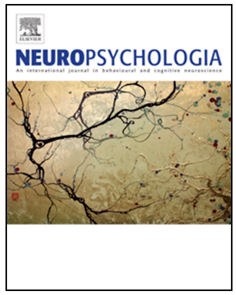
Journal Pre-proof



Developmental prosopagnosics and super recognizers rely on the same facial features used by individuals with normal face recognition abilities for face identification

Naphtali Abudarham, Sarah Bate, Brad Duchaine, Galit Yovel

PII: S0028-3932(21)00216-5

DOI: https://doi.org/10.1016/j.neuropsychologia.2021.107963

Reference: NSY 107963

To appear in: *Neuropsychologia*

Received Date: 19 February 2021

Revised Date: 13 May 2021

Accepted Date: 16 July 2021

Please cite this article as: Abudarham, N., Bate, S., Duchaine, B., Yovel, G., Developmental prosopagnosics and super recognizers rely on the same facial features used by individuals with normal face recognition abilities for face identification, *Neuropsychologia*, https://doi.org/10.1016/j.neuropsychologia.2021.107963.

CRediT author statement

Napthali Abudarham: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Software; Validation; Visualization; Roles/Writing - original draft;
Sarah Bates: Data curation; Writing - review & editing.
Brad Duchaine: Data curation; Writing - review & editing
Galit Yovel: Conceptualization; Formal analysis; Investigation; Methodology; Project administration; Resources; Supervision; Validation; Visualization; Roles/Writing - original draft;

Developmental prosopagnosics and super recognizers rely on the same facial features used by individuals with normal face recognition abilities for face identification

Naphtali Abudarham[1], Sarah Bate[2], Brad Duchaine[3] & Galit Yovel[1,4]

1 School of Psychological Sciences, Tel Aviv University, Tel Aviv, Israel
2 Department of Psychology, Bournemouth University, UK
3 Department of Psychological and Brain Sciences, Dartmouth College, Hanover, NH, USA
4 Sagol School of Neuroscience, Tel Aviv University, Tel Aviv, Israel

Corresponding regarding this manuscript
Galit Yovel
gality@tauex.tau.ac.il
School of Psychological Sciences
Sagol School of Neuroscience
Tel Aviv University

Abstract

Face recognition depends on the ability of the face processing system to extract facial features that define the identity of a face. In a recent study we discovered that altering a subset of facial features changed the identity of the face, indicating that they are *critical* for face identification. Changing another set of features did not change the identity of a face, indicating that they are *not critical* for face identification. In the current study, we assessed whether developmental prosopagnosics (DPs) and super recognizers (SRs) rely more heavily on critical features than non-critical features for face identification. To that end, we presented to DPs and SRs faces in which either the critical or the non-critical features were manipulated. In Study 1, we presented SRs with a famous face recognition task. We found that overall SRs recognized famous faces that differ in either critical or non-critical features better than controls. Similar to controls, changes in critical features had a larger effect on SRs' face recognition than changes in non-critical features. In Study 2, we presented an identity matching task to DPs and SRs. Similar to controls, DPs and SRs perceived faces that differed in critical features as more different than faces that differed in non-critical features. Taken together, our results indicate that SRs and DPs use the same critical features as normal individuals for face identification. These findings emphasize the fundamental role of this subset of features for face identification.

Keywords: Face recognition, Face perception, Super recognizers, Developmental Prosopagnosia, Features.

Highlights

- Super recognizers (SRs) and developmental prosopagnosics (DPs) were tested.
- SRs and DPs use the same critical features that controls use for face identification.
- SRs were less sensitive than controls to changes in non-critical features.
- These critical features play a central role in face identification.

Introduction

Face identification is a computationally challenging task that depends on the ability to generalize across different appearances of the same identity and discriminate between different individuals. Computational and neural investigations have shown that the face processing system represents faces in a hierarchical manner, starting with view-specific features in posterior areas, followed by the processing of high-level, view-invariant features in more anterior areas (Chang & Tsao, 2017; Freiwald & Tsao, 2010; Riesenhuber & Poggio, 1999). These high-level features enable identification of faces across different head views (Chang & Tsao, 2017, Abudarham et al, 2020).

