Disentangling the mechanisms underlying infant fixation durations in scene perception: A computational account

CrossMark

Irati R. Saez de Urabain [a], Antje Nuthmann [b], Mark H. Johnson [a], Tim J. Smith [a,*]

[a] Centre for Brain and Cognitive Development, Birkbeck, University of London, UK
[b] School of Philosophy, Psychology and Language Sciences, Psychology Department, University of Edinburgh, UK

ABSTRACT

The goal of this article is to investigate the unexplored mechanisms underlying the development of saccadic control in infancy by determining the generalizability and potential limitations of extending the CRISP theoretical framework and computational model of fixation durations (FDs) in adult scene-viewing to infants. The CRISP model was used to investigate the underlying mechanisms modulating FDs in 6-month-olds by applying the model to empirical eye-movement data gathered from groups of infants and adults during free-viewing of naturalistic and semi-naturalistic videos. Participants also performed a gap-overlap task to measure their disengagement abilities. Results confirmed the CRISP model's applicability to infant data. Specifically, model simulations support the view that infant saccade programming is completed in two stages: an initial labile stage, followed by a non-labile stage. Moreover, results from the empirical data and simulation studies highlighted the influence of the material viewed on the FD distributions in infants and adults, as well as the impact that the developmental state of the oculomotor system can have on saccade programming and execution at 6 months. The present work suggests that infant FDs reflect on-line perceptual and cognitive activity in a similar way to adults, but that the individual developmental state of the oculomotor system affects this relationship at 6 months. Furthermore, computational modeling filled the gaps of psychophysical studies and allowed the effects of these two factors on FDs to be simulated in infant data providing greater insights into the development of oculomotor and attentional control than can be gained from behavioral results alone.

© 2017 Elsevier Ltd. All rights reserved.

1. Introduction

From the moment we are born most daily activities involve constant decisions about where and when to move our eyes next. During active visual sampling our eyes may remain stable at a point (fixations, during which visual encoding occurs) or perform fast ballistic movements (saccadic eye movements; Matin, 1974; Ross, Morrone, Goldberg, & Burr, 2001).

In infants, FDs have been associated with the developmental state of the oculomotor system (Bronson, 1994; Johnson, 1990) and with visual and cognitive processes such as attention, information processing, memory or anticipation (e.g., Harris, Hainline, Abramov, Lemerise, & Camenzuli, 1988; Papageorgiou et al., 2014). Nevertheless, due to practical and technical limitations in testing young infants not much is known about the development of the mechanisms underlying FDs and saccade programming.

Computational modeling allows us to describe, predict and explain data that is itself unobservable (Lewandowsky & Farrell, 2011). This method is particularly useful in developmental science, where it permits the investigation of aspects of eye-movement control that could not be analyzed otherwise (Reichle et al., 2013). The present article aims to utilize a theoretical framework and computational model of FDs in scenes and determine its generalizability to infants (CRISP, a timer (C)ontrolled (R)andom-walk with (I)nhibition for (S)accade (P)lanning model; Nuthmann & Henderson, 2012; Nuthmann, Smith, Engbert, & Henderson, 2010), in order to investigate the mechanisms underlying FDs in 6-month-olds. Notably, the model assumes that saccades are programmed in two stages, an initial labile phase during which saccade programs can be altered or cancelled and a subsequent non-labile phase in which programs cannot be cancelled. Whether infant saccadic programming operates via these two phases is not known. In addition, the CRISP model will be used to examine whether FDs at this age are affected by developmental aspects of the oculomotor system and/or by visual and cognitive processing. For this purpose, we report fixation-

duration data from 6-month-old infants and adults who each viewed dynamic scenes, and two simulation studies that will test whether the data from both infants and adults can be explained by a single model architecture, with age-specific and task-specific influences realized by differences in parameter settings. In the following sections, we review the past research on FDs in adults and infants, and introduce the background literature on modeling FDs with CRISP.

1.1. FDs and saccade latency in adults

The relationship between FDs and visual and cognitive processing has been extensively investigated in skilled adult reading (Rayner, 1998, for review) and, more recently, also in the context of scene viewing (Nuthmann, 2016, for review). For instance, factors such as the viewing task (search vs. memorization; Castelhano, Mack, & Henderson, 2009; Nuthmann et al., 2010), the visual characteristics of the stimulus (e.g., luminance, image degradation; Loftus, 1985; Walshe & Nuthmann, 2014), the semantics of the scene (e.g., Henderson, Weeks, & Hollingworth, 1999; Loftus & Mackworth, 1978; Wu, Wick, & Pomplun, 2014), or familiarity (e.g., Althoff & Cohen, 1999) can affect gaze control and FDs. Collectively, these findings demonstrate that visual and cognitive processing demands are associated with differences in FDs.

