The primary finding is that chaining SGD cues effectively guides gaze to multiple targets by prolonging runs during visual foraging. Despite using subtle flickering modulation (using a 50% detection threshold), participants followed the cues. This is specifically the case for the *close chaining* variant, where participants selected modulated targets in 44%

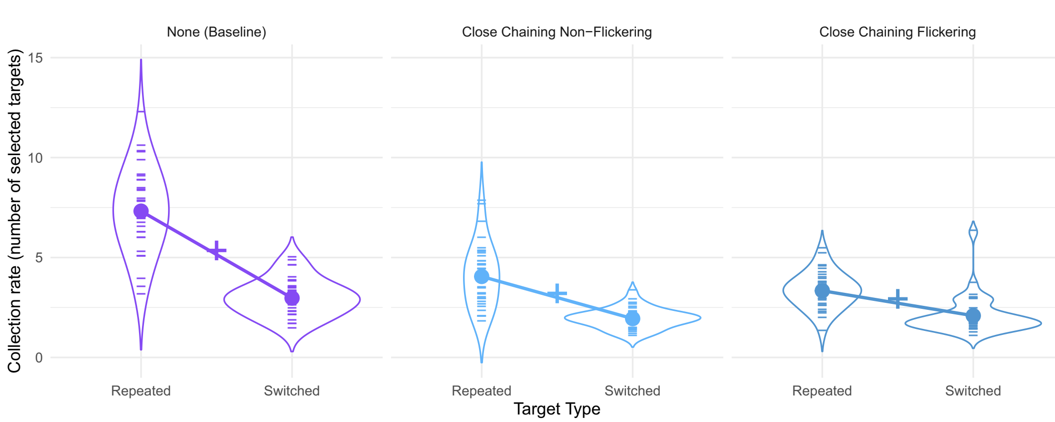


Fig. 10. Collection rate ratio between targets of repeated and switched target types.

(a) Inter-target times (ITTs)

(b) Inter-target lengths (ITLs)

Fig. 11. ITTs and ITLs for repeated and switched targets comparing flickering and non-flickering targets against the baseline.

Table 7
Summary statistics for collection rate: grouped by chaining variant and target type.

Chaining Variant	Target Type	N	Collection Rate	ITT (ms)	ITL (m)
None (Baseline)	Repeated	35	$M = 7.32$ ($Sd = 2.08$)	$M = 706$ ($Sd = 196$)	$M = 0.97$ ($Sd = 0.11$)
None (Baseline)	Switched	35	$M = 2.97$ ($Sd = 0.90$)	$M = 816$ ($Sd = 265$)	$M = 1.00$ ($Sd = 0.10$)
Close Chaining	Repeated	35	$M = 7.06$ ($Sd = 2.17$)	$M = 689$ ($Sd = 177$)	$M = 0.97$ ($Sd = 0.10$)
Close Chaining	Switched	35	$M = 3.27$ ($Sd = 1.24$)	$M = 899$ ($Sd = 357$)	$M = 1.03$ ($Sd = 0.17$)
Distant Chaining	Repeated	35	$M = 7.32$ ($Sd = 2.14$)	$M = 692$ ($Sd = 173$)	$M = 0.98$ ($Sd = 0.10$)
Distant Chaining	Switched	35	$M = 3.04$ ($Sd = 0.97$)	$M = 852$ ($Sd = 352$)	$M = 1.04$ ($Sd = 0.20$)
Random Chaining	Repeated	35	$M = 7.36$ ($Sd = 2.02$)	$M = 702$ ($Sd = 184$)	$M = 0.98$ ($Sd = 0.08$)
Random Chaining	Switched	35	$M = 2.99$ ($Sd = 1.02$)	$M = 928$ ($Sd = 362$)	$M = 1.06$ ($Sd = 0.17$)

Note. ITT stands for inter-target time, ITL stand for inter-target length.

of selections. Furthermore, chaining not only increases the likelihood of directing gaze but also does so faster than unguided search.

6.1. What about finding your keys, learning to drive, and following Magnus Carlsen's gaze?

To illustrate our findings practically, we refer to the examples envisioned in the Introduction.

Fig. 12 illustrates the effect of the technique on the chess example.

For the object searching example (which we envision to be a combination of close and distant chaining), *random chaining* would draw your gaze to relevant objects by 15%. In this example, a person would be searching for relevant objects before leaving the house, which is beneficial for the technique as it prolongs runs that support a person's strategies.

Finally, referring back to the driving lessons, chaining subtle gaze cues would increase the likelihood of looking at vulnerable road users in a specific order. In this scenario, flickering would be applied in the periphery, since it would scan for people on sidewalks or crossing the street from the side. Peripheral flickering was effective 9% of the time, suggesting that it could redirect one's gaze, while allowing for enough agency and self-determined observation to avoid making you "dependent" on guidance to recognize critical situations.

To clarify, our findings suggest the likelihood of supporting experiences such as those we discussed. Given that we conducted a controlled study, the ideas sketched out above are aspirational, not fully demonstrated by the study's results. The given examples involve complex cognitive functions, not only visual search, but also memory and spatial reasoning, for example. To start realizing them, we need to begin at the most low-level cognitive process, which is visual perception. Our study showed that we can redirect a person's gaze by influencing low-level processing through visually manipulating the search targets. To realize the examples, the next steps should involve incorporating more cognitive processes in the study task and observing how well the influence on visual search transfers to those tasks. Yet, it is not only desirable but necessary to start with the most fundamental cognitive process (visual perception).

Next, we discuss our findings on a more concrete level.

6.2. Chaining techniques affect the collection rate of flickering targets only

While chaining did not significantly increase overall performance as we initially expected, it is important to note that participants were already highly proficient, taking just over a second to select a target. This proficiency likely contributed to the consistent ITTs across all conditions. The only other foraging study using eye gaze as a selection mechanism employed a dwell time of 100 ms to select targets (Jóhannesson et al., 2016), which is not directly comparable to our point-and-click selection. A task with harder-to-identify targets, either due to more distractor objects, more conjunction features, or fewer targets, might reveal stronger effects.

We observed differences across the three *chaining variants* in collection rate and ITTs. Participants collected more flickering targets and were faster with *close chaining*, while *distant chaining* led to the fewest targets and longest times due to larger movement costs. Despite this, even the most challenging chaining technique was effective in 9%. These results suggest that the influence of flicker is limited by the cost of performing a motor response (i.e., turning one's head). Therefore, when designing flicker-based chaining techniques, the position of the target in a person's foveal or peripheral view should be factored in.

While we calibrated for a 50% detection threshold, we only saw close to this value for the close condition. We take this as an indication that the task-driven allocation of attention impacted the selection of flickering cues, and incongruence of the random and distant chaining with natural patterns led to large effects here. Whether this means a higher detection threshold is facilitated by the task, or the detection remains the same but a stronger effect is required to elicit a response to the cue requires further investigation.

6.3. Chaining strongly interacts with run behavior

We observed strong interactions between *chaining techniques* and run behavior. Consequently, the effects of chaining need to be seen in the context of foraging behavior. Supporting participants' strategies (i.e., in runs/when selecting repeated targets) increased run length and sped up

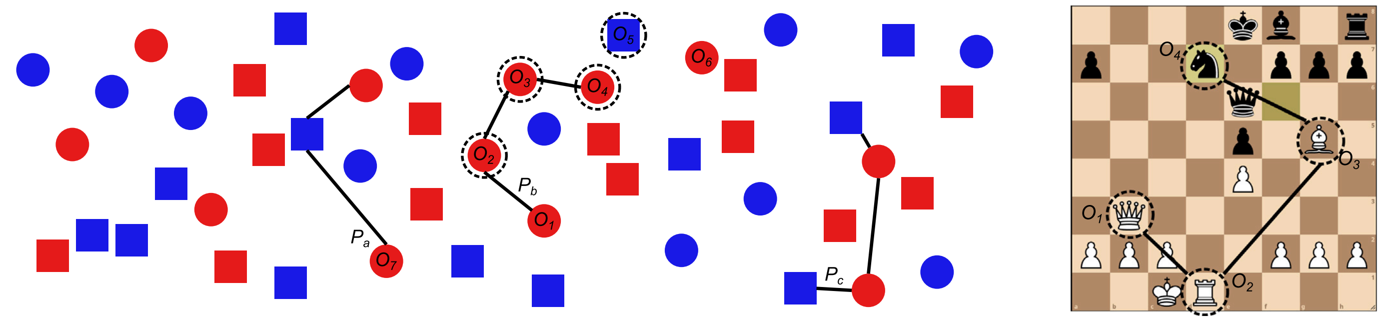


Fig. 12. Practical illustration of the effects in the results applied to the chess example. **On the left,** three potential selection paths are shown. Assuming red spheres and blue cubes are targets and sequential application of flicker (dotted lines), path P_b is most likely when flicker occurs mid-run (e.g., on O_2 and O_3). The length is representative of our results, as for mid-run flickering runs, the average length was 4.59 targets. After selecting O_4, users are more likely to switch to O_5 than O_6, initiating a new sequence with a different target type. Paths P_a and P_c are less likely, as they involve more frequent switches without a preceding run. Similarly, from O_1, users are more likely to select the closer flickering target O_2 than the distant O_7. **On the right**, an illustration of how chaining could be applied practically is shown. Sequential flickering (indicated through dotted lines) increases the likelihood that users follow this pattern. This is particularly effective for intermediate players, as mid-run cues reinforce attention on correct pieces. In this situation, *close chaining* would prolong runs and influence a player to follow Magnus' gaze more accurately (and for longer). Additionally, the findings indicate that when a person is ready to shift focus to a new target, flickering a significant piece increases the likelihood of initiating a new sequence with more task-relevant targets (based on the decreased ratio of repeated/switched targets with chaining). Furthermore, *close chaining* would speed up the guidance to relevant pieces.

target selection. Additionally, the reduced ITLs indicate that participants did not engage in searching for more distant targets.

