Vision Research xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Vision Research

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

A face in a (temporal) crowd

Catrina M. Hacker[a], Emily X. Meschke[a], Irving Biederman[a,b,*]

[a] Program in Neuroscience, University of Southern California, United States
[b] Department of Psychology, University of Southern California, United States

ARTICLE INFO

Number of Reviews = 1

Keywords:
Face recognition
RSVP
Prosopagnosia
Negative detection
Face familiarity
Celebrity recognition
Unconscious recognition

ABSTRACT

Familiar objects, specified by name, can be identified with high accuracy when embedded in a rapidly presented sequence of images at rates exceeding 10 images/s. Not only can target objects be detected at such brief presentation rates, they can also be detected under high uncertainty, where their classification is defined negatively, e.g., "Not a Tool." The identification of a familiar speaker's voice declines precipitously when uncertainty is increased from one to a mere handful of possible speakers. Is the limitation imposed by uncertainty, i.e., the number of possible individuals, a general characteristic of processes for person individuation such that the identifiability of a familiar face would undergo a similar decline with uncertainty? Specifically, could the presence of an unnamed celebrity, thus *any* celebrity, be detected when presented in a rapid sequence of unfamiliar faces? If so, could the celebrity be identified? Despite the markedly greater physical similarity of faces compared to objects that are, say, not tools, the presence of a celebrity could be detected with moderately high accuracy (~75%) at rates exceeding 7 faces/s. False alarms were exceedingly rare as almost all the errors were misses. Detection accuracy by moderate congenital prosopagnosics was lower than controls, but still well above chance. Given the detection of the presence of a celebrity, all subjects were almost always able to identify that celebrity, providing no role for a covert familiarity signal outside of awareness.

1. Introduction

Rapid serial visual presentation (RSVP) paradigms have been extensively employed to assess the temporal limits of object recognition. In a typical version of the task, observers search for a target, specified by name. The target is visually masked in both the forwards and backwards directions by the preceding and following images and may be present in only half the sequences, imposing high perceptual and attentional demands as the observer must maintain attentional scrutiny throughout the sequence until a target is detected or the sequence terminates without a target detected. Moreover, the observer is faced with high uncertainty; not only does the observer not know if there will be a target in the sequence and, if one is present, where in the sequence it will occur, but (typically) does not know what its exact instantiation might be, e.g., the specific shape and pose of the object.

Uncertainty can be greatly increased with a "negative detection" version of the RSVP task, first studied by Intraub (1981), in which all the images in the sequence are from a common category, say "tools," and the observer is to detect an object that is *not* a tool. The set of objects that are not tools is, essentially, infinite. Intraub reported (1981) that at a duration of 114 ms/image, accuracy dropped from 71% when the target was specified by name, e.g., a "chair", to 35% when specified negatively, e.g., "Not a Tool."

Can faces be recognized in the extremely high uncertainty of a negative detection RSVP task? Most studies of face recognition performance require a same-different response to a single unfamiliar face, perhaps where the faces are rotated in depth (or translated or varied in size) to assess invariance, and they are presented either simultaneously (as in a match-to-sample task) or several seconds earlier as in the Cambridge Face Memory Test (CFMT) (Duchaine & Nakayama, 2006). Given the apparent difficulty of the CFMT, where even normal subjects find the test challenging, it is not unreasonable to hypothesize that face recognition under high uncertainty and extremely short masked exposure durations is impossible. The instantiation of such a paradigm in the present investigation presented sequences of unfamiliar faces in which the face of a familiar celebrity was present in half the sequences. The subject's task was to detect whether a celebrity headshot was present (i.e., to find the face that is *not* that of a *non*-celebrity) and, if so, to identify the celebrity. The issue of person recognition under high uncertainty is of some interest in that the ability to identify a familiar celebrity voice declines markedly as the number of possible celebrities is increased from one to only a handful (Legge, Grosman, & Pieper, 1984; Shilowich & Biederman, 2016; Xu et al., 2015). With an unlimited set of possible familiar celebrities, voice identification is almost

* Corresponding author at: Department of Psychology and Program in Neuroscience, University of Southern California, United States.
E-mail address: bieder@usc.edu (I. Biederman).

https://doi.org/10.1016/j.visres.2018.02.007
Received 28 July 2017; Received in revised form 25 February 2018; Accepted 27 February 2018

impossible. If the RSVP identification of faces shows a similar decline (compared to, say, identification of objects), then there may be a general decline in person perception under high uncertainty that is not evident when perceiving objects.

