Journal of Experimental Child Psychology 166 (2018) 293–309

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

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

Semantic content outweighs low-level saliency in determining children's and adults' fixation of movies

Andrew T. Rider [a], Antoine Coutrot [b], Elizabeth Pellicano [c,d], Steven C. Dakin [a,e], Isabelle Mareschal [f,*]

[a] UCL Institute of Ophthalmology, University College London, London WC1E 6BT, UK
[b] Centre for Mathematics and Physics in Life Sciences and Experimental Biology (CoMPLEX), University College London, London WC1E 6BT, UK
[c] Centre for Research in Autism and Education (CRAE), Department of Psychology and Human Development, UCL Institute of Education, University College London, London WC1H 0AL, UK
[d] School of Psychology, University of Western Australia, Crawley, Perth, Western Australia 6009, Australia
[e] School of Optometry and Vision Science, University of Auckland, Auckland 1010, New Zealand
[f] Department of Psychology, Queen Mary University, London E1 4NS, UK

ARTICLE INFO

Article history:
Received 11 August 2016
Revised 21 August 2017

Keywords:
Eye movements
Visual attention
Development
Gaze
Dynamic
Faces
Saliency

ABSTRACT

To make sense of the visual world, we need to move our eyes to focus regions of interest on the high-resolution fovea. Eye movements, therefore, give us a way to infer mechanisms of visual processing and attention allocation. Here, we examined age-related differences in visual processing by recording eye movements from 37 children (aged 6–14 years) and 10 adults while viewing three 5-min dynamic video clips taken from child-friendly movies. The data were analyzed in two complementary ways: (a) gaze based and (b) content based. First, similarity of scanpaths within and across age groups was examined using three different measures of variance (dispersion, clusters, and distance from center). Second, content-based models of fixation were compared to determine which of these provided the best account of our dynamic data. We found that the variance in eye movements decreased as a function of age, suggesting common attentional orienting. Comparison of the different models revealed that a model that relies on faces generally performed better than the other models tested, even for the youngest age group (<10 years). However, the best predictor of a given participant's eye movements was the

* Corresponding author.
E-mail addresses: a.rider@ucl.ac.uk (A.T. Rider), i.mareschal@qmul.ac.uk (I. Mareschal).

https://doi.org/10.1016/j.jecp.2017.09.002
0022-0965/

average of all other participants' eye movements both within the same age group and in different age groups. These findings have implications for understanding how children attend to visual information and highlight similarities in viewing strategies across development.

Introduction

Eye tracking is increasingly being used to try to infer what people are doing (Hayhoe & Ballard, 2005) or thinking (Kardan, Berman, Yourganov, Schmidt, & Henderson, 2015) based solely on how they looked at a visual scene. One advantage of this method over, for example, self-report or psychophysical testing is that it can readily be applied to populations that are difficult to evaluate, from young babies (Jones, Kalwarowsky, Atkinson, Braddick, & Nardini, 2014) to clinical populations, including autistic people (e.g., see Papagiannopoulou, Chitty, Hermens, Hickie, & Lagopoulos, 2014, for a review) and patients with Alzheimer's disease (Crutcher et al., 2009). Although there has been a great deal of work looking at modeling patterns of fixations in adults, particularly looking at top-down and bottom-up influences (Itti, Koch, & Niebur, 1998; Xu, Jiang, Wang, Kankanhalli, & Zhao, 2014), and there has been some important work examining how these influences develop during a child's first months and years (e.g., Amso, Haas, & Markant, 2014; Franchak, Heeger, Hasson, & Adolph, 2016; Frank, Vul, & Johnson, 2009), there has as yet been no systematic examination of different models applied to viewing dynamic scenes in school-age children compared with adults. We addressed this gap in the literature in this study.

Development of fixation behavior

Certain viewing behaviors, such as looking at faces and stimuli with social relevance, develop so early in childhood that they appear to be largely innate. For example, it is well established that newborn babies preferentially track faces and face-like stimuli in simple displays (Farroni et al., 2005; Johnson, Dziurawiec, Ellis, & Morton, 1991). Using static (complex) images, several studies have shown that infants aged 6 months and older orient to faces in images that contain non-face distractors (Di Giorgio, Turati, Altoè, & Simion, 2012; Gliga, Elsabbagh, Andravizou, & Johnson, 2009; Gluckman & Johnson, 2013). Frank, Vul, and Saxe (2012) used videos of objects, faces, children playing with toys, and complex social scenes with young children aged 3–30 months. They showed further that facial and bodily features that have social relevance, such as eyes, mouths, and hands, capture infants' and toddlers' attention and that this capacity to direct their attention to the stimuli that are potentially the most socially informative increases with age. In another study, Frank, Amso, and Johnson (2014) reported an age-related increase in looking at faces in complex videos (clips from *Peanuts* [Charlie Brown] and *Sesame Street*) in 3- to 9-month-old infants, which correlated with increased attentional orienting using a visual search task, suggesting that the extent to which infants show social preferences may well be underpinned by their ability to detect socially relevant stimuli in otherwise complex dynamic scenes. Frank et al. (2014) finding is also consistent with a recent report showing that infants over 4 months of age tended to look first and longest at faces, whereas 4-month-olds tended to look at the most salient object in a display (Kwon, Setoodehnia, Baek, Luck, & Oakes, 2016).

Less is known, however, about the developments in fixation behavior that take place beyond early childhood. Kirkorian, Anderson, and Keen (2012) showed children (aged 1–4 years) and adults 20-min clips of television shows and found that younger children fixated more regions over a larger area than did older children. This variability was greatest immediately following scene cuts, which these authors proposed is due to an inability to suppress attention to irrelevant features (from the previous scene). More recently, Helo, Pannasch, Sirri, and Rämä (2014) examined differences in scanning behavior in adults and children aged 2–10 years and reported that fixation durations decreased and saccade