Probing short-term face memory in developmental prosopagnosia

Punit Shah\textsuperscript{a,b}, Anne Gaule\textsuperscript{a,c}, Sebastian B. Gaigg\textsuperscript{a}, Geoffrey Bird\textsuperscript{b,d}, & Richard Cook\textsuperscript{a*}

\textsuperscript{a}Department of Psychology, City University London, London, U.K.

\textsuperscript{b}MRC Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology, and Neuroscience, King’s College London, University of London, London, U.K.

\textsuperscript{c}Division of Psychology and Language Sciences, University College London, University of London, London, U.K.

\textsuperscript{d}Institute of Cognitive Neuroscience, University College London, University of London, London, U.K.

*Corresponding author: Richard.Cook.1@city.ac.uk

Department of Psychology,
City University London,
Whiskin Street
London, EC1R OJD
Tel: +44 20 7040 8644, Fax: +44 20 7040 8877
HIGHLIGHTS

- 16 developmental prosopagnosics and 16 controls completed a match-to-sample task
- Face-specific impairments were evident following short and long retention-intervals
- Face-matching performance was insensitive to the duration of the retention-interval
- Prosopagnosics form stable, albeit inaccurate perceptual descriptions of faces
ABSTRACT
It has recently been proposed that the face recognition deficits seen in neurodevelopmental disorders may reflect impaired short-term face memory. For example, introducing a brief delay between the presentation of target and test faces seems to disproportionately impair matching or recognition performance in individuals with Autism Spectrum Disorders. The present study sought to determine whether deficits of short-term face memory contribute to impaired face recognition seen in Developmental Prosopagnosia. To determine whether developmental prosopagnosics exhibit impaired short-term face memory, the present study used a six-alternative-forced-choice match-to-sample procedure. Memory demand was manipulated by employing a short or long delay between the presentation of the target face, and the six test faces. Crucially, the perceptual demands were identical in both conditions, thereby allowing the independent contribution of short-term face memory to be assessed. Prosopagnosics showed clear evidence of a category-specific impairment for face-matching in both conditions; they were both slower and less accurate than matched controls. Crucially however, the prosopagnosics showed no evidence of disproportionate face recognition impairment in the long-interval condition. While individuals with developmental prosopagnosia may have problems with the perceptual encoding of faces, it appears that their representations are stable over short durations. These results suggest that the face recognition difficulties seen in developmental prosopagnosia and autism may be qualitatively different, attributable to deficits of perceptual encoding and perceptual maintenance, respectively.

Key words: face perception, body perception, Developmental Prosopagnosia, short-term face memory, neurodevelopmental disorders
1. INTRODUCTION

Developmental Prosopagnosia (DP) is a neurodevelopmental disorder characterized by impaired face recognition, despite normal intelligence, low-level vision, and broader social cognition (Behrmann & Avidan, 2005; Duchaine & Nakayama, 2006b; McConachie, 1976; Susilo & Duchaine, 2013). While individuals with DP typically learn to recognize others using cues such as voice and hairstyle, the condition is often associated with detrimental psychosocial consequences (Yardley, McDermott, Pisarski, Duchaine, & Nakayama, 2008). Current estimates, inferred from performance on computerised tasks (Bowles et al., 2009) and self-report measures (Kennerknecht et al., 2006), suggest that the prevalence of DP in the general population is approximately 2% (Susilo & Duchaine, 2013). While its origins remain poorly understood, DP frequently runs in families, indicating a genetic component (Dobel, Bölte, Aicher, & Schweinberger, 2007; Duchaine, Germine, & Nakayama, 2007; Grüter et al., 2007; Johnen et al., 2014). Differences in cortical structure (Behrmann, Avidan, Gao, & Black, 2007; Garrido et al., 2009), structural (Thomas et al., 2009) and functional connectivity (Avidan & Behrmann, 2009) have been observed in inferotemporal regions including the fusiform gyrus, a region crucial for face processing.