To discover which features define the identity of a face, previous studies have manipulated various facial features and examined the effects of these manipulations on performance in face perception or identification tasks. For example, many studies have examined the effect of changing facial parts or the spacing among parts on face matching tasks. Whereas earlier studies emphasized the critical role of spacing among facial features (Le Grand, Mondloch, Maurer, & Brent, 2001; Maurer, Le Grand, & Mondloch, 2002; Mondloch, Le Grand, & Maurer, 2002), later studies have indicated that both spacing and parts are important for face identification (Mckone & Yovel, 2009; Sergent, 1984; Tanaka & Sengco, 1997;Yovel & Duchaine, 2006; Yovel & Kanwisher, 2004). Other studies have examined the roles of external vs. internal facial features, indicating the importance of external features for unfamiliar face identification and internal features for familiar face identification (Ellis, Shepherd, & Davies, 1979; Kramer, Towler, Reynolds, & Burton, 2018). Another approach to reveal which information is critical for face recognition has employed the “bubbles” technique, which assesses the importance of different regions of the face by revealing small regions in a random manner and assessing their influence on performance on a face identification task (Butler, Blais, Gosselin, Bub, & Fiset, 2010; Gosselin & Schyns, 2001). These studies indicated that the eye, eyebrow and mouth regions are important for face identification. Taken together, studies that employed different paradigms have emphasized the importance of different facial features for face identity. Nevertheless, none of these studies have provided an account for how the critical facial information enables the generation of a view-invariant face representation.

In a recent study, Abudarham and Yovel (2016) used a different reverse-engineering approach to reveal which features are critical for face identification. Participants were asked to rate the similarity of a set of twenty facial features between pairs of different identity faces that were presented from the same head view or different head views (e.g., which face has thicker lips? closer eyes?). They found that perceptual sensitivity to detect differences varied across facial features (Abudarham & Yovel, 2016; Abudarham, Shkiller, & Yovel, 2019). There were some facial features for which humans showed high perceptual sensitivity within and across head-views. These features included hair, eyebrow thickness, eye shape, eye color and lip thickness. They further found that replacing these features with features taken from other faces changed the identity of a face (Figure 1A, top row). These findings indicate that these features are critical for face identification. In contrast, replacing another set of features, for which participants showed lower perceptual sensitivity on the same feature similarity rating task, did not change the identity of the face. These non-critical features included eye distance, face proportion, nose size, mouth size, and skin color (Figure 1A, bottom). Importantly, faces that differ in critical and non-critical features were matched for low-level image-based differences. Taken together, these findings suggest that humans rely on a subset of facial features that enable face identification across different head views (see Figure 1B).

The findings reported above are based on the performance of individuals with face recognition abilities within the normal range. However, there are great individual differences in face recognition, ranging from individuals who suffer from developmental prosopagnosia (DPs) and perform poorly on standard face recognition tasks (Bate, Bennetts, Tree, Adams, & Murray, 2019; Duchaine & Nakayama, 2006) to super recognizers (Bobak, Hancock, & Bate, 2016; Ramon, Bobak, & White, 2019; Russell, Duchaine, & Nakayama, 2009) who excel with face recognition in daily life and on lab-based tasks. Are these individuals, in the extreme ends of the face recognition ability spectrum, sensitive to the same facial information used by individuals within the normal range? A recent study that used the bubbles technique showed that face recognition depends on information from the eyebrow, eye, and mouth regions. Furthermore, the usage of this information was correlated with face recognition abilities, with super recognizers (SRs) showing increased reliance on these features, whereas DPs showed less reliance on this facial information (Tardif et al., 2019).

The bubbles technique reveals which facial regions are used for face recognition but is unable to tell which features are extracted from these regions. For example, whereas the bubbles technique indicates that the mouth region is informative for face identification, Abudarham and Yovel (2016) revealed that lip thickness, but not mouth size, was a critical feature for face identity. Similarly, the bubbles technique showed that the eyebrow region is critical for face identification, whereas Abudarham and Yovel (2016) showed that eyebrow thickness, but not eyebrow shape, was a critical feature for face identity. Thus, examination of the sensitivity of these features goes beyond methods that point to specific regions on the face while also providing a mechanistic account for view-invariant face processing. Thus, in the current study, we assessed whether SRs and DPs are also sensitive to changes in critical and non-critical features in face identification tasks.