Flickering also facilitated switching between target types, suggesting reduced switch costs and improved guidance to new target sets when a patch is exhausted. Since switching is typically costly, this may yield cumulative benefits in longer tasks. Our findings align with prior work showing faster switching under time constraints (Kristjánsson et al., 2018), and suggest that flickering further enhances this effect.

When participants encountered a flickering target during a run (and it was not the last), they significantly prolonged the run and were more likely to select subsequent flickering targets, showing that chaining is effective when it supports the current search strategy. These findings highlight an interaction between top-down (task-driven) and bottom-up (flicker-driven) attention. Task-driven processes appeared stronger, as participants often followed pre-planned selection paths that were not overridden by flicker. However, flicker was most effective when guiding attention to new targets that had not yet been identified. It is important to consider here our premise of chaining guidance cues. Rather than increasing competence at a given task (e.g., foraging), we intended to affect gaze behavior in a controllable manner. While sequential flicker can influence how users explore a scene, the strong role of top-down processes suggests that guidance should be aligned with users' existing strategies.

6.4. What do the results mean for continuous gaze guidance?

In the following, we discuss the results' significance for continuous gaze guidance considering how likely the technique is to influence a person's viewing order. The prerequisite for our technique was to develop a subtle technique, calibrating subtly through a psychophysics staircase procedure to 50%. *close chaining* guided 44% of selections, close to the expected optimum. Even *distant chaining* achieved a 9% rate notable given the cognitive and physical effort (e.g., head movement) required to follow peripheral cues. Naturally, subtle cues will always involve a trade off between effectiveness and noticeability, which has been confirmed by prior work (Sutton et al., 2022, 2024; Grogorick and Magnor, 2020).

For *distant chaining*, the collection rate could likely be increased by trading some of the subtlety for noticeability. However, this would shift the technique beyond our intended design goals. Potential reasons for its lower effectiveness may stem from higher cognitive and motor costs or visual artifacts in the VR HMD that could have interacted with the flickering stimulus (e.g., reduced peripheral resolution in the Varjo XR headset). Furthermore, we reiterate that we used a cognitively demanding search task including both conjunction search (Egeth et al., 1984) (i.e., a target was defined by two features) and hybrid foraging (Wolfe et al., 2019) (i.e., the task involved finding multiple target objects among multiple distractor objects). Participants were less likely to be influenced by a stimulus in their periphery, which would have come at a greater cost to move away from their current strategy – an effect that was less severe for the *close* and *random* techniques. It would be interesting to investigate distant chaining in a simpler task, such as feature foraging. Here, people are more likely to switch between target types and forage less systematically (Kristjánsson et al., 2022), as it requires less cognitive effort. A comparison would clarify whether the lower effectiveness of *distant chaining* stems from reduced peripheral visibility or from the cost of deviating from systematic foraging.

Our results showed chaining to be particularly successful when applied in conjunction with their current search strategy. To work in conjunction with the current gaze behavior, the ability to extend runs by placing a flicker near (temporally) to the point of transition can be used to extend behavior in a certain direction. Similarly, by utilizing flicker at transition points (such as switching) a new set of points (like a new patch) can be directed. However, top-down attention will influence where the user currently is. Therefore, rather than dictate a gaze path to a user, as initially premised, future designers and researchers should consider subtle flicker chaining as a means to promote behaviors within the current gaze behavior, and view guidance in regard to making the intended areas more appealing to the visual system (by reducing costs such as switching).

Finally, we discuss the implications of our controlled study for applied use cases. Prior work has shown that visual attention is important for typing behavior (Hu et al., 2025). For instance, visual cues on keyboards can significantly improve text entry performance, for example, by increasing typing speed through enlarging the font on keys (Magnien et al., 2004). Mutasim et al. (2025) show another approach using color to guide attention to relevant letters. These tasks can be interpreted as forms of conjunctive foraging, where multiple features (e.g., letter type and visual properties such as color or size) guide selection. The main difference in our approach is that it provides subtle, largely implicit guidance, guiding gaze without drawing conscious awareness. This distinction may help mitigate issues such as visual clutter or over-reliance on the above overt cues. Future work could directly compare existing explicit guidance techniques with chaining SGD guidance to assess differences in performance and user experience. Notably, sequential gaze guidance may be particularly effective when it supports users' ongoing search, for example by subtly highlighting keys they are already likely to select. Another practical application of our technique could be in gaze-adaptive user interfaces where interface size is adapted based on where a person is looking (Piening et al., 2021). However, interface adaptation is likely more noticeable than SGD. Additionally, while not intentionally designed for selection, there might be interesting interaction effects of SGD chaining and gaze-based selection (Sidenmark et al., 2023; Wagner et al., 2024).

An important difference to note is whether a task or use case involves an order. In our task all targets were equally relevant to users as they did not need to be collected in a specific order. However, the interactions between guidance and user strategies on effectiveness, let us make assumptions about how the technique might perform in tasks with a specific order. In text entry, for example, sequential order is highly important. Therefore, guiding users away from an intended word (e.g., because of limitations in text prediction, where a wrongly recognized letter might be highlighted), should be even harder and make the technique potentially more robust for text entry than other techniques. To confirm such results future work should investigate the technique in visual search tasks that involve an order of targets.

7. Limitations and future work

This section discusses the limitations of our work and suggests future avenues opened by it.

We found that participants chose flickering targets faster when they were presented in the central field of view than when the next flickering object was in their peripheral vision. Following the assumption of high-frequency flickering being less visible in the central field of view and more effective, we discuss what might have caused these effects. First, turning to the periphery to react to a peripheral cue costs a motor response in our study, which reduces the likelihood of reacting to it. Secondly, we suspect there might have been visual artifacts caused by the VR HMD that were interacting with the flickering, potentially making it less noticeable in the periphery than expected. Thirdly, while we expected participants to have a baseline tendency to select adjacent targets before switching to more distant ones (Kristjánsson et al., 2020a), it is interesting that this behavior is reinforced in trials with the close flickering variant. This could be the result of participants overlooking adjacent targets in the baseline, when they do not fit into their run, but the flickering could prompt them to switch. Furthermore, we focused much of our analysis on the close chaining condition because we had the most observations there, making it statistically more relevant. Future studies should collect more data to support robust analyzes for the other two variants. The study investigated the influence of chaining subtle gaze direction cues on low-level visual processing. To realize the examples in the introduction, future work should build on our findings and extend the task to more complex settings, involving more complex cognitive functions. Moreover, our study sample is specific to young university students with VR experience. While we do not expect VR experience to be a factor influencing a person's reaction to flicker, there might be other sample-related influences that limit the generalizability of our results, such as age where reaction times might be slower with increased age. Future work should investigate these influences on a larger sample, representative of the general population. Finally, we chose to calibrate to a 50% perception threshold for the stimulus subtly. We saw that with the *close chaining* variant, which works in line with search patterns, we reached effectiveness close to this threshold. However, for less congruent approaches (e.g., the *distant chaining* variant), we noted a reduction from this value. We cannot calculate the percentage of instances where participants either did not detect the stimulus or did not react to it, as our study results do not allow us to differentiate between detection and response. While increasing the detection threshold could potentially lead to a higher response rate, it would also reduce the subtlety of the perceived effect, making the guidance more overt with potential effects on the presented examples.

8. Conclusion

An important application of extended reality is to guide our visual attention in the world; in particular, we envision allowing your gaze to be guided in a manner similar to how experts see the world. To develop the groundwork for this vision, we have explored the potential of chaining subtle gaze direction cues to improve foraging in extended reality. Our results show that a close chaining variant, in which users' gaze is directed to nearby targets, can support users' search strategies and help them follow sequences of targets. They also show that gaze guidance is still fragile for complex tasks such as foraging. Ultimately, however, close chaining appears to be a good basis for developing more effective and engaging XR applications.