It has recently been proposed that face recognition deficits seen in neurodevelopmental disorders may reflect impaired short-term face memory (STFM; Weigelt, Koldewyn, & Kanwisher, 2012). Where face recognition difficulties are seen in Autism Spectrum Disorder (ASD), tasks often require participants to retain faces in memory (Arkush, Smith-Collins, Fiorentini, & Skuse, 2013; Boucher & Lewis, 1992; Hedley, Brewer, & Young, 2011; Weigelt, Koldewyn, & Kanwisher, 2013). Introducing a delay of a few seconds between target and test faces seems to disproportionately impair matching or recognition performance (Weigelt et al., 2012). Nevertheless, participants with ASD often demonstrate broadly typical face perception, exhibiting inversion effects (Scherf, Behrmann, Minshew, & Luna, 2008), behavioural markers of holistic representation (Nishimura, Rutherford, & Maurer, 2008), and intact memory for non-face stimuli (Arkush et al., 2013; Boucher & Lewis, 1992; Weigelt, Koldewyn, & Kanwisher, 2013).

The view that faces recruit domain-specific perceptual processing has proved controversial (Diamond & Carey, 1986; McKone & Robbins, 2011). The suggestion that
STFM can be selectively impaired is important because it raises the further possibility that face-specific neurocognitive mechanisms are also seen in the domain of memory. One possibility is that domain-specific mechanisms responsible for maintaining face percepts are dysfunctional in some neurodevelopmental populations. Consequently, initially accurate perceptual representations may be less stable and rapidly degrade. The implied dissociation between perceptual processes responsible for encoding, and memory processes responsible for maintaining face representations, is consistent with evidence that face memory follows a different developmental trajectory relative to perceptual memory for other objects (Weigelt, Koldewyn, Dilks et al., 2014).

The present study sought to determine whether aberrant STFM, specifically impaired perceptual maintenance, contributes to face recognition difficulties in DP. Many neurodevelopmental disorders, including ASD and prosopagnosia (Barton et al., 2004), are thought to co-occur, suggestive of common causal factors (Bird & Cook, 2013; Rutter et al., 2011; Visser, 2003). However, unlike individuals with ASD, most DPs have problems with the perceptual encoding of faces. Participants with DP often perform poorly on tasks that tax perception in the absence of a memory demand (Duchaine et al., 2007). Similarly, DPs often show reduced inversion effects (Behrmann, Avidan, Marotta, & Kimchi, 2005) and evidence of diminished holistic representation (Avidan, Tanzer, & Behrmann, 2011). However, surprisingly little is known about STFM in DP. Individuals with DP typically score well below controls on tasks that require participants to memorise faces for subsequent test (Duchaine & Nakayama, 2006a; Duchaine, Parker, & Nakayama, 2003; Duchaine, Yovel, Butterworth, & Nakayama, 2006). However, it is unclear whether these difficulties reflect impaired encoding, perceptual maintenance, or both.

Consistent with possible deficits of STFM, cases of DP have been reported where delayed face recognition is disproportionately impaired, relative to performance on perceptual face-matching tasks (McKone et al., 2011). When DPs are required to retain faces in memory for brief periods, functional magnetic resonance imaging (fMRI) reveals wider activation in prefrontal regions implicated in working memory, relative to controls (Avidan, Hasson, Malach, & Behrmann, 2005), suggesting that STFM may be effortful. Finally, developmental cases exist who show atypical fMRI adaptation to faces. Repeated presentation of unfamiliar faces typically elicits attenuated responses in the Fusiform Face
Area (FFA), indicative of short-term learning. However, Case C exhibited no repetition suppression, suggesting that her FFA may support unstable face representations (Williams, Berberovic, & Mattingley, 2007).

To determine whether DPs exhibit impaired STFM, the present study used a six-alternative-forced-choice (6AFC) match-to-sample procedure. Whereas previous studies employing match-to-sample designs have employed a single interval (Dobel et al., 2007; Lobmaier, Bölte, Mast, & Dobel, 2010), the present study manipulated memory demand by varying the delay between the presentation of target and test faces. The perceptual demands of the resulting short- and long-interval conditions were identical, allowing the independent contribution of STFM to be assessed. Should DPs have problems maintaining face percepts over short durations, differences between controls and DPs should be larger at longer retention-intervals. If deficits reflect impairments of a domain-specific mechanism, differences should be seen with faces, but not for other within-class discriminations.