To examine the importance of critical features in face identification, we used both a face recognition task (Study 1) and a face identity matching task (Study 2). A face recognition task enables us to assess the importance of critical features in the representation of familiar faces in memory. A face matching task enables us to assess their importance in the perceptual representation of face identity. In our previous study (Abudarham et al., 2019), we found that these critical features were important for both face recognition and identity matching, indicating that they are used both when matching the identity of faces perceptually as well as in the representation of familiar faces in memory. To generalize these findings to groups with extreme face recognition abilities, we examined sensitivity to critical features on a face recognition task in SRs (Study 1). Because DPs show poor face recognition abilities, we tested their sensitivity to critical features only on a perceptual identity matching task (Study 2) and compared to the performance of both controls and SRs.

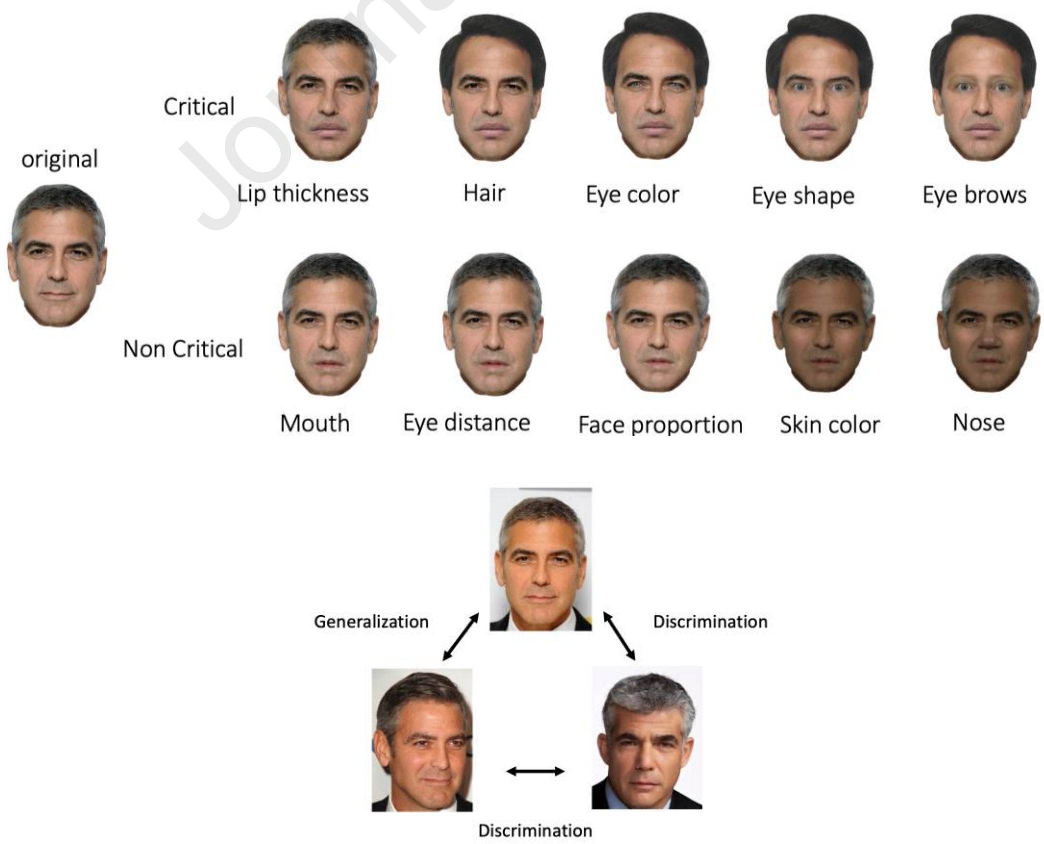


Figure 1: Top. An example of the face of George Clooney that was manipulated by changing five critical features (top row) or five non-critical features (bottom row). Bottom. Features that are critical for face recognition are view-invariant, whereas features that are not critical for face recognition vary across head views. For example, hair, eyebrow thickness, and lip thickness can be easily matched across different head views, whereas face proportion and eye distance are hard to match across head views, making them less useful for face identification (Abudarham et al., 2016)

Study 1 – Critical features for face recognition in SRs

To examine the role of critical and non-critical features in SRs, we used a face recognition task. Participants were asked to recognize celebrity faces that were modified by changing five critical or five non-critical features in a gradual manner (Figure 1, see also Abudarham et al., 2019). We then computed the proportion of faces that were recognized for each of the five feature manipulations. In a previous study, we found that changing four or five critical features radically impaired face recognition, whereas changing four or five non-critical features resulted in a more gradual and shallower drop in performance, in participants with face-recognition abilities in the normal range.