CRediT authorship contribution statement

Teresa Hirzle: Writing – review & editing, Writing – original draft, Visualization, Software, Project administration, Methodology, Formal analysis, Data curation, Conceptualization. **Jonathan Sutton:** Writing – review & editing, Writing – original draft, Software, Methodology, Conceptualization. **Thomas Edward Larsen:** Writing – review & editing, Writing – original draft, Software, Investigation. **Lise A. Bruun:** Writing – review & editing, Writing – original draft, Visualization, Investigation. **Kasper Hornbæk:** Writing – review & editing, Writing – original draft, Methodology, Conceptualization.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used Microsoft Copilot to improve readability and check spelling and grammar. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the article.

Funding

This research was supported by the Pioneer Center for AI, DNRF grant number P1 and by the Villum Fonden grant VIL-50108.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Results

A.1. Run behavior

A.1.1. Run length

We analyzed the effect of *run category* on run length. We fitted a GLMM model with the independent variable *run category* as a fixed effect and included *participant* as a random intercept. *Run length* was the outcome variable. The model was specified with a gamma distribution and an inverse link function. It was fitted using maximum likelihood estimation with Laplace approximation. The intercept was set to the *baseline* chaining variant. The model specifications are shown in Table A.8.

The model reported significant main effects of both the flicker-unsupported runs ($Estimate = 0.23$, $SE = 0.016$, $t - value = 14.48$, $p < 2e - 16$) and the flicker-supported runs ($Estimate = -0.027$, $SE = 0.01$, $t - value = -2.55$, $p = 0.0109$) in the close condition compared to the baseline condition. We continued with pairwise comparisons using estimated marginal means with Tukey adjustment, finding that indeed there is a significant difference between all three means of run number, also revealing a significant difference in run length between flicker-supported and flicker-unsupported runs in the close condition ($Estimate = 0.26$, $SE = 0.02$, $df = Inf$, $z.ratio = 16.55$, $p < 0.0001$). See specification in Table A.9.

Table A.8
GLMM for analysing effects of *run length category* on run length.

	Estimate	SE	t-value	Pr(>|z|)	Exp. Coef.
(Intercept)	0.37355	0.01902	19.644	<2e-16 ***	1.4528870
Close Chaining - Non-Flicker	0.22936	0.01584	14.478	<2e-16 ***	1.2577897
Close Chaining - Flicker	−0.02735	0.01074	−2.547	0.0109 *	0.9730252

Note. The model was specified as: $RunLength \sim RunLengthCateogry + (1|ParticipantID)$. The intercept is set to *None (baseline)*. $AIC = 131.6348$, $BIC = 144.9046$. We compared the model against a random intercept model specified as $RunLength \sim 1 + (1|ParticipantID)$ $AIC_{\text{simple model}} = 270.9196$ $BIC_{\text{simple model}} = 278.8815$. ANOVA between the simple and our model: $\chi^2 = 143.28, df = 2, p < 2.2e-16$ ***. Family: Gamma (inverse). Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

Table A.9
Interaction effects for *run length category* on run length.

Chaining-Flicker pairwise	Estimate	SE	df	z-ratio	p-value
Baseline - Close Non-Flickering	−0.2294	0.0158	Inf	−14.478	<0.0001
Baseline - Close Flickering	0.0273	0.0107	Inf	2.547	0.0109
Close Non-Flickering - Close Flickering	0.2567	0.0155	Inf	16.547	<0.0001

Note. Using Tukey adjustments. Results are given on the log (not the response) scale.

Table A.10
GLMM for analysing effects of *run length category* on run number.

	Estimate	SE	t-value	Pr(>| z |)	Exp. Coef.
(Intercept)	1.51062	0.05581	27.06	<2e-16 ***	4.5295274
Close Chaining - Non-Flicker	−0.80368	0.04426	−18.157	<2e-16 ***	0.4476800
Close Chaining - Flicker	−0.33923	0.04436	−7.648	2.05e-14 *	0.7123181

Note. The model was specified as: $RunNumber \sim RunLengthCateogry + (1|ParticipantID)$. The intercept is set to *None (baseline)*. $AIC = 229.5377$, $BIC = 242.8075$. We compared the model against a random intercept model specified as $RunNumber \sim 1 + (1|ParticipantID)$ $AIC_{\text{simple model}} = 361.9546$ $BIC_{\text{simple model}} = 369.9165$. ANOVA between the simple and our model: $\chi^2 = 136.42, df = 2, p < 2.2e - 16$ ***. Family: Gamma (log). Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

A.1.2. Run number

We analyzed an effect of *run category* on run number (see Fig. 7b). We also fitted a GLMM model with the independent variable *run category* as a fixed effect and included *participant* as a random intercept for *run number* as the outcome variable. The model was specified with a gamma distribution and a log link function. It was fitted using maximum likelihood estimation with Laplace approximation. The intercept was set to the *baseline* chaining variant. The model specifications are shown in Table A.10.

The model reported significant main effects of both the flicker-unsupported runs ($Estimate = -0.80$, $SE = 0.04$, $t - value = -18.16$, $p < 2e-16$) and the flicker-supported runs ($Estimate = -0.34$, $SE = 0.04$, $t - value = -7.65$, $p = 2.05e - 14$) of the close condition compared to the baseline condition. We continued with pairwise comparisons using estimated marginal means with Tukey adjustment, finding that indeed there is a significant difference between all three means of run number, also revealing a significant difference in run number between flicker-supported and flicker-unsupported runs in the close condition ($Estimate = -0.46$, $SE = 0.04$, $df = Inf$, $z.ratio = -10.41$, $p < 0.0001$). See specification in Table A.11.

Table A.11
Interaction contrasts for *run length category* on run number.

Chaining-Flicker pairwise	Estimate	SE	df	z-ratio	p-value
Baseline - Close Non-Flickering	0.804	0.0443	Inf	18.157	<0.0001
Baseline - Close Flickering	0.339	0.0444	Inf	7.648	<0.0001
Close Non-Flickering - Close Flickering	−0.464	0.0446	Inf	−10.412	<0.0001

Note. Using Tukey adjustments. Results are given on the log (not the response) scale.

Table A.12
GLMM for analysing effects of *run length category* on ITT.

	Estimate	SE	t-value	Pr(>| z |)	Exp. Coef.
(Intercept)	6.60560	0.05686	116.180	<2e-16 ***	739.2226101
Close Chaining - Non-Flicker	0.12781	0.01848	6.917	4.6e-12 ***	1.1363393
Close Chaining - Flicker	−0.03671	0.01847	−1.987	0.0469 *	0.9639532

Note. The model was specified as: $ITT \sim RunLengthCateogry + (1|ParticipantID)$. The intercept is set to *None (baseline)*. $AIC = 1238.035$, $BIC = 1251.305$. We compared the model against a random intercept model specified as $ITT \sim 1 + (1|ParticipantID)$ $AIC_{\text{simple model}} = 1290.28$ $BIC_{\text{simple model}} = 1298.242$. ANOVA between the simple and our model: $\chi^2 = 56.245, df = 2, p < 6.117e - 13$ ***. Family: Gamma (log). Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

Table A.13
Interaction contrasts for the pairwise comparisons of *run length category* on ITT.

Chaining-Flicker pairwise	Estimate	SE	df	z-ratio	p-value
Baseline - Close Non-Flickering	−0.1278	0.0185	Inf	−6.917	<0.0001
Baseline - Close Flickering	0.0367	0.0185	Inf	1.987	<0.0469
Close Non-Flickering - Close Flickering	0.1645	0.0185	Inf	8.896	<0.0001

Note. Using Tukey adjustments. Results are given on the log (not the response) scale.

A.1.3. ITT

We analyzed the effect of *run category* on inter-target times (see Fig. 8(a). We also fitted a GLMM model with the independent variable *run category* as a fixed effect and included *participant* as a random intercept for *ITT* as the outcome variable. The model was specified with a gamma distribution and a log link function. It was fitted using maximum likelihood estimation with Laplace approximation. The intercept was set to the *baseline* chaining variant. The model specifications are shown in Table A.12.

The model reported significant main effects of both the flicker-unsupported runs ($Estimate = 0.13$, $SE = 0.02$, $t - value = 6.92$, $p = 4.6e - 12$) and the flicker-supported runs ($Estimate = -0.04$, $SE = 0.02$, $t - value = -1.99$, $p = 0.05$) in the close condition compared to the baseline condition. We continued with pairwise comparisons using estimated marginal means with Tukey adjustment, finding that indeed there is a significant difference between all three means of ITT, also revealing a significant difference in ITTs between flicker-supported and flicker-unsupported runs in the close condition ($Estimate = 0.16$, $SE = 0.02$, $df = Inf$, $z.ratio = 8.90$, $p < 0.0001$). See specification in Table A.13.

Table A.14
GLMM for analysing effects of *run length category* on ITL.