2. METHOD
2.1 Participants
Participants were 32 right-handed adults, 16 with (12 males; \(M_{\text{age}} = 47.2\) years, \(SD_{\text{age}} = 17.8\) years) and 16 without DP (11 males; \(M_{\text{age}} = 45.5\) years, \(SD_{\text{age}} = 14.3\) years). The DP and control groups did not differ significantly in age \([t(30) = .295, p = .770]\) or proportion of females \([\chi^2(1) = .08, p = .777]\).

2.2 Diagnostic Procedures
Participants completed a series of computer-based tasks testing their face recognition and wider visual abilities. Figure 1 shows the performance of both groups on the diagnostic procedures. The scores of each DP are shown in Table 1. While diagnostic evidence accumulated across a number of procedures, each the DPs scored less than two SDs below the control group mean on the Cambridge Face Memory Test.

Table-1
2.2.1. *Famous Face Recognition Test.* This test assesses recognition of familiar faces. Participants had to identify 34 international celebrities (actors, singers, sporting stars, politicians), from cropped photographic images, by providing their name or other identifying information about the individual. Scores reflect the number of correct identifications expressed as a proportion of the total number of celebrities with which each participant was familiar. The DP group ($M = 34.0\%, SD = 15.5\%$) scored significantly worse [$t(30) = 7.039, p < .001$] than controls ($M = 74.8\%, SD = 17.2\%)$.

2.2.2. *Cambridge Memory Tests.* The Cambridge Face Memory Test (CFMT; Duchaine & Nakayama, 2006a) and the Cambridge Car Memory Test (CCMT; Dennett et al., 2011) assess recognition of unfamiliar faces and cars. The tests, each comprising 72 trials, employ identical formats. Participants are required to learn exemplars in a training phase and then identify the trained exemplars in a 3AFC procedure. Difficulty is varied by presenting items from different viewpoints and through the addition of visual noise. The DP group was disproportionately impaired on the CFMT [$F(1,30) = 15.532, p < .001$]. Whereas controls’ scores ($M = 85.2\%, SD = 10.5\%$) exceeded those of the DPs ($M = 53.4\%, SD = 9.7\%$) on the CFMT [$t(30) = 8.941, p < .001$], the performance of the controls ($M = 79.2\%, SD = 14.3\%$) and DPs ($M = 70.7\%, SD = 14.7\%$) on the CCMT did not differ significantly [$t(30) = 1.665, p = .106$].

2.2.3. *Cambridge Face Perception Test* (CFPT; Duchaine et al., 2007). This test assesses face perception ability in such a way as to minimize the memory demand. Trials present a target face and a series of six faces that resemble the target to varying degrees. Participants have 60 secs to sort the six in order of target-face similarity. Eight trials present the target and test faces upright, eight present the faces inverted. Trials are scored by calculating deviations from the correct order. Controls showed a greater advantage for upright presentation, than the DPs [$F(1,30) = 23.779, p < .001$]. Whereas controls ($M = 30.4, SD = 11.0$) made significantly fewer errors than DPs ($M = 59.9, SD = 22.8$) in the upright condition [$t(30) = 4.665, p < .001$], performance of the DPs ($M = 69.3, SD = 15.1$) and controls ($M = 64.8, SD = 14.1$) was comparable in the inverted condition [$t(30) = .859, p = .397$].
2.3 Experimental Stimuli & Materials

Four stimulus sets were used: Caucasian male faces, Caucasian male hands, butterflies and wooden dining chairs (Figure 2a). The hands provided a non-face body-part control condition, whereas the butterflies and chairs provided additional animate and inanimate control conditions. Each set comprised 34 exemplars presented in greyscale. The faces were taken from the Radboud (Langner et al., 2010) and Karolinska (Lundqvist, Flykt, & Öhman, 1998) face databases. Faces were cropped so that hairline and external features were not visible. All had neutral expressions. The chair and butterfly stimuli were downloaded from various online sources. The hand stimuli were purposely created for the study. Faces, hands, butterflies and chairs subtended 8°, 8°, 6°, and 9° vertically, when viewed at a distance of 60 cm. The experimental program was written and presented in Matlab using Psychtoolbox (Brainard, 1997; Pelli, 1997).