Methods

Participants

Thirty-three SRs participated in the experiment (mean age: 37 (22-50), 24 females). Sixteen participants performed the critical feature version of the face recognition task and 17 performed the non-critical feature version. Forty aged-matched control participants were recruited through Amazon Mechanical Turk to perform the same tasks. Four participants were excluded because they completed less than 10% of the trials. Thus, the SRs data were compared to 36 control participants (mean age: 37 (32-46), 9 females); 19 performed the critical feature task and 17 the non-critical feature tasks. The study was approved by the ethics committee of Tel Aviv University.

Selection criteria of SRs

The SRs were selected based on their scores on three face tasks: the Cambridge Face Memory Test - long form (CFMT+), the Pairs Matching Test (PMT), and the Models Memory Test (MMT). Their scores were 1.96SD above the mean on the CFMT+ and at least on one of the two other tasks. The cutoff score was 90 (87%) for the CFMT+, 73 (81%) for the MMT and 40 (83%) for the PMT (Bate et al., 2018). The average CFMT+ scores of the SRs that were assigned to the critical feature condition was similar (M=95 (93%)) to those who were assigned to the non-critical feature condition (M=93 (91%)). SRs that took part in the study were the people who responded to our invitation to participate in the study from a total pool of 66 who met the criteria for SRs indicated above.

Stimuli

We used the same face stimuli as Abudarham et al. (2019). The faces were frontal images of 10 American male celebrities, with neutral facial expression, no glasses or facial hair, in uniform lighting and good resolution. The images of each identity were modified in two different ways: by either changing five critical features or changing five non-critical features. The features were modified by replacing each feature with a feature copied from a different identity (different features were copied from different identities). Critical features were modified in the following order: lip thickness, hair, eye color, eye shape and eyebrow thickness (see Figure 1, top row). Non-critical features were replaced in the following order: mouth size, eye distance, face proportion, skin color, nose size (see Figure 1, bottom row). For more details about the feature substitution technique see Abudarham et al. (2019) and Abudarham & Yovel (2016).

Procedure

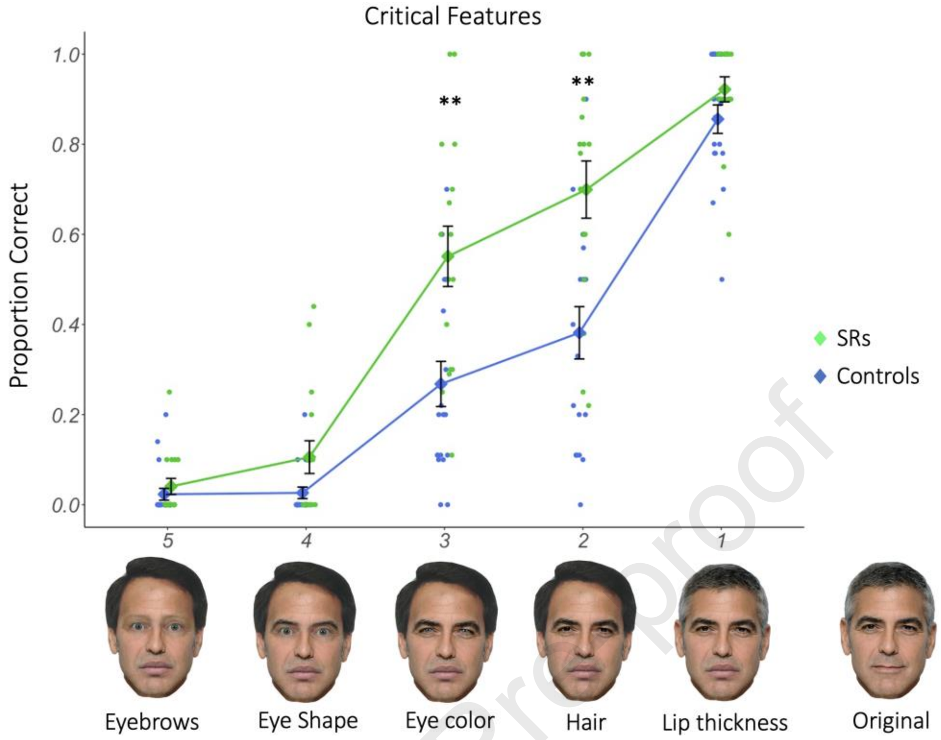
Because the same identities were used for the critical and non-critical changes, a between-subject design was used. Participants were presented with faces that were modified with either critical or non-critical features. The images were presented one at a time, in the following order: The first 10 celebrity faces with five feature changes were presented one after the other in a random order. Then the same identities with four feature changes were presented in a random order, and so on until the final set of 10 original, unchanged faces were presented. Thus, a total of 60 faces were presented one after the other. Participants were asked to type the name of the person on the screen if they recognized him, or any information that they may know about him if they could not recall the celebrity's name. If they could not recognize the face, they selected the answer "I cannot recognize this face". After participants made their response, they pressed a key to initiate the presentation of the next face. The faces were presented on the screen until the participants pressed the key to move on to the next face.