	Estimate	SE	t-value	Pr(>| z |)	Exp. Coef.
(Intercept)	1.13048	0.02767	40.856	<2e-16 ***	3.0971492
Close Chaining - Non-Flicker	−0.09971	0.02355	−4.235	2.29e-05 ***	0.9050963
Close Chaining - Flicker	0.05395	0.02524	2.137	0.0326 *	1.0554336

Note. The model was specified as: $ITL \sim RunLengthCateogry + (1|ParticipantID)$. The intercept is set to *None (baseline)*. $AIC = -180.8667$, $BIC = -167.5969$. We compared the model against a random intercept model specified as $ITL \sim 1 + (1|ParticipantID)$ $AIC_{\text{simple model}} = -150.3136$ $BIC_{\text{simple model}} = -142.3517$. ANOVA between the simple and our model: $\chi^2 = 34.553, df = 2, p < 3.14e-08$ ***. Family: Gamma (inverse). Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

Table A.15
Interaction contrasts for the pairwise comparisons of *run length category* on ITL.

Chaining-Flicker pairwise	Estimate	SE	df	z-ratio	p-value
Baseline - Close Non-Flickering	0.0997	0.0235	Inf	4.235	<0.0001
Baseline - Close Flickering	−0.0540	0.0252	Inf	−2.137	0.0326
Close Non-Flickering - Close Flickering	−0.1537	0.0242	Inf	−6.354	<0.0001

Note. Using Tukey adjustments. Results are given on the log (not the response) scale.

A.1.4. ITL

We analyzed an effect of *run category* on inter-target lengths (see Fig. 8(b). We also fitted a GLMM model with the independent variable *run category* as a fixed effect and included *participant* as a random intercept for *ITL* as the outcome variable. The model was specified with a gamma distribution and an inverse link function. It was fitted using maximum likelihood estimation with Laplace approximation. The intercept was set to the *baseline* chaining variant. The model specifications are shown in Table A.14.

The model reported significant main effects of both the flicker-unsupported runs ($Estimate = -0.10$, $SE = 0.02$, $t-value = -4.24$, $p = 2.29e-05$) and the flicker-supported runs ($Estimate = 0.05$, $SE = 0.03$, $t-value = -2.14$, $p = 0.03$) of the close condition compared to the baseline condition. We continued with pairwise comparisons using estimated marginal means with Tukey adjustment, finding that indeed there is a significant difference between all three means of ITL, also revealing a significant difference in ITLs between flicker-supported and flicker-unsupported runs in the close condition ($Estimate = -0.15$, $SE = 0.02$, $df = Inf$, $z.ratio = -6.36$, $p < 0.0001$). See specification in Table A.15.

A.2. Flicker-supported runs

A.2.1. Number of flickering targets per run

We fitted a GLMM model with fixed effect *chaining variant* and random effect *participant* on the number of flickering targets per run for only the flicker-supported runs to identify whether there were any differences in how many flickering targets observers collected in a run for each chaining variant. The model reports significant main effects of the *distant* condition ($Estimate = 0.38$, $SE = 0.03$, $t-value = 12.88$, $p < 2e-16$) and the *random* condition ($Estimate = 0.24$, $SE = 0.02$, $t-value = 12.23$, $p < 2e-16$), confirming that the highest number of flickering targets per run was in the close condition. See the specification in Table A.16.

Pairwise comparisons further confirm the significant differences between all three *chaining variants*. See Table A.17 for details.

Table A.16
GLMM for analysing effects of *chaining technique* on number of flickering targets per run.

	Estimate	SE	t-value	Pr(>| z |)	Exp. Coef.
(Intercept)	0.61767	0.01752	35.25	<2e-16 ***	1.854593
Distant Chaining	0.38090	0.02956	12.88	<2e-16 ***	1.463601
Random Chaining	0.23502	0.01922	12.23	<2e-16 ***	1.264935

Note. The model was specified as: $CountFlickeringTargets \sim ChainingTechnique + (1|ParticipantID)$. The intercept is set to *Close Chaining*. $AIC = -59.5981$, $BIC = -47.26856$. We compared the model against a random intercept model specified as $CountFlickeringTargets \sim 1 + (1|ParticipantID)$ $AIC_{\text{simple model}} = -49.588266$ $BIC_{\text{simple model}} = 56.98599$. ANOVA between the simple and our model: $\chi^2 = 113.19, df = 2, p < 2.2e-16$ ***. Family: Gamma (inverse). Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

Table A.17
Interaction contrasts for the pairwise comparisons of *chaining technique* on the mean count of flickering targets per run.

Chaining-Flicker pairwise	Estimate	SE	df	z-ratio	p-value
Close - Distant	−0.381	0.0296	Inf	−12.885	<0.0001
Close - Random	−0.235	0.0192	Inf	−12.230	<0.0001
Distant - Random	0.146	0.0315	Inf	4.634	<0.0001

Note. Using Tukey adjustments. Results are given on the log (not the response) scale.

A.2.2. Position of flickering target in run - count of runs

We fitted a GLMM with the independent variable *flicker position* (non-flicker, start, middle, end) as a fixed effect and *participant* as a random intercept with the *count of runs* per participant as the outcome variable. The model was specified with a gamma distribution and an inverse link function. It was fitted using maximum likelihood estimation with Laplace approximation. The intercept was set to the *non-flicker* category of *flicker position*. The model reported significant main effects for all levels of *flicker position*. See Table A.18. Following up with pairwise comparisons using estimated marginal means with Tukey adjustment (see Table A.19), we find significant differences between all categories. The means for the end and middle categories are higher compared to non-flicker runs in terms of geometric means but lower in terms of log-transformed means. The start category has a lower geometric mean but a higher log-transformed mean compared to non-flicker.

As we are interested in multiplicative effects and relative changes, we interpret the findings aligned with the geometric means derived from the model. These indicate that participants had more runs in the end and middle categories compared to the non-flicker category. Compared to the non-flicker category, participants had fewer runs starting with a flickering target.

A.2.3. Position of flickering target in run - run length

GLMM models for the effects of position in flickering run on run length (see Table A.20) and interaction contrasts (see Table A.21).

A.2.4. Position of flickering target in run - ITTs

GLMM models for the effects of position in flickering runs on ITTs (see Table A.22) and interaction contrasts (see Table A.23).

A.3. Collection rate and ITTs for flickering targets grouped by condition

A.3.1. Collection rate

We analyzed the effects with a GLMM model with *chaining variant* as a fixed effect and *participant* as a random intercept. The model was specified with a gamma distribution and a log link function to

Table A.18
GLMM for analysing effects of *flicker position* on run count.

	Estimate	SE	t-value	Pr(>| z |)	Exp. Coef.
(Intercept)	0.50230	0.02563	19.595	<2e-16 ***	1.6525157
FlickeringPositionCategoryEnd	0.31522	0.03374	9.343	<2e-16 ***	1.3705626
FlickeringPositionCategoryMiddle	0.29223	0.03301	8.852	<2e-16 ***	1.3394133
FlickeringPositionCategoryStart	−0.06236	0.02278	−2.738	0.00619	0.9395479

Note. The model was specified as: $CountRunsPerParticipant \sim FlickeringPositionCategory + (1|ParticipantID)$. The intercept is set to *Non-Flicker*. $AIC = 145.0769$, $BIC = 162.7268$. We compared the model against a random intercept model specified as $CountRunsPerParticipant \sim 1 + (1|ParticipantID)$ $AIC_{\text{simple model}} = 276.5325$ $BIC_{\text{simple model}} = 285.3575$. ANOVA between the simple and our model: $\chi^2 = 137.46, df = 3, p < 2.2e - 16$ ***. Family: Gamma (inverse). Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

Table A.19
Interaction contrasts for the pairwise comparisons of *flicker position* on run count.

Chaining-Flicker pairwise	Estimate	SE	df	z-ratio	p-value
(Non-Flicker) - End	−0.3152	0.0337	Inf	−9.343	<0.0001
(Non-Flicker) - Middle	−0.2922	0.0330	Inf	−8.852	<0.0001
(Non-Flicker) - Start	0.0624	0.0228	Inf	2.738	0.0062
End - Middle	0.146	0.0404	Inf	0.570	0.5690
End - Start - Start	0.3776	0.0326	Inf	11.575	<0.0001
Middle - Start - Start	0.3546	0.0319	Inf	11.127	<0.0001

Note. Using Tukey adjustments. Results are given on the log (not the response) scale.

account for the positively skewed data. The model was fitted using maximum likelihood estimation with Laplace approximation. See Table A.24 (Tables A.25–A.27).

A.3.2. ITT

A.4. Interaction effects of target type and chaining variant on collection rate

We analyzed the effects with a GLMM model with *chaining variant* and *target type* as fixed effects and *participant* as a random intercept. The model was specified with a gamma distribution and a log link function. The model was fitted using maximum likelihood estimation with Laplace approximation. The intercept was set to the *none (baseline)* chaining variant and the *repeated* target type. The model specifications are shown in Table A.28. As anticipated, the fitted GLMM model suggests that *switched target type* had a significant negative impact on the number of targets collected ($Estimate = -0.86$, $SE = 0.03$, $t - value = -28.43$, $p < 2e - 16$), indicating that on average participants collected 6.99 targets of *repeated* target type and 2.95 targets of *switched* target type.

Table A.20
GLMM for analysing effects of *flicker position* on run length.