2.4 Experimental Procedure

The experiment employed a 6AFC match-to-sample procedure (Figure 2b). Trials began with a target stimulus presented for 1 sec, followed by a retention-interval of 2 secs (low-demand) or 8 secs (high-demand), during which a mask of high-frequency greyscale noise was presented. The mask was replaced by six test stimuli, one of which was identical to the target. Participants were required to identify which of the six images was the target. The remaining five stimuli were foils chosen at random from the same stimulus set. Test stimuli were presented at the same scale as the target, and were visible until participants responded with a keypress. A given exemplar could only appear as a target once in each retention condition. Response times (RTs) were measured from the onset of the test arrays, to the register of the keypress response. Participants completed 6 practice trials, followed by 224 trials (28 trials × 4 stimulus classes × 2 retention-intervals). Trial type was interleaved randomly within mini-blocks of 56 trials.

Figure-2

3. RESULTS

The accuracy achieved by each participant in the eight conditions was computed together with mean RTs (Figure 3). Overall matching accuracy for faces correlated closely with both the CFMT \( r = .733, p < .001 \) and CFPT \( r = .669, p < .001 \). No significant
correlations were observed between matching accuracy for hands, chairs, or butterflies, and either the CFMT or CFPT (all $r$’s < .30, $p$’s > .10). RTs exceeding 3 SDs of a participant’s mean RT were excluded. In total 1.76% and 2.09% of RTs were excluded for controls and prosopagnosics, respectively. The resulting distributions were analysed using mixed model ANOVAs with stimulus (faces, hands, butterflies, chairs) and retention-interval (long, short) as within-subjects factors, and group (controls, prosopagnosics) as a between-subjects factor.

3.1 Accuracy
The accuracy analysis revealed significant main effects of retention-interval [$F(1,30) = 6.569, p = .016$] and stimulus [$F(3,90) = 42.561, p < .001$]. Accuracy was better on short-interval trials ($M = 84.8\%$, $SD = 7.1\%$) than on long-interval trials ($M = 82.8\%, SD = 8.2\%$), confirming the effectiveness of the manipulation. Participants were less accurate at recognising hands than faces [$t(31) = 4.281, p < .001$], butterflies [$t(31) = 10.798, p < .001$] and chairs [$t(31) = 13.192, p < .001$]. Accuracy for faces was also worse than for chairs [$t(31) = 3.018, p = .005$] and marginally worse than for butterflies [$t(31) = 1.971, p = .058$]. There was no retention-interval × stimulus interaction [$F(3,90) = 1.672, p = .179$].

A main effect of group was also observed [$F(1,30) = 7.150, p = .012$] whereby controls ($M = 86.9\%, SD = 5.7\%$) were more accurate than prosopagnosics ($M = 80.7\%, SD = 7.4\%$). Crucially however, a significant group × stimulus interaction was observed [$F(3,90) = 8.211, p < .001$]. Prosopagnosics showed a marked reduction in face-matching accuracy [$t(30) = 3.880, p = .001$], seen on both short- [$t(31) = 4.443, p < .001$] and long-interval trials [$t(31) = 3.258, p = .005$]. Their performance was comparable to controls for hands [$t(30) = .982, p = .334$], butterflies [$t(30) = .245, p = .808$], and chairs [$t(30) = .798, p = .431$]. Finally, neither a group × retention-interval [$F(1,30) = .006, p = .938$], nor a group × retention-interval × stimulus interaction was observed [$F(3,90) = .578, p = .631$], indicating that effects of retention-interval were comparable for both groups.

