ARTICLE IN PRESS

Vision Research xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Vision Research

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

Short and long-term visual deprivation leads to adapted use of audiovisual information for face-voice recognition

Stefania S. Moro[a,b], Adria E.N. Hoover[a], Jennifer K.E. Steeves[a,b,*]

[a] Department of Psychology and Centre for Vision Research, York University, Toronto, Canada
[b] The Hospital for Sick Children, Toronto, Canada

ARTICLE INFO

Number of Reviews = 3

Keywords:
Monocular enucleation
Visual deprivation
Face
Voice
Object recognition
Audiovisual

ABSTRACT

Person identification is essential for everyday social interactions. We quickly identify people from cues such as a person's face or the sound of their voice. A change in sensory input, such as losing one's vision, can alter how one uses sensory information. We asked how people with only one eye, who have had reduced visual input during postnatal maturation of the visual system, use faces and voices for person identity recognition. We used an old/new paradigm to investigate unimodal (visual or auditory) and bimodal (audiovisual) identity recognition of people (face, voice and face-voice) and a control category, objects (car, horn and car-horn). Participants learned the identity of 10 pairs of faces and voices (Experiment 1) and 10 cars and horns (Experiment 2) and were asked to identify the learned face/voice or car/horn among 20 distractors. People with one eye were more sensitive to voice identification compared to controls viewing binocularly or with an eye-patch. However, both people with one eye and eye-patched viewing controls use combined audiovisual information for person identification more equally than binocular viewing controls, who favour vision. People with one eye were no different from controls at object identification. The observed visual dominance for binocular controls is larger for person compared to object identification, indicating that faces (vision) play a larger role in person identification and that person identity processing is unique from that for objects. People with long-term visual deprivation from the loss of one eye may have adaptive strategies, such as placing less reliance on vision to achieve intact performance, particularly for face processing.

1. Introduction

We typically experience a combination of different sensory stimuli at the same time each day. Social interaction is often an important part of one's day ranging from chatting with friends to identifying whether an approaching individual is a friend or foe. Multisensory cues, such as seeing a person's face and hearing their voice provide important information that contribute to distinguishing an individual person's identity. If all of our sensory systems are intact, we use them to our full advantage.

Humans with complete visual deprivation have shown evidence for changes in other sensory systems that indicate enhanced abilities with their remaining senses following a complete loss of a sensory system. For instance, congenitally blind individuals have shorter response times for auditory discrimination tasks (Röder et al., 1999), faster processing of language (Röder, Stock, Bien, Neville, Rösler, & 2002), enhanced sound localization (Lessard, Paré, Lepore, & Lassonde, 1998) and enhanced tactile perception (Sathian, 2000) compared to sighted individuals. This suggests underlying physiological changes within the systems responsible for these senses to support these behavioural enhancements. If visual input were reduced by half, as in people with one functioning eye, it seems reasonable to expect that the other intact sensory systems should function to the best of their ability in order to adapt and compensate for the partial loss of vision.

Monocular blindness, resulting from the surgical removal of one eye (enucleation) represents a unique human model for examining the consequences of the loss of binocularity. It is unlike other more common forms of monocular deprivation such as amblyopia or strabismus since the removal of the end organ eliminates all forms of visual input to the brain from that eye (Kelly, Moro, & Steeves, 2012). People with one eye have enhanced sound processing ability compared to controls (Hoover, Harris, & Steeves, 2012). When presented with auditory stimuli along the horizontal azimuth, people with one eye have consistently more accurate spatial localization within a field of +/- 78 degrees compared to control participants who were monocular viewing, binocular viewing or with their eyes closed (Hoover et al., 2012).

Typically, when we are simultaneously presented with auditory and visual stimuli, the visual information is processed preferentially over

* Corresponding author at: 1032 Sherman Health Science Research Centre, York University, 4700 Keele St., Toronto, ON, Canada.
E-mail address: steeves@yorku.ca (J.K.E. Steeves).

https://doi.org/10.1016/j.visres.2018.01.009
Received 27 July 2017; Received in revised form 11 January 2018; Accepted 14 January 2018
0042-6989/

auditory information. This is also known as the Colavita visual dominance effect (Colavita, 1974). People with one eye do not demonstrate the Colavita effect when asked to categorize rapidly presented audiovisual targets consisting of line drawings of common objects paired with common sounds (Moro & Steeves, 2012). Instead, people with one eye process auditory and visual components of audiovisual targets equally. Together with the enhanced auditory spatial localization (Hoover et al., 2012), these results suggest that auditory information may be processed more reliably for people with one eye and could be a form of sensory compensation for the loss of binocularity.

A common illusion occurs when the auditory and visual components of an audiovisual stimulus are spatially displaced relative to one another, resulting in the perception of a single event typically displaced towards the visual component, known as the ventriloquism effect (Welch & Warren, 1980). People with one eye show similar audiovisual localization for spatially disparate audiovisual stimuli compared to controls viewing binocularly or with one eye-patched (Moro, Harris, & Steeves, 2014). However, unlike binocular and eye-patched controls, people with one eye take longer to localize unimodal visual stimuli compared to unimodal auditory stimuli (Moro et al., 2014).