	Estimate	SE	t-value	Pr(>| z |)	Exp. Coef.
(Intercept)	0.53199	0.04396	12.10	<2e-16 ***	1.702317
FlickeringPositionCategoryEnd	0.45614	0.03302	13.81	<2e-16 ***	1.577976
FlickeringPositionCategoryMiddle	0.96993	0.03302	29.38	< 2e-16 ***	2.637759
FlickeringPositionCategoryStart	0.34628	0.03304	10.48	< 2e-16 ***	1.413794

Note. The model was specified as: $RunLength \sim FlickeringPositionCategory + (1|ParticipantID)$. The intercept is set to *Non-Flicker*. $AIC = 172.9667$, $BIC = 190.6165$. We compared the model against a random intercept model specified as $RunLength \sim 1 + (1|ParticipantID)$ $AIC_{\text{simple model}} = 431.932$ $BIC_{\text{simple model}} = 440.757$. ANOVA between the simple and our model: $\chi^2 = 264.97, df = 3, p < 2.2e - 16$ ***. Family: Gamma (log). Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

A.4.1. Interaction effects of target type and flickering on collection rate

We fitted a GLMM model with the new independent variable *chaining-flicker* and *target type* as fixed effects and *participant* as a random intercept. *Collection rate* was the outcome variable. The model was specified with a gamma distribution and a log link function and it was fitted using maximum likelihood estimation with Laplace approximation. The intercept was set to the *none (baseline)* chaining variant and the *repeated* target type. The model specifications are shown in Table A.29.

The model reported significant main effects for both *close chaining - non-flickering targets* ($Estimate = -0.59$, $SE = 0.06$, $t - value = -9.05$, $p < 2e - 16$) and *close chaining - flickering targets* ($Estimate = -0.78$, $SE = 0.06$, $t - value = -12.7133$, $p < 2e - 16$), as well as for the *switched target type* ($Estimate = -0.89$, $SE = 0.06$, $t - value = -13.73$, $p < 2e - 16$). This confirmed that in both cases fewer targets were selected (which was expected, as the number of targets was split between the two). Furthermore, the model found a significant interaction effect of *close chaining - flickering targets* and *switched target type* ($Estimate = 0.42$, $SE = 0.09$, $t - value = 4.54$, $p = 5.64e - 06$), hinting at the ratio between the number of collected targets being significantly smaller when selecting flickering targets in the close condition than during the baseline. To find out if there are significant differences between the ratios of repeated/switched targets, we then explored the interaction effects further through pairwise comparisons using estimated marginal means with Bonferroni adjustment. The contrasts indicate significant differences between *baseline* and *close chaining - flickering targets* ($Estimate = 0.42$, $SE = 0.09$, $df = Inf$, $z.ratio = 4.54$, $p < 0.0001$) and between *close chaining - flickering targets* and *close chaining - non-flickering targets* ($Estimate = 0.25$, $SE = 0.09$, $df = Inf$, $z.ratio = 2.73$, $p = 0.006$). The details for the interaction contrast test are given in Table A.30.

Table A.21
Interaction contrasts for the pairwise comparisons of *flicker position* on run length.

Chaining-Flicker pairwise	Estimate	SE	df	z-ratio	p-value
(Non-Flicker) - End	−0.456	0.033	Inf	−13.814	<0.0001
(Non-Flicker) - Middle	−0.970	0.033	Inf	−29.376	<0.0001
(Non-Flicker) - Start	−0.346	0.033	Inf	−10.480	< 0.0001
End - Middle	−0.514	0.033	Inf	−15.563	<0.0001
End - Start - Start	0.110	0.033	Inf	3.326	0.0049
Middle - Start - Start	0.624	0.033	Inf	18.900	<0.0001

Note. Using Tukey adjustments. Results are given on the log (not the response) scale.

A.4.2. Interaction effects of target type and flickering on ITT

For ITTs, the model was specified with a gamma distribution and a log link function and it was fitted using maximum likelihood estimation with Laplace approximation. The intercept was set to the *none (baseline)* chaining variant and the *repeated* target type. The model specifications are shown in Table A.31.

The model reported significant main effects for both *close chaining - flickering targets* ($Estimate = -0.11$, $SE = 0.03$, $t - value = -3.64$, $p = 0.000277$) and for the *switched target type* ($Estimate = -0.14$, $SE = 0.03$, $t - value = 4.71$, $p = 2.53e - 06$). This indicates that participants were significantly faster when selecting flickering targets compared to the baseline condition. The results also confirm that selecting switched

Table A.22
GLMM for analysing effects of *flicker position* on ITTs.

	Estimate	SE	t-value	Pr(>| z |)	Exp. Coef.
(Intercept)	6.73416	0.05708	117.986	<2e-16 ***	840.6357950
FlickeringPositionCategoryEnd	−0.23150	0.03271	−7.077	1.48e-12 ***	0.7933401
FlickeringPositionCategoryMiddle	−0.24560	0.03276	−7.498	6.50e-14 ***	0.7822335
FlickeringPositionCategoryStart	−0.14605	0.03272	−4.463	8.08e-06 ***	0.8641148

Note. The model was specified as: $ITT \sim FlickeringPositionCategory + (1|ParticipantID)$. The intercept is set to *Non-Flicker*. $AIC = 1742.482$, $BIC = 1770.132$. We compared the model against a random intercept model specified as $ITT \sim 1 + (1|ParticipantID)$ $AIC_{\text{simple model}} = 1798.818$ $BIC_{\text{simple model}} = 1807.643$. ANOVA between the simple and our model: $\chi^2 = 52.337, df = 3, p = 2.539e - 11$ $***$. Family: Gamma (log). Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

Table A.23
Interaction contrasts for the pairwise comparisons of *flicker position* on ITT.

Chaining-Flicker pairwise	Estimate	SE	df	z-ratio	p-value
(Non-Flicker) - End	0.2315	0.0327	Inf	7.077	<0.0001
(Non-Flicker) - Middle	0.2456	0.0328	Inf	7.498	<0.0001
(Non-Flicker) - Start	0.1460	0.0327	Inf	4.463	< 0.0001
End - Middle	0.0141	0.0328	Inf	0.430	0.6675
End - Start - Start	−0.0855	0.0328	Inf	−2.608	0.0091
Middle - Start - Start	−0.0996	0.0327	Inf	−3.043	0.0023

Note. Using Tukey adjustments. Results are given on the log (not the response) scale.

Table A.24
GLMM for comparing collection rate for flickering targets grouped by condition.

	Estimate	SE	t-value	Pr(>| z |)	Exp. Coef.
(Intercept)	1.6163	0.0570	28.3	<2e-16 ***	5.034
Distant Chaining	−1.6615	0.0595	−27.9	<2e-16 ***	0.190
Random Chaining	−1.0746	0.0452	−23.8	<2e-16 ***	0.341

Note. The model was specified as: $CollectionRate \sim Condition + (1|ParticipantID)$. The intercept is set to *Close Chaining*. $AIC = 155.4154$, $BIC = 167.7450$, $AIC_{\text{simple model}} = 351.7121$ $BIC_{\text{simple model}} = 359.1098$, ANOVA between the simple and our model: $\chi^2 = 100, df = 2, p =< 2e - 16$ $***$. Family: Gamma (log). Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

Table A.25
Interaction contrasts for chaining technique on collection rate.

Chaining Variant	Estimate	SE	df	z-ratio	p-value
Close - Distant	1.661	0.0595	Inf	27.950	<0.0001
Close - Random	1.075	0.0452	Inf	23.780	<0.0001
Distant - Random	−0.587	0.0596	Inf	−9.850	<0.0001

Note. Using Tukey adjustments. Results are given on the log (not the response) scale.

Table A.26
GLMM for comparing collection rate for flickering targets grouped by condition.

	Estimate	SE	t-value	Pr(>| z |)	Exp. Coef.
(Intercept)	6.4889	0.0539	120.30	<2e-16 ***	657.80
Distant Chaining	0.4818	0.0512	9.40	<2e-16 ***	1.62
Random Chaining	0.0812	0.0383	2.12	0.034 *	1.08

Note. The model was specified as: $CollectionRate \sim Condition + (1|ParticipantID)$. The intercept is set to *Close Chaining*. $AIC = 1128.132$, $BIC = 1140.462$, $AIC_{\text{simple model}} = 1185.036$ $BIC_{\text{simple model}} = 1192.433$, ANOVA between the simple and our model: $\chi^2 = 60.9, df = 2, p = 6e - 14$ $***$. Family: Gamma (log). Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

Table A.27
Interaction contrasts for chaining technique on ITTs.

Chaining Variant	Estimate	SE	df	z-ratio	p-value
Close - Distant	−0.482	0.0512	Inf	−9.400	<0.0001
Close - Random	−0.081	0.0383	Inf	−2.120	0.0342
Distant - Random	0.401	0.0513	Inf	7.810	<0.0001

Note. Using Tukey adjustments. Results are given on the log (not the response) scale.

Table A.28
GLMM for analysing effects of chaining technique and target type on collection rate.