3.2 Response times
The analysis of participants’ RTs revealed significant main effects of retention-interval [$F(1,30) = 64.325, p < .001$] and stimulus [$F(3,90) = 36.725, p < .001$]. Participants responded slower on long-interval trials ($M = 4.30$ secs, $SD = 1.31$ secs) than on short-
interval trials ($M = 3.54$ secs, $SD = .95$ secs), providing additional evidence for the effectiveness of the manipulation. Participants were slower to respond to hands than to faces [$t(31) = 2.095, p = .044$], butterflies [$t(31) = 10.840, p < .001$] and chairs [$t(31) = 12.910, p < .001$]. Responses on face trials were also slower than for chairs [$t(31) = 4.158, p < .001$] and butterflies [$t(31) = 3.939, p < .001$]. There was no retention-interval × stimulus interaction [$F(3,90) = 1.638, p = .168$].

No main effect of group was observed for RT [$F(1,30) = 1.602, p = .215$], however a significant group × stimulus interaction was revealed [$F(3,90) = 6.995, p < .001$]. Prosopagnosics were slower than controls on face trials [$t(30) = 2.688, p = .012$], a difference seen on both short- [$t(30) = 2.625, p = .014$] and long-interval trials [$t(30) = 2.523, p = .017$]. In contrast, RTs for hands [$t(30) = .515, p = .611$], butterflies [$t(30) = .302, p = .765$], and chairs [$t(30) = .872, p = .390$] were comparable. Neither a group × retention-interval [$F(1,30) = 1.301, p = .263$] nor a group × retention-interval × stimulus interaction was observed [$F(3,90) = 1.189, p = .318$].

**4. DISCUSSION**

The present study sought to determine whether aberrant STFM, specifically impaired perceptual maintenance, contributes to face recognition difficulties in DP. Conditions of high and low memory demand were created by varying the interval between target and test faces, a manipulation that keeps the perceptual demands constant. Should DPs have problems with perceptual maintenance, any group difference observed on short- interval trials, should increase under conditions of greater retention demand. The results confirmed the effectiveness of the memory manipulation: participants were slower and less accurate at identifying the targets in the long-interval condition. Moreover, prosopagnosics showed clear evidence of a category-specific impairment for face-matching in both conditions. Crucially however, DPs showed no evidence of disproportionate face recognition impairment at longer retention-intervals.

While the perceptual encoding of faces may be impaired in DP; for example descriptions of target or test faces may be less accurate or less differentiated, representations appear to
be stable over short durations. Should perceptual maintenance of face percepts be aided by domain-specific neurocognitive mechanisms, we find no evidence that these processes are impaired in DP. While face recognition difficulties in ASD may be associated with impaired STFM (Weigelt et al., 2012), the problems seen in DP may more frequently relate to perceptual encoding. It is important to note, however, that DP is a heterogeneous condition (Johnen et al., 2014; Susilo & Duchaine, 2013), and individuals may be identified who exhibit impaired STFM (Williams et al., 2007). Nevertheless, while exceptions may be identified, impaired maintenance of face percepts does not appear to be characteristic of DP.

The suggestion that STFM processes are unimpaired in DP may seem counterintuitive given previous reports of poor performance on face memory tests (Duchaine & Nakayama, 2006a; Duchaine et al., 2003). Indeed the CFMT is widely regarded as a key diagnostic tool in DP research (Bowles et al., 2009; McKone et al., 2011). However, because tests of ‘face memory’ typically incorporate perceptual encoding, maintenance and retrieval demands, the locus of impairment is unclear. The present data suggest that difficulties with perceptual encoding may contribute substantially to the low scores of DPs on these tasks. Whether DPs exhibit additional deficits of long-term face learning remains an open empirical question.

To distinguish between perceptual encoding and maintenance, identical images were used during target presentation and test. The observation of a face-specific deficit on such a simple matching task is a striking finding. Matching identical instances of a target face is undeniably easier than matching different instances (Burton & Jenkins, 2011). This finding confirms however, that simple face matching recruits face-specific perceptual ability, particularly when external features are occluded (cf. Megreya & Burton, 2006). CFMT and CFPT scores correlated closely with face matching performance, but not with matching of non-face stimuli, further confirming the validity of our face matching measure.

