ELSEVIER

Archives of Clinical Neuropsychology 23 (2008) 175–187

Archives of CLINICAL NEUROPSYCHOLOGY

The Philadelphia Face Perception Battery

Amy L. Thomas [a], Kathy Lawler [a], Ingrid R. Olson [b], Geoffrey K. Aguirre [a,b,*]

[a] *University of Pennsylvania, Neurology Department, United States*
[b] *University of Pennsylvania, Center for Cognitive Neuroscience, United States*

Accepted 23 October 2007

Abstract

The Philadelphia Face Perception Battery (PFPB) tests four aspects of face perception: discrimination of facial similarity, attractiveness, gender, and age. Calibration with 116 neurologically intact subjects yielded average performance of ∼90%. Across subjects, there was a low correlation (<0.22) in performance between the tests (with the exception of the attractiveness and age discrimination tests) suggesting that the tests measure independent aspects of face perception. There were modest effects of subject demographic factors upon performance, and test–retest reliability scores (between 0.37 and 0.75) were comparable to other neuropsychological batteries. Modification of the stimuli to obscure internal facial features lowered performance on the age, gender, and attractiveness discrimination tests between 2 and 4 standard deviations. The clinical sensitivity of the battery was demonstrated by testing a patient with acquired prosopagnosia. She showed performance impairments of between 2 and 4 standard deviations on all sub-tests. The PFPB is freely available for non-commercial use.

Keywords: Face perception; Prosopagnosia; Facial beauty; Neuropsychological testing

1. Introduction

Neuropsychological tests provide standardized measures of cognitive processes. Face perception ability, and its neural correlates, has been the subject of intense study and frequently the target of neuropsychological investigation. Well described is the clinical deficit of prosopagnosia: an inability to perceive facial information despite ostensibly intact rudimentary visual perception. Initially understood as an acquired deficit following focal, ventral occipito-temporal lesions, prosopagnosia is now also recognized as a developmental condition in both general (Duchaine & Nakayama, 2006a) and special (e.g. autistic) populations (Blair, Frith, Smith, Abell, & Cipolotti, 2002; Cipolotti, Robinson, Blair, & Frith, 1999). Despite the great research interest in face processing deficits, formal neuropsychological tests to detect impairments in these core abilities of human social behavior have been fairly limited. The two most commonly used neuropsychological tests that employ face stimuli are Warrington's Recognition Memory Test (RMT) and the Benton Facial Recognition Test (BFRT). Despite their widespread application these tests have specific limitations for the measurement of face perception ability per se.

The RMT uses face stimuli and was well standardized with 300 intact and impaired individuals. The test involves presenting faces and words to a subject and asking them to judge if each stimulus is pleasant or unpleasant. Following a delay, the subjects are then asked to determine which of two faces or words they had been shown previously. This

* Corresponding author at: 3400 Spruce St., 3W. Gates, Philadelphia, PA 19104, United States. Tel.: +1 215 662 3390; fax: +1 215 349 8260.
E-mail address: aguirreg@mail.med.upenn.edu (G.K. Aguirre).

0887-6177/$ – see front matter
doi:10.1016/j.acn.2007.10.003

test was originally designed to evaluate deficits of verbal vs. non-verbal memory (Sweet, Demakis, Ricker, & Millis, 2000; Warrington, 1984), although recent evidence has suggested that the RMT may not be reliably sensitive to these deficits (Sweet et al., 2000). As the test employs face stimuli, it has often been used to measure impairments in face perception (e.g., Nunn, Postma, & Pearson, 2001). By design, however, the test is clearly not a process-pure measure of face perception, as impairments in memory alone could contribute to impaired performance. The 3rd edition of the Wechsler Memory Scale includes a subtest of face memory and similarly does not provide a pure measure of face perception.

The BFRT was designed to evaluate unfamiliar face perception. It is easily administered to a clinical population and is also well standardized (Benton, Sivan, Hamsher, Varney, & Spreen, 1983). Both short (27 items) and long form (54 items) versions of the test are used. The patient is first shown gray scale photographs of several faces under severe lighting conditions designed to eliminate the perception of hairstyle and other features not intrinsic to the face. Initially, the task is to identify a face that matches a target face from a set of six. Additional trials in the long form involve identifying three of the six faces that match a target face, but are displayed in different orientations. The BFRT is commonly used to measure face perception deficits (Nunn et al., 2001; Weniger, Boucsein, & Irle, 2004). While ostensibly a more direct measure of face perception per se, there is evidence that the test may be solved without reference to the internal facial features. Prosopagnosia is generally conceived as a specific impairment of perception of internal facial features (eyes, nose, mouth) and it has long been recognized that such patients can demonstrate some intact performance with faces by using other perceptual features of the head (e.g., hairline, shape of the jaw).

Duchaine and Weidenfeld (2003) administered a modified form of the BFRT (and the RMT) that occluded internal facial features to normal undergraduate volunteers. These subjects were able to maintain performance within the normal range using only non-internal face features, such as hairline and eyebrows. Duchaine and Nakayama (2004) then tested developmental prosopagnosics (DP), individuals with a congenital deficit of facial perception, on the BFRT and ascertained that they are able to achieve scores within the normal range.

In response to these findings, and to assist in the identification of neurologically intact individuals who have developmental impairments in face perception, Duchaine and Nakayama (2006a) created a test of face memory focused upon internal facial features. The Cambridge Face Memory Test (CFMT) first presents six target faces to the participant and then 72 target detection, three-alternative forced choice trials, each involving one target face with views identical to the target stimuli, novel views, or novel views with noise. In tests of eight subjects who self-describe as having developmental prosopagnosia, six were found to be impaired on the CFMT.

While a substantial advance over extant methods, the CMFT is not suitable for all applications. First, the CMFT relies upon a single aspect of face perception to identify impairments. LeGrand et al. (2006) have argued that sensitivity to developmental impairments in face perception are best identified using multiple measures of face information. They tested DP patients on a series of tasks to evaluate their sensitivity to global form and motion, face detection, holistic face processing, perception of facial identity, the ability to determine gender from face stimuli, and whether attractiveness ratings were similar to controls. Their results demonstrated that DP patients are sensitive to global motion and form and perform normally in detection of faces, holistic processing, and gender discriminations. But they were impaired in processing of facial identity and judgments of attractiveness. LeGrand et al. (2006) concluded that facial deficits, specifically DP, cannot be reliably determined by poor performance on any single task.

Second, the CMFT is explicitly a memory test, and has fairly complicated instructions and motor response requirements. While ideal for screening an otherwise normal population for isolated face perception deficits, it is less well suited to neurologically impaired populations. For example, it could not be used to examine perceptual impairments in patients with organic dementing illnesses such as Alzheimer's disease.

We describe here the creation of a set of tests to examine face perception ability. Ideally, these tests would be sensitive to and specific for impairments in the perception of internal facial features. To allow broad application to neurologically impaired populations, the tests would have minimal memory, instruction, and motor response requirements, with a roughly 30 min administration time. Following the lead of Le Grand and colleagues, the measurement of several different aspects of face perception might provide for improved sensitivity. Human observers can rapidly derive from facial appearance information about facial identity, gender, age, and relative attractiveness. Using photo-realistic, synthetic face stimuli, we created a set of two-alternative forced-choice discrimination tasks that draw upon these aspects of facial perception. As many as 116 subjects from a broad range of demographic backgrounds completed the tasks, allowing the identification of a subset of trials that provide good test performance. We then demonstrated the relative independence of performance on each subcomponent of the test, the good test–retest reliability of the