	Estimate	SE	t-value	Pr(>| z |)	Exp. Coef.
(Intercept)	1.94	0.058611	33.175	<2e-16 ***	6.9895265
Close Chaining	0.029538	0.042149	0.701	0.483	1.0299784
Distant Chaining	0.012014	0.042123	0.285	0.775	1.0120868
Random Chaining	0.002317	0.042130	0.055	0.956	1.0023201
Switched Target Type	−0.862828	0.030350	=28.429	<2e-16 ***	0.4219673

Note. The model was specified as: $CollectionRate \sim Condition + TargetType + (1|ParticipantID)$. The intercept is set to *None (baseline)* and the *switched* target type. $AIC = 926.36$, $BIC = 951.81$. We compared the model against a more complex model specified as $CollectionRate \sim Condition * TargetType + (1|ParticipantID)$ $AIC_{\text{complex model}} = 929.34$ $BIC_{\text{complex model}} = 965.69$. Family: Gamma (log). Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

targets resulted in longer ITTs. We also found an interaction effect of *close-non-flickering* and *switched target type* ($Estimate = 0.16$, $SE = 0.04$, $t - value = 3.77$, $p = 0.000165$), indicating that participants took significantly longer in selecting targets when they were not flickering in the *close chaining* condition compared to the baseline. We then explored the interaction effects further through pairwise comparisons using estimated marginal means with Tukey adjustment.

The contrasts indicate significant differences between *baseline* and *close chaining - non-flickering targets* ($Estimate = 0.16$, $SE = 0.04$, $df = Inf$, $z.ratio = 3.77$, $p = 0.0002$) and between *close chaining - flickering targets* and *close chaining - non-flickering targets* ($Estimate == 0.11$, $SE = 0.04$, $df = Inf$, $z.ratio = 2.48$, $p = 0.0133$). The details for the interaction contrast test are given in Table A.32.

A.4.3. Interaction effects of target type and flickering on ITL

For ITLs, the model was specified with a gamma distribution and a inverse link function and it was fitted using maximum likelihood estimation with Laplace approximation. The intercept was set to the

Table A.29
GLMM for analysing effects of the combination of chaining technique and flickering targets and target type on collection rate.

	Estimate	SE	t-value	Pr(>|z|)	Exp. Coef.
(Intercept)	1.96229	0.06018	32.605	<2e-16 ***	7.1155715
Close Chaining - Non-Flickering	−0.58535	0.06466	−9.053	<2e-16 ***	0.5569121
Close Chaining - Flickering	−0.78349	0.06462	−12.125	<2e-16 ***	0.4568082
Switched	−0.89148	0.06491	−13.734	<2e-16 ***	0.4100486
Close Chaining - Non-Flickering: Switched	0.16667	0.09139	1.824	0.0682	1.1813675
Close Chaining - Flickering: Switched	0.41716	0.09190	4.539	5.64e-06 ***	1.5176522

Note. The model was specified as: $CollectionRate \sim Chaining - Flicker * TargetType + (1|ParticipantID)$. The intercept is set to *None (baseline)* and the ***repeated*** target type. $AIC = 575.2358$, $BIC = 601.9744$. We compared the model against a simpler model specified as $CollectionRate \sim Chaining - Flicker + TargetType + (1|ParticipantID)$ $AIC_{\text{complex model}} = 591.0829$ $BIC_{\text{simple model}} = 611.1369$, which yielded a better fit for the complex model. ANOVA between the simple and our model: $\chi^2 = 19.9, df = 2, p = 4.9e-05$ ***. Family: Gamma (log). Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

Table A.30
Interaction contrasts for the combination of chaining technique and flickering targets and target type on collection rate.

Chaining-Flicker pairwise	Target type pairwise	Estimate	SE	df	z-ratio	p-value
Baseline - Close Non-Flickering	Repeated-Switched	0.167	0.0914	Inf	1.824	0.0682
Baseline - Close Flickering	Repeated-Switched	0.417	0.0919	Inf	4.539	<0.0001
Close Non-Flickering - Close Flickering	Repeated-Switched	0.250	0.0918	Inf	2.728	0.0064

Note. Using Tukey adjustments. Results are given on the log (not the response) scale.

Table A.31
GLMM for analysing effects of the combination of chaining technique and flickering targets and target type on ITTs.

	Estimate	SE	t-value	Pr(>| z |)	Exp. Coef.
(Intercept)	6.52841	0.05539	117.872	<2e-16 ***	684.3065796
Close Chaining - Non-Flickering	0.02315	0.03012	0.769	0.442161	1.0234216
Close Chaining - Flickering	−0.10954	0.03012	−3.636	0.000277	0.8962486
Switched	0.14180	0.03013	4.706	2.53e-06 ***	1.1523458
Close Chaining - Non-Flickering: Switched	0.16058	0.04262	3.768	0.000165 ***	1.1741890
Close Chaining - Flickering: Switched	0.05449	0.04279	1.274	0.202792	1.0560063

Note. The model was specified as: $ITT \sim Chaining - Flicker * TargetType + (1|ParticipantID)$. The intercept is set to *None (baseline)* and the *repeated* target type. $AIC = 2569.758$, $BIC = 2596.497$. We compared the model against a simpler model specified as $ITT \sim Chaining - Flicker + TargetType + (1|ParticipantID)$ $AIC_{\text{simple model}} = 2579.902$ $BIC_{\text{simple model}} = 2599.956$. ANOVA between the simple and our model: $\chi^2 = 14.144, df = 2, p = 0.0008486$. Family: Gamma (log). Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

Table A.32
Interaction contrasts for the combination of chaining technique and flickering targets and target type on ITTs.

Chaining-Flicker pairwise	Target type pairwise	Estimate	SE	df	z-ratio	p-value
Baseline - Close Non-Flickering	Repeated-Switched	0.1606	0.0426	Inf	3.768	0.0002
Baseline - Close Flickering	Repeated-Switched	0.0545	0.0428	Inf	1.274	<0.2028
Close Non-Flickering - Close Flickering	Repeated-Switched	−0.1061	0.0428	Inf	−2.476	0.0133

Note. Using Tukey adjustments. Results are given on the log (not the response) scale.

Table A.33
GLMM for analysing effects of the combination of chaining technique and flickering targets and target type on ITLs.

	Estimate	SE	t-value	Pr(>| z |)	Exp. Coef.
(Intercept)	1.0525	0.0167	62.86	<2e-16 ***	2.865
Close Chaining - Non-Flickering	−0.1930	0.0115	−16.73	<2e-16 ***	0.825
Close Chaining - Flickering	0.3452	0.0151	22.82	<2e-16 ***	1.412
Switched	−0.0620	0.0103	−6.05	1.5e-09 ***	0.940

Note. The model was specified as: $ITL \sim Chaining - Flicker * TargetType + (1|ParticipantID)$. The intercept is set to *None (baseline)* and the *repeated* target type. $AIC = -464.9856$, $BIC = -438.2469$. We compared the model against a simpler model specified as $ITL \sim Chaining - Flicker + TargetType + (1|ParticipantID)$ $AIC_{\text{simple model}} = -463.9167$ $BIC_{\text{simple model}} = -443.8627$. ANOVA between the simple and our model: $\chi^2 = 5.0688, df = 2, p = 0.07931$, showing that the simple model has a better fit. Family: Gamma (inverse). Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

Table A.34
GLMM for analysing effects of *chaining technique* on run length.

	Estimate	SE	t-value	Pr(>| z |)	Exp. Coef.
(Intercept)	0.3831513	0.0182757	20.965	<2e-16 ***	1.4668999
Close Chaining	0.0155549	0.0083890	1.854	0.0637	1.0156765
Distant Chaining	−0.0005212	0.0081643	−0.064	0.9491	0.9994790
Random Chaining	−0.0021621	0.0081416	−0.266	0.7906	0.9978402

Note. The model was specified as: $RunLength \sim Condition + (1|ParticipantID)$. The intercept is set to *None (baseline)*. $AIC = 127.1835$, $BIC = 144.8333$. We compared the model against a random intercept model specified as $RunLength \sim 1 + (1|ParticipantID)$ $AIC_{\text{simple model}} = 126.8683$ $BIC_{\text{simple model}} = 135.6932$. ANOVA between the simple and our model: $\chi^2 = 5.6848, df = 3, p = 0.128$. Family: Gamma (inverse). Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

none (baseline) chaining variant and the *repeated* target type. The model specifications are shown in Table A.33.

The model reported two significant main effects for *close chaining - non-flickering* ($Estimate = -0.17$, $SE = 0.02$, $t - value = -10.25$, $p < 2e-16$) and *close chaining - flickering* ($Estimate = 0.38$, $SE = 0.02$, $t - value = 17.44$, $p < 2e-16$) (Tables A.34–A.36).

Table A.35
GLMM for analysing effects of *chaining technique* on run number.