We have focussed on a particular aspect of STFM, the perceptual maintenance of faces over short durations, thought to be deficient in ASD (Weigelt et al., 2012). However, it remains important to determine whether other aspects of STFM are intact in DP. For
example, percept manipulation distinguishes visual working memory from simple perceptual maintenance (Baddeley, 1992). Because matching across viewpoints requires an additional mental rotation process, this manipulation might allow future comparison of perceptual maintenance and working memory in DP. It is also important to study perceptual maintenance further. Future studies could probe the decay of percepts over longer retention-intervals and under conditions of load induced by the encoding of multiple targets.

Finally, these data provide further indication that face recognition can be selectively impaired, in the absence of wider deficits of within-class discrimination (Duchaine & Nakayama, 2006b; Susilo & Duchaine, 2013). Of particular interest, DPs showed impaired face-matching, despite normal hand-matching. This is noteworthy for two reasons. First, the hands task was more challenging than the face condition, indicated by poorer accuracy and slower responses. That the DPs performed normally confirms that the face recognition deficits observed were not an artefact of task difficulty. Second, there are notable parallels between the perceptual mechanisms recruited by faces and bodies (Peelen & Downing, 2007). However, these results, together with previous findings (Pitcher, Charles, Devlin, Walsh, & Duchaine, 2009; Susilo, Yovel, Barton, & Duchaine, 2013), indicate that these mechanisms dissociate.
REFERENCES


Weigelt, S., Koldewyn, K., & Kanwisher, N. (2013). Face recognition deficits in autism spectrum disorders are both domain specific and process specific. *PLoS One, 8*(9), e74541.


FOOTNOTES

1 We use the term Developmental Prosopagnosia in preference to Congenital Prosopagnosia to reflect the possibility that the condition emerges during development, and may not necessarily be present from birth.

2 No concurrent task was employed during the interval to guard against unwanted interactions with stimulus type. It was reasoned that additional verbal, visual or numerical task demands could affect retention of some stimulus classes more than others. The delay interval of two secs used in the low-demand condition is in line with the threshold suggested by Weigelt et al. (2012). An interval of eight secs was employed in the high-demand condition to constrain the duration of the procedure.

3 The use of identical images during the encoding and test phases of trials allowed us to study perceptual encoding and maintenance without the additional demands of perceptual manipulation; i.e., without the need to resolve lighting or viewpoint disparities. The study of perceptual maintenance is a necessary first step in elucidating STFM in DP as abnormalities at this fundamental stage will impact on related processes, including percept manipulation.
FIGURES

Figure 1: Results from diagnostic tests conducted on the sample of developmental prosopagnosics and matched controls. (a) Lower scores on the Famous Face Recognition test indicate impaired recognition of familiar faces. (b) Results from the Cambridge Memory Tests indicate impaired recognition of unfamiliar faces despite typical recognition of unfamiliar cars. (c) The pattern of errors on the Cambridge Face Perception Test indicates that perception of upright faces is impaired in the sample of developmental prosopagnosics, relative to controls, while the perception of inverted faces is comparable in the two groups.
Figure 2: (a) Examples of the four classes of stimulus – Caucasian male faces, Caucasian male hands, butterflies and wooden dining chairs – used in the experiment. (b) Illustration of the six-alternative-match-to-sample procedure. Memory demand was manipulated by increasing the interval between the presentation of the target and the onset of the response display. Because the perceptual demands of the long- and short-interval conditions are identical, this paradigm allows the contribution of STFM to be isolated and assessed.
Figure 3: (a) Accuracy data for the prosopagnosic and control groups. The dashed line indicates chance performance (16.7%) in the 6-alternative-forced-choice match-to-sample procedure. (b) The mean response times for the prosopagnosic and control groups. Both the accuracy and response time analyses indicate that prosopagnosics were selectively impaired at faces, and that this deficit is comparable following short and long retention-intervals.
TABLES

Table 1: The scores achieved on the diagnostic tests by each member of the prosopagnosic sample. Scores on the Famous Faces Test (FF), the Cambridge Face Memory Test (CFMT), and Cambridge Car Memory Test (CCMT), reflect % correct. Scores in the Cambridge Face Perception Test (CFPT) reflect total deviation errors.