The mechanisms underlying saccadic control in adults have been greatly investigated by studying saccadic responses in simple saccade-targeting tasks. In the double-step paradigm, for instance, participants are instructed to follow a target while it makes two successive movements or steps that are separated by a varying temporal gap (Becker & Jürgens, 1979; Camalier et al., 2007; Findlay & Harris, 1984; Westheimer, 1954). Findings from double-step studies have provided evidence for parallel programming of saccades in which saccade programming occurs in two stages: an initial labile stage which is subject to cancellation, followed by a non-labile stage that cannot be cancelled. Becker and Jürgens (1979) showed that participants' performance is best predicted by the time D elapsing between the onset of the second target step and the onset of the first saccade. If D is short (<70 ms), the response saccade is directed to the first target location. The saccade program to this location was already fully specified; in other words, it was in its non-labile stage of development when it could no longer be altered. As D increases, an amplitude transition function emerges, with the first saccade landing progressively closer to the second, final location of the target. When the first saccade lands on the second target, the oculomotor system began programming a saccade to the second target location while the saccade program for the first target location was still in its labile stage of development. In this situation, the first program is cancelled and only the second program is executed, prolonging the saccade latency and hence the duration of the fixation. Recently, Walshe and Nuthmann (2015) adopted the double-step paradigm to a scene-viewing context and showed that saccade cancellation processes generalize to scene viewing, and that cancelling a saccade prolongs FDs. The general finding that saccades are programmed in two stages has been adopted in computational models of fixation behavior in reading (Engbert, Longtin, & Kliegl, 2002; Reichle, Pollatsek, Fisher, & Rayner, 1998) and scene viewing (Nuthmann et al., 2010).

In infant research, video-based double-step paradigms have been used to investigate developmental changes in spatial remapping of saccade trajectories across the two saccades, but the paradigm has yet to be used to isolate the timing of the transition from labile to non-labile stage of saccade programming in infants (Brown et al., 2003; Gilmore & Johnson, 1997a, 1997b).

1.2. FDs and visual and cognitive processing in infancy

The mechanisms that control FDs in infants remain poorly understood. Evidence suggests that as early as 3- to 4-month-old infants' looking behavior (e.g., looking times to a particular stimulus) can be influenced by cognitive factors related to the visual input such as expectations of spatiotemporal object continuity and causality (e.g., Leslie & Keeble, 1987; Spelke, 1990). However, such studies examine rather coarse shifts in attention (look at the scene vs. do not look) with relatively few studies paying attention to the micro-dynamics of visual and cognitive processing (such as FDs).

For instance, some studies have investigated infants' scanning abilities and FDs when presented with familiar and non-familiar complex dynamic stimuli. Hunnius and Geuze (2004) followed infants between the ages of 6 and 26 weeks and presented them with a video of their mother's face, and an abstract video. They found that infants only adapted their eye-movements according to the type of stimulus from 14 weeks on, showing longer mean FDs for the abstract unfamiliar condition. Additionally, the median fixation duration did not stabilize before 18 weeks, which is slightly later than what has been reported for static stimuli (Bronson, 1990). These findings suggest that FDs in infancy can also reflect the visual and cognitive processing of the visual input, even though it is still unclear whether these factors have the same influence in infants and adults.

1.3. Neural mechanisms underlying eye-movement control in infancy

Whilst the subcortical structures involved in saccadic generation are relatively developed at birth (e.g., superior colliculus), cortical pathways associated with the generation of more complex eye-movements (e.g., the frontal eye fields) remain underdeveloped until 3 to 4 months of age (e.g., Atkinson, 2000; Bronson, 1974; Johnson, 1990, 2011). At around 1 month postnatal age, unregulated tonic inhibition of the superior colliculus prevents infants from consistently moving their eyes from a point of foveation. This phenomenon is commonly known as "sticky fixation" or "obligatory attention" (Atkinson, 2000; Braddick et al., 1992; Farroni, Simion, Umiltà, & Barba, 1999; Frick, Colombo, & Saxon, 1999; Johnson, 2011) and is thought to diminish from 3 to 4 months with the increasing cortical control over saccades. Sticky fixation is thought to occur due to problems with "disengagement", defined as the difficulty in generating an eye-movement after a fixation (Johnson, 1990).

The ability to disengage from a central target to shift the gaze to a peripheral target has traditionally been evaluated using the gap-overlap paradigm (Atkinson, Hood, Wattam-Bell, & Braddick, 1992; Farroni et al., 1999; Hood & Atkinson, 1990, 1993; Johnson, Posner, & Rothbart, 1991). Reaction times are usually faster on gap trials, where the central target disappears and after a temporal gap (e.g., 200 ms) the peripheral target appears. During overlap trials, the central target stays on after the peripheral target appears. Sometimes, baseline trials are additionally included, in which the peripheral target appears without a temporal gap after the central target disappears (e.g., Elsabbagh et al., 2009; Wass, Porayska-Pomsta, & Johnson, 2011). In this case, disengagement latencies can be calculated by subtracting the baseline latencies from the overlap latencies.

In infancy research, longer disengagement latencies have been associated with greater immaturity of the visual system (e.g., Butcher, Kalverboer, & Geuze, 2000; Matsuzawa & Shimojo, 1997), particularly during the first 6 months of life, when the neurological structures involved are thought to develop rapidly and approach their adult form (Rothbart, Posner, & Rosicky, 1994). Nevertheless, evidence from various neurophysiological and behavioral studies suggests that the neural mechanisms involved