We assessed the ability of controls, moderate prosopagnosics, and two extreme prosopagnosics, one congenital (GJ) and the other acquired (MJH), to perform negative detection RSVP tasks with objects and celebrity and non-celebrity headshots. Studies of prosopagnosia typically compare individuals who, on the basis of some standardized tests, are at the extremes; either clearly prosopagnosic or not. However, the diagnostic tests for prosopagnosia yield graded scoring with individuals intermediate on a dimension of face recognition ability. If faces can be detected in a negative RSVP task and if the task engages the same processes that are deficient in prosopagnosia, then those classified as intermediate in face recognition might be expected to perform at a level intermediate between controls and those more extreme on the tests for prosopagnosia. We designate such an intermediate group as moderate congenital prosopagnosics (mCPs) and those at the more extreme as extreme prosopagnosics (xCPs). Although the term "Developmental Prosopagnosia" has been used to designate those individuals who are deficient in face recognition but who have no history of neurological insult or detectable lesions in face selective areas as distinct from those "Acquired Prosopagnosics" whose deficiency is a likely consequence of lesion or disease, we prefer the term "Congenital Prosopagnosia" as there is no evidence that early childhood experience can lead to prosopagnosia. Indeed, twin studies show a higher correlation in face recognition ability between monozygotic than dizygotic twins, suggesting a genetic linkage (Wilmer et al., 2010).

2. Method

2.1. Participants

Fifty-four students from the University of Southern California (52 right handed, 34 female, mean age of 20.3 years, range 18–38 years) participated for course credit or monetary compensation. From this distribution, 47 students served as controls, while six subjects (age range 19–21, three female) were classified as moderate congenital prosopagnosics (mCPs) given a) their unremarkable neurological history and b) a level of performance of one standard deviation below the mean on at least four of five diagnostic tests (Table 1, raw scores in appendix). The PI20 served as a self-report measure of face recognition ability. The CFMT assessed face perception, short-term face memory, and invariance to orientation (Duchaine & Nakayama, 2006). The USC Face Perception Test (USCFPT, link: testable.org/t/3732942e7), is a minimal match-to-sample task with a triangular display of three computer-generated faces, a sample (on top), with one of the two lower faces an exact match to the sample, the other being the distractor (Biederman, Margalit, Maarek, Meschke, & Shilowich, 2017). The subject indicates by key press whether the left or right face is the exact match. The display remains in view for 5 s or until the subject responds providing a relatively pure test of face perception, with virtually no contribution of memory. The Famous Faces Test (http://www.faceblind.org) and USC Celebrity Test (http://bit.ly/2Bd2dyP) are celebrity recognition tasks, reflecting long-term memory for faces.

Additionally, two "extreme" prosopagnosic subjects, xPros, were run. GJ, a 33-year-old male, was classified as a congenital prosopagnosic, xCP, on the basis of self-report, an interview, his survey scores, and no evidence of neurological incidents. The other was MJH, a 53-year-old male who is an acquired prosopagnosic, xAP, with bilateral lesions to OFA and FFA suffered as a result of a fall at the age of 5 (reported in Xu & Biederman, 2014). While MJH's simultagnosia and mild object agnosia reflect more generalized perceptual deficits than those presented by congenital prosopagnosics, his inclusion provides additional assessment of the RSVP task to differentiate those with face recognition difficulties from controls. All subjects reported normal or corrected-to-normal vision. The work was carried out in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki). Informed consent was obtained for experimentation with human subjects.

2.2. Stimuli

Stimuli were colored photographic images of either objects or faces (headshots) obtained from a Google image search. In the object sequences, all the images, but possibly one, belonged to a single category: tools, animals, modes of transportation, plants, or articles of clothing. All the images in the face sequences, but possibly one, were of non-celebrities or contained one of the 50 most familiar celebrities (half female) as rated by USC undergraduates in prior studies of voice recognition. Non-celebrity images were taken from websites with headshots of aspiring actors and business executives. In the judgment of the experimenters and several other observers, there was no discernible difference in image quality between the celebrities and the non-celebrities. The backgrounds were removed from all images and replaced with a homogeneous gray before being scaled to 800 by 800 pixels. Images of faces were cropped to show the full face and top of the shoulders of each person and images of objects were scaled to fit, approximately, within the center of a 19.3° square given the typical distance of the subject to the screen.

2.3. Design and Procedure

2.3.1. Familiarity ratings

Prior to the experimental trials, subjects rated their familiarity with the faces of 50 celebrities, listed by name, on a scale of 1–5 (unfamiliar, slightly familiar, moderately familiar, very familiar, most familiar). To assess whether providing the names of celebrities influenced their subsequent RSVP detection performance, eight additional subjects were run with their familiarity ratings made after their RSVP trials. As will be discussed later, there was no evidence of a detection benefit from rating the familiarity of the celebrities prior to their presence in the RSVP sequences.

Table 1
Mean Scores on Five Diagnostic Tests Distinguishing Controls, mCPs, and Two xPros.

Subject Classification	USC Face Perception Test[1]	Famous Faces Test	Cambridge Face Memory Test (CFMT)	USC Celebrity Test	PI20[2]
Controls	86.9% (72%–99%)	81.4% (27%–100%)	79.3% (57%–97%)	85.7% (52%–100%)	38.3 (23–72)
mCPs	71.2% (55%–89%)	38.3% (18%–65%)	56.2% (40%–65%)	50.3% (33%–77%)	65.3 (52–84)
xCP (GJ)	45%	18%	28%	25%	90
xAP (MJH)	52%	3%	38%	26%	83

[1] Chance on the USC Face Perception Test is 50%.
[2] All scores but those on the PI20 are percent correct. Scores on the PI20 are self-ratings with higher scores indicating greater difficulty in face recognition.