<table>
<thead>
<tr>
<th>Case</th>
<th>Age</th>
<th>Gender</th>
<th>FF (%)</th>
<th>CFMT (%)</th>
<th>CCMT (%)</th>
<th>CFPT Upright</th>
<th>CFPT Inverted</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>20</td>
<td>M</td>
<td>52**</td>
<td>60**</td>
<td>82</td>
<td>50*</td>
<td>60</td>
</tr>
<tr>
<td>2</td>
<td>24</td>
<td>M</td>
<td>59</td>
<td>60**</td>
<td>50**</td>
<td>60**</td>
<td>78</td>
</tr>
<tr>
<td>3</td>
<td>27</td>
<td>M</td>
<td>44*</td>
<td>63**</td>
<td>69</td>
<td>46*</td>
<td>60</td>
</tr>
<tr>
<td>4</td>
<td>31</td>
<td>M</td>
<td>31**</td>
<td>56**</td>
<td>65</td>
<td>30</td>
<td>70</td>
</tr>
<tr>
<td>5</td>
<td>33</td>
<td>F</td>
<td>46*</td>
<td>60**</td>
<td>74</td>
<td>78***</td>
<td>84*</td>
</tr>
<tr>
<td>6</td>
<td>36</td>
<td>M</td>
<td>24**</td>
<td>57**</td>
<td>56*</td>
<td>42*</td>
<td>60</td>
</tr>
<tr>
<td>7</td>
<td>42</td>
<td>M</td>
<td>44*</td>
<td>58**</td>
<td>93</td>
<td>52*</td>
<td>50</td>
</tr>
<tr>
<td>8</td>
<td>45</td>
<td>M</td>
<td>15***</td>
<td>51***</td>
<td>94</td>
<td>86***</td>
<td>54</td>
</tr>
<tr>
<td>9</td>
<td>48</td>
<td>F</td>
<td>30**</td>
<td>58**</td>
<td>86</td>
<td>34</td>
<td>52</td>
</tr>
<tr>
<td>10</td>
<td>51</td>
<td>F</td>
<td>42*</td>
<td>46***</td>
<td>64*</td>
<td>74***</td>
<td>94**</td>
</tr>
<tr>
<td>11</td>
<td>57</td>
<td>M</td>
<td>48*</td>
<td>61**</td>
<td>53*</td>
<td>32</td>
<td>52</td>
</tr>
<tr>
<td>12</td>
<td>59</td>
<td>M</td>
<td>3***</td>
<td>49***</td>
<td>82</td>
<td>56**</td>
<td>64</td>
</tr>
<tr>
<td>13</td>
<td>67</td>
<td>M</td>
<td>10***</td>
<td>28***</td>
<td>47**</td>
<td>92***</td>
<td>78</td>
</tr>
<tr>
<td>14</td>
<td>69</td>
<td>F</td>
<td>32**</td>
<td>36***</td>
<td>76</td>
<td>100***</td>
<td>92*</td>
</tr>
<tr>
<td>15</td>
<td>73</td>
<td>M</td>
<td>34**</td>
<td>60**</td>
<td>67</td>
<td>42*</td>
<td>70</td>
</tr>
<tr>
<td>16</td>
<td>73</td>
<td>M</td>
<td>30**</td>
<td>53***</td>
<td>72</td>
<td>84***</td>
<td>90*</td>
</tr>
<tr>
<td><strong>Control mean</strong></td>
<td>74.8</td>
<td>85.2</td>
<td>79.2</td>
<td>30.4</td>
<td>64.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Control SD</strong></td>
<td>17.2</td>
<td>10.5</td>
<td>14.3</td>
<td>11.0</td>
<td>14.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Best control</strong></td>
<td>100</td>
<td>99</td>
<td>100</td>
<td>10</td>
<td>36</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Worst control</strong></td>
<td>34</td>
<td>67</td>
<td>56</td>
<td>52</td>
<td>84</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: *differs from control mean one SD; **differs from control mean two SDs; ***differs from control mean three SDs.