	Estimate	SE	t-value	Pr(>| z |)	Exp. Coef.
(Intercept)	1.510187	0.059940	25.195	<2e-16 ***	4.5275760
Close Chaining	0.040409	0.023586	1.713	0.0867	1.0412369
Distant Chaining	−0.003188	0.023554	−0.135	0.8923	0.9968171
Random Chaining	−0.020314	0.023558	−0.862	0.3885	0.9798912

Note. The model was specified as: $RunNumber \sim Condition + (1|ParticipantID)$. The intercept is set to *None (baseline)*. $AIC = 127.1835$, $BIC = 144.8333$. We compared the model against a random intercept model specified as $RunLength \sim 1 + (1|ParticipantID)$ $AIC_{\text{simple model}} = 126.8683$ $BIC_{\text{simple model}} = 135.6932$. ANOVA between the simple and our model: $\chi^2 = 5.6848, df = 3, p = 0.128$. Family: Gamma (log). Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

Table A.36
GLMM for analysing effects of *chaining technique* on run number.

	Estimate	SE	t-value	Pr(>| z |)	Exp. Coef.
(Intercept)	1.510187	0.059940	25.195	<2e-16 ***	4.5275760
Close Chaining	0.040409	0.023586	1.713	0.0867	1.0412369
Distant Chaining	−0.003188	0.023554	−0.135	0.8923	0.9968171
Random Chaining	−0.020314	0.023558	−0.862	0.3885	0.9798912

Note. The model was specified as: $RunNumber \sim Condition + (1|ParticipantID)$. The intercept is set to *None (baseline)*. $AIC = 127.1835$, $BIC = 144.8333$. We compared the model against a random intercept model specified as $RunLength \sim 1 + (1|ParticipantID)$ $AIC_{\text{simple model}} = 126.8683$ $BIC_{\text{simple model}} = 135.6932$. ANOVA between the simple and our model: $\chi^2 = 5.6848, df = 3, p = 0.128$. Family: Gamma (log). Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

Data availability

The research data, including eye tracking validation, flickering intensity calibration, study logs of the object states and selections, and the R analysis scripts are available at this link: https://osf.io/h8q2z/?view_only=f5327677e24846209a3ac29a52d3e44c.

References

Abrams, R.A., Christ, S.E., 2003. Motion onset captures attention. Psychol. Sci. 14, 427–432. https://doi.org/10.1111/1467-9280.01458. PMID: 12930472.

Bailey, R., McNamara, A., Sudarsanam, N., Grimm, C., 2009. Subtle gaze direction. ACM Trans. Graph. 28, https://doi.org/10.1145/1559755.1559757

Bailey, R., McNamara, A., Costello, A., Sridharan, S., Grimm, C., 2012. Impact of subtle gaze direction on short-term spatial information recall. In: Proceedings of the Symposium on Eye Tracking Research and Applications ETRA '12. Association for Computing Machinery, New York, NY, USA, pp. 67–74, https://doi.org/10.1145/2168556.2168567.

Bochynska, A., Laeng, B., 2015. Tracking down the path of memory: eye scanpaths facilitate retrieval of visuospatial information. Cogn. Process. 16, 159–163. https://doi.org/10.1007/s10339-015-0690-0

Booth, T., Sridharan, S., McNamara, A., Grimm, C., Bailey, R., 2013. Guiding attention in controlled real-world environments. In: Proceedings of the ACM Symposium on Applied Perception, SAP '13. Association for Computing Machinery, New York, NY, USA, pp. 75–82, https://doi.org/10.1145/2492494.2492508.

Brandt, S.A., Stark, L.W., 1997. Spontaneous eye movements during visual imagery reflect the content of the visual scene. J. Cogn. Neurosci. 9, 27–38. https://doi.org/10.1162/jocn.1997.9.1.27

Chiquet, S., Martarelli, C.S., Mast, F.W., 2021. Eye movements to absent objects during mental imagery and visual memory in immersive virtual reality. Virtual Real. 25, 655–667. https://doi.org/10.1007/s10055-020-00478-y

Cornsweet, T.N., 1962. The staircase-method in psychophysics. Am. J. Psychol. 75, 485–491. http://www.jstor.org/stable/1419876.

Egeth, H.E., Virzi, R.A., Garbart, H., 1984. Searching for conjunctively defined targets. J. Exp. Psychol. Hum. Percept. Perform. 10, 32–39. https://doi.org/10.1037/0096-1523.10.1.32, Place: US Publisher: American Psychological Association.

Engbert, R., Kliegl, R., 2003. Microsaccades uncover the orientation of covert attention. Vis. Res. 43, 1035–1045. https://doi.org/10.1016/S0042-6989(03)00084-1. https://www.sciencedirect.com/science/article/pii/S0042698903000841.

Feick, M., Kleer, N., Zenner, A., Tang, A., Krüger, A., 2021. Visuo-haptic illusions for linear translation and stretching using physical proxies in virtual reality. In: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, CHI '21. Association for Computing Machinery, New York, NY, USA, pp. 1–13, https://doi.org/10.1145/3411764.3445456.

Grant, E.R., Spivey, M.J., 2003. Eye movements and problem solving: guiding attention guides thought. Psychol. Sci. 14, 462–466. https://doi.org/10.1111/1467-9280.02454. PMID: 12930477.

Grogorick, S., Magnor, M., 2020. Subtle Visual Attention Guidance in VR. Springer International Publishing, Cham, pp. 272–284. https://doi.org/10.1007/978-3-030-41816-8_11

Grogorick, S., Stengel, M., Eisemann, E., Magnor, M., 2017. Subtle gaze guidance for immersive environments. In: Proceedings of the ACM Symposium on Applied Perception, SAP '17. Association for Computing Machinery, New York, NY, USA, pp. 1–7, https://doi.org/10.1145/3119881.3119890.

Grogorick, S., Albuquerque, G., Tauscher, J.-P., Magnor, M., 2018. Comparison of unobtrusive visual guidance methods in an immersive dome environment. ACM Trans. Appl. Percept. 15, https://doi.org/10.1145/3238303

Hu, J., Dudley, J.J., Kristensson, P.O., 2025. Seeing and touching the air: unraveling eye-hand coordination in mid-air gesture typing for mixed reality. In: Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems, CHI '25. Association for Computing Machinery, New York, NY, USA, https://doi.org/10.1145/3706598.3713743.

Ian M. Thornton, C.D., Kristjánsson, Á., 2019. The influence of selection modality, display dynamics and error feedback on patterns of human foraging. Vis. Cogn. 27, 626–648. https://doi.org/10.1080/13506285.2019.1658001

Jarodzka, H., van Gog, T., Dorr, M., Scheiter, K., Gerjets, P., 2013. Learning to see: guiding students' attention via a model's eye movements fosters learning. Learn. Instr. 25, 62–70. https://doi.org/10.1016/j.learninstruc.2012.11.004. https://www.sciencedirect.com/science/article/pii/S0959475212000990.

Jóhannesson, .I., Thornton, I.M., Smith, I.J., Chetverikov, A., Kristjánsson Á., 2016. Visual foraging with fingers and eye gaze. i-Percept. 7, 2041669516637279. https://doi.org/10.1177/2041669516637279. PMID: 27433323.

Johansson, R., Johansson, M., 2014. Look here, eye movements play a functional role in memory retrieval. Psychol. Sci. 25, 236–242. https://doi.org/10.1177/0956797613498260. PMID: 24166856.

Krekhov, A., Cmentowski, S., Waschk, A., Krüger, J., 2020. Deadeye visualization revisited: investigation of preattentiveness and applicability in virtual environments. IEEE Trans. Vis. Comput. Graph. 26, 547–557. https://doi.org/10.1109/TVCG.2019.2934370

Kristjánsson Á., Jóhannesson, .I., Thornton, I.M., 2014. Common attentional constraints in visual foraging. PLoS One 9, 1–9. https://doi.org/10.1371/journal.pone.0100752

Kristjánsson, T., Thornton, I.M., Kristjánsson Á., 2018. Time limits during visual foraging reveal flexible working memory templates. J. Exp. Psychol. Hum. Percept. Perform. 44, 827–835. https://doi.org/10.1037/xhp0000517, Place: US Publisher: American Psychological Association.

Kristjánsson Á., Ólafsdóttir, I.M., Kristjánsson, T., 2020. Visual foraging tasks provide new insights into the orienting of visual attention: methodological considerations. In: Pollmann, S. (Ed.), Spatial Learning and Attention Guidance. Springer US, New York, NY, pp. 3–21. https://doi.org/10.1007/7657_2019_21

Kristjánsson, Á., Ólafsdóttir, I.M., Kristjánsson, T., 2020. Visual Foraging Tasks Provide New Insights Into the Orienting of Visual Attention: Methodological Considerations. Springer US, New York, NY, pp. 3–21. https://doi.org/10.1007/7657_2019_21

Kristjánsson, T., Draschkow, D., Pálsson, Á., Haraldsson, D., Örn Jónsson, P., Kristjánsson, Á., 2022. Moving foraging into three dimensions: feature- versus conjunction-based foraging in virtual reality. Q. J. Exp. Psychol. 75, 313–327. https://doi.org/10.1177/1747021820937020. PMID: 32519926.