For the most part, spatial visual ability is intact or somewhat better than controls in people with the loss of one eye early in life despite a 50% reduction of visual input to the visual brain. For example, the ability to discriminate low-contrast global shape, a more complex visual process (hyperacuity), is enhanced in people with one eye compared to eye-patched controls and controls viewing dichoptically (Steeves, Wilkinson, González, Wilson, & Steinbach, 2004). Face processing, however, has emerged as an exception. People with one eye do not show the composite face effect and they take longer to process the shape and spacing between internal facial features (Kelly, Gallie, & Steeves, 2012). This mild impairment is face-specific and was not replicated when performing the same tasks on stimuli from other visual image categories, namely houses (Kelly et al., 2012).

Face perception is a unique and multi-faceted aspect of vision (Leopold & Rhodes, 2010). The ability to identify a person is facilitated when face information is integrated with voice information through crosstalk between the unimodal visual and unimodal auditory percepts (Campanella & Belin, 2008). Previous exposure to combined face-voice information during person identity encoding facilitates identification of that individual when only unimodal cues (face or voice) are available (Ellis, Jones, & Mosdell, 1997; Schweinberger, Herholz, & Sommer, 1997; Sheffert & Olson, 2004; von Kriegstein et al., 2008). Bimodal identity recognition shows interference effects (increased reaction time and decreased accuracy) due to the addition of auditory information to visual information when identifying bimodal (face-voice) stimuli (Joassin, Maurage, Bruyer, Crommelinck, & Campanella, 2004). This suggests that because unimodal face recognition is superior to voice recognition, the addition of voice information interferes with the efficient processing of the face (during bimodal face-voice pairings). When face stimuli are degraded and therefore less reliable relative to the voices, bimodal stimulus presentations led to an enhancement effect indicating that the more reliable sensory information (face or voice) has greater influence on person identity recognition (Joassin, Maurage, & Campanella, 2008).

Face-voice integration effects, similar to those found in healthy controls (Campanella & Belin, 2008), are seen in infants as young as 4 months of age (Bahrick, Netto, & Hernandez-Reif, 1998) and in non-human primates (Izumi & Kojima, 2004). This suggests that face-voice identification improves with development and experience as older infants (7 months of age) have better face-voice matching compared to younger infants (4 months of age) (Bahrick et al., 1998). People who have had one eye removed early in life experience monocular deprivation during postnatal visual system maturation. The long-term consequences of this abnormal visual experience during development may result in altered use of auditory and visual sensory information when tested later in life, as mature adults.

The current study investigates how people with one eye use auditory and visual information for person and object identity recognition compared to binocular and eye-patched viewing controls. Since people with one eye have half of the visual input to the brain, will this alter face-voice integration? Will audiovisual integration be affected more generally and also alter the identification of audiovisual objects (cars paired with horns)? Using the same paradigm as Hoover, Démonet, and Steeves (2010) we measure visual, auditory and audiovisual recognition of people and objects in people with one eye compared to eye-patched and binocular viewing controls. The addition of an eye-patched control group compares whether the effects observed in people with one eye are simply due to monocular viewing or whether they are the result of long-term visual deprivation from eye enucleation. In Experiment 1, we quantify person recognition for faces, voices and face-voice pair combinations. We predict that people with one eye will have poorer person identification compared to controls as a result of reduced sensitivity to faces since previously we have shown mild face perception impairments in this monocular group (Kelly et al., 2012). In Experiment 2, we quantify object recognition for cars, car horns and car-car horn pair combinations. We predict no difference in object identification between groups. We have previously shown that higher-level aspects of spatial form vision are intact (Kelly et al., 2012) or somewhat enhanced (Steeves et al., 2004) for this group compared to eye-patched viewing. We also compare across experiments to assess whether person recognition is unique compared to object recognition.

2. Experiment 1: Person identity recognition

2.1. Materials and methods

2.1.1. Participants

2.1.1.1. People with one eye (monocular enucleation, ME). Eleven adult participants who had undergone early monocular eye enucleation (ME) at The Hospital For Sick Children (Toronto) participated in this study (mean age = 34 years, SD = 12). All ME participants had been unilaterally eye enucleated (7 right eye removed) due to retinoblastoma, a rare childhood cancer of the retina. Age at enucleation ranged from 4 to 66 months of age (mean age at enucleation = 22 months, SD = 16).

2.1.1.2. Binocular viewing control participants (BV). Twenty-five binocularly intact controls with a mean age of 27 years (SD = 7) were tested while viewing stimuli binocularly.

2.1.1.3. Patched viewing control participants (MV). Twenty-five binocularly intact participants, with a mean age of 24 years (SD = 3), completed the experiments with one eye patched. Participants' non-preferred eye (determined using the Porta test) was patched with a semi-opaque eye covering and translucent tape (12 right-eye covered).

All participants (ME, BV, MV) reported normal hearing and normal or corrected-to-normal acuity as assessed by an EDTRS eye chart (Precision Vision™, La Salle, IL) and wore optical correction if needed. All participants gave informed consent prior to their inclusion in the study, which was approved by York University Office of Research Ethics.

2.1.2. Stimuli

Stimuli were previously used in Hoover et al. (2010) for assessing face and object identities in a patient with visual agnosia. See Hoover et al. (2010) for more detailed information on how stimuli were created. In short, visual stimuli consisted of 110 greyscale female face images that were cropped within an oval aperture and distinguishable features such as beauty marks removed. Auditory stimuli were 20 s in duration and were played through headphones. Each consisted of a short neutral passage spoken in English by one of 110 female voices.

Participants sat 45 cm from the display in a dimly lit room with the