Kwok, T.C., Kiefer, P., Schinazi, V.R., Adams, B., Raubal, M., 2019. Gaze-guided narratives: adapting audio guide content to gaze in virtual and real environments. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, CHI '19. Association for Computing Machinery, New York, NY, USA, pp. 1–12, https://doi.org/10.1145/3290605.3300721.

Lange, D., Stratmann, T.C., Gruenefeld, U., Boll, S., 2020. HiveFive: immersion preserving attention guidance in virtual reality. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, CHI '20. Association for Computing Machinery, New York, NY, USA, pp. 1–13, https://doi.org/10.1145/3313831.3376803.

Lankes, M., Ramirez Gomez, A., 2022. GazeCues: exploring the effects of gaze-based visual cues in virtual reality exploration games. Proc. ACM Hum.-Comput. Interact. 6, https://doi.org/10.1145/3549500

Lu, W., Duh, B.-L.H., Feiner, S., 2012. Subtle cueing for visual search in augmented reality. In: 2012 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), pp. 161–166, https://doi.org/10.1109/ISMAR.2012.6402553

Magnien, L., Bouraoui, J.L., Vigouroux, N., 2004. Mobile text input with soft keyboards: optimization by means of visual clues. In: Brewster, S., Dunlop, M. (Eds.), Mobile Human-Computer Interaction - MobileHCI 2004. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 337–341.

Martarelli, C.S., Mast, F.W., 2013. Eye movements during long-term pictorial recall. Psychol. Res. 77, 303–309. https://doi.org/10.1007/s00426-012-0439-7

McNamara, A., Bailey, R., Grimm, C., 2008. Improving search task performance using subtle gaze direction. In: Proceedings of the 5th Symposium on Applied Perception in Graphics and Visualization, APGV '08. Association for Computing Machinery, New York, NY, USA, pp. 51–56, https://doi.org/10.1145/1394281.1394289.

McNamara, A., Booth, T., Sridharan, S., Caffey, S., Grimm, C., Bailey, R., 2012. Directing gaze in narrative art. In: Proceedings of the ACM Symposium on Applied Perception, SAP '12. Association for Computing Machinery, New York, NY, USA, pp. 63–70, https://doi.org/10.1145/2338676.2338689.

Mutasim, A.K., Batmaz, A.U., Hudhud Mughrabi, M., Stuerzlinger, W., 2025. The influence of eye gaze interaction technique expertise and the guided evaluation method on text entry performance evaluations. Proc. ACM Hum.-Comput. Interact. 9, https://doi.org/10.1145/3725842

Piening, R., Pfeuffer, K., Esteves, A., Mittermeier, T., Prange, S., Schröder, P., Alt, F., 2021. Looking for info: evaluation of gaze based information retrieval in augmented reality. In: Ardito, C., Lanzilotti, R., Malizia, A., Petrie, H., Piccinno, A., Desolda, G., Inkpen,

K. (Eds.), Human-Computer Interaction – INTERACT 2021. Springer International Publishing, Cham, pp. 544–565.

Rivu, R., Mäkelä, V., Hassib, M., Abdelrahman, Y., Alt, F., 2021. Exploring how saliency affects attention in virtual reality. In: Ardito, C., Lanzilotti, R., Malizia, A., Petrie, H., Piccinno, A., Desolda, G., Inkpen, K. (Eds.), Human-Computer Interaction – INTERACT 2021. Springer International Publishing, Cham, pp. 147–155.

Rothe, S., Althammer, F., Khamis, M., 2018. GazeRecall: using gaze direction to increase recall of details in cinematic virtual reality. In: Proceedings of the 17th International Conference on Mobile and Ubiquitous Multimedia, MUM '18. Association for Computing Machinery, New York, NY, USA, pp. 115–119, https://doi.org/10.1145/3282894.3282903.

Seeliger, A., Merz, G., Holz, C., Feuerriegel, S., 2021. Exploring the effect of visual cues on eye gaze during AR-guided picking and assembly tasks. In: 2021 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), pp. 159–164, https://doi.org/10.1109/ISMAR-Adjunct54149.2021.00041

Sidenmark, L., Prummer, F., Newn, J., Gellersen, H., 2023. Comparing gaze, head and controller selection of dynamically revealed targets in head-mounted displays. IEEE Trans. Vis. Comput. Graph. 29, 4740–4750. https://doi.org/10.1109/TVCG.2023.3320235

Spivey, M.J., Geng, J.J., 2001. Oculomotor mechanisms activated by imagery and memory: eye movements to absent objects. Psychol. Res. 65, 235–241. https://doi.org/10.1007/s004260100059

Sutton, J., Langlotz, T., Plopski, A., Zollmann, S., Itoh, Y., Regenbrecht, H., 2022. Look over there! Investigating saliency modulation for visual guidance with augmented reality glasses. In: Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology, UIST '22. Association for Computing Machinery, New York, NY, USA, pp. 1–15, https://doi.org/10.1145/3526113.3545633.

Sutton, J., Langlotz, T., Plopski, A., Hornbæk, K., 2024. Flicker augmentations: rapid brightness modulation for real-world visual guidance using augmented reality. In: Proceedings of the CHI Conference on Human Factors in Computing Systems, CHI '24. Association for Computing Machinery, New York, NY, USA, pp. 1–19, https://doi.org/10.1145/3613904.3642085.

Todd, R.M., Manaligod, M.G.M., 2018. Implicit guidance of attention: the priority state space framework. Cortex 102, 121–138. https://doi.org/10.1016/j.cortex.2017.08.001. https://www.sciencedirect.com/science/article/pii/S0010945217302484.

Tünnermann, J., Chelazzi, L., Schubö, A., 2021. How feature context alters attentional template switching. J. Exp. Psychol. Hum. Percept. Perform. 47, 1431–1444. https://doi.org/10.1037/xhp0000951, Place: US Publisher: American Psychological Association.

Tünnermann, J., Á.Kristjánsson , Petersen, A., Schubö, A., Scharlau, I., 2022. Advances in the application of a computational theory of visual attention (TVA): moving towards more naturalistic stimuli and game-like tasks. Open Psychol. 4, 27–46. https://doi.org/10.1515/psych-2022-0002

von der Malsburg, T., 2015. Saccades: detection of saccades and fixations in raw eyetracking data. https://github.com/tmalsburg/saccades.

Wagner, U., Albrecht, M., Jacobsen, A.A., Wang, H., Gellersen, H., Pfeuffer, K., 2024. Gaze, wall, and racket: combining gaze and hand-controlled plane for 3D selection in virtual reality. Proc. ACM Hum.-Comput. Interact. 8, https://doi.org/10.1145/3698134

Waldner, M., Le Muzic, M., Bernhard, M., Purgathofer, W., Viola, I., 2014. Attractive flicker — guiding attention in dynamic narrative visualizations. IEEE Trans. Vis. Comput. Graph. 20, 2456–2465. https://doi.org/10.1109/TVCG.2014.2346352

Wolfe, J.M., 2013. When is it time to move to the next raspberry bush? Foraging rules in human visual search. J. Vis. 13, 10. https://doi.org/10.1167/13.3.10

Wolfe, J.M., 2014. Approaches to visual search: feature integration theory and guided search. Oxf. handb. atten. 11, 35–44.

Wolfe, J.M., 2020. Visual search: how do we find what we are looking for? Annu. Rev. Vis. Sci. 6, 539–562. https://doi.org/10.1146/annurev-vision-091718-015048

Wolfe, J.M., Horowitz, T.S., 2017. Five factors that guide attention in visual search. Nat. Hum. Behav. 1, 0058. https://doi.org/10.1038/s41562-017-0058

Wolfe, J.M., Aizenman, A.M., Boettcher, S.E., Cain, M.S., 2016. Hybrid foraging search: searching for multiple instances of multiple types of target. Vis. Res. 119, 50–59. https://doi.org/10.1016/j.visres.2015.12.006. https://www.sciencedirect.com/science/article/pii/S0042698915003739.

Wolfe, J.M., Cain, M.S., Alaoui-Soce, A., 2018. Hybrid value foraging: how the value of targets shapes human foraging behavior. Atten. Percept. Psychophys. 80, 609–621. https://doi.org/10.3758/s13414-017-1471-x

Wolfe, J.M., Cain, M.S., Aizenman, A.M., 2019. Guidance and selection history in hybrid foraging visual search. Atten. Percept. Psychophys. 81, 637–653. https://doi.org/10.3758/s13414-018-01649-5

Woodworth, J.W., Yoshimura, A., Lipari, N.G., Borst, C.W., 2023. Design and evaluation of visual cues for restoring and guiding visual attention in eye-tracked VR. In: 2023 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW), pp. 442–450, https://doi.org/10.1109/VRW58643.2023.00096

Zenner, A., Karr, C., Feick, M., Ariza, O., Krüger, A., 2024. Beyond the blink: investigating combined saccadic & blink-suppressed hand redirection in virtual reality. In: Proceedings of the CHI Conference on Human Factors in Computing Systems CHI '24. Association for Computing Machinery, New York, NY, USA, pp. 1–14, https://doi.org/10.1145/3613904.3642073