Data Analysis

For each participant, we calculated the proportion of faces correctly recognized, as a function of the number of feature changes for critical and non-critical features. Correct or incorrect recognition was assessed manually based on the subjects' textual response. For example, for the face of Mark Zuckerberg, the response "Mark Zuckerberg" as well as the response "The Facebook guy" were both accepted as correct responses. We averaged the recognition rate separately for SRs and for controls. Only trials in which participants recognized the celebrity faces in their original forms were included. The scores for the original faces were not included in the analysis because their variance was 0.

Results

Figure 2 shows the average proportion of faces that were recognized for each number of feature changes for critical and non-critical feature changes in SRs and Controls. A mixed 2x2x5 ANOVA with Group (SR, Controls), Feature Type (Critical, Non Critical) as between-subject factors and number of feature changes (one to five) as a within-subject factor revealed a main effect of Group ($F(1,65) = 23.5$, $p < .001$, $\eta^2_p = .27$), a main effect of Feature Type ($F(1,65) = 222.6$, $p < .001$, $\eta^2_p = .77$) and a significant interaction of Group, Feature Type, and Number of feature changes $F(4,260) = 17.5$, $p < .001$, $\eta^2_p = .21$). We examined the interaction between Group x Number of feature changes separately for critical (Figure 2, Top) and non-critical features (Figure 2, Bottom). The interaction was significant for Critical features $F(4,132) = 8.47$, $p < .001$, $\eta^2_p = .20$). Post-hoc comparisons (Bonferroni corrected) revealed significantly better performance for SRs than controls when there were two (($t(33) = 5.33$, $p < .001$, Cohen's $d = 1.25$) and three feature changes ($t(33) = 4.76$, $p < .001$, Cohen's $d = 1.17$). Because performance for the four and five critical feature changes was at floor level, we conducted another analysis of these data including only 1, 2, and 3 feature changes. Results were similar with a main effect of Group ($F(1,33) = 13.15$, $p < .001$, $\eta^2_p = .28$) and an interaction between Group and number of feature changes ($F(2,66) = 7.95$, $p < .001$, $\eta^2_p = .19$), indicating better recognition for SRs than controls. The interaction between Group x Number of feature changes was significant also for Non-Critical features $F(4,128) = 13.9$, $p < .001$, $\eta^2_p = .30$). Post-hoc comparisons (Bonferroni corrected) revealed significant better performance for SRs then controls for four ($t(32) = 3.97$, $p < .01$, Cohen's $d = 1.32$) and five feature changes ($t(32) = 6.37$, $p < .001$, Cohen's $d = 1.71$). None of the other Group comparisons was significant. Finally, a relatively weaker interaction effect of Group x Feature type $F(4,260) = 3.74$, $p < .01$, $\eta^2_p = .05$) indicates that the gap in performance between critical and non-critical features was somewhat larger in SRs than controls. Nevertheless, both groups showed much lower recognition rates for famous faces following critical than non-critical feature changes.


Critical Features
Proportion Correct
1.0
0.8
0.6
0.4
0.2
0.0
**
**
5
4
3
2
1
SRs
Controls
Eyebrows
Eye Shape
Eye color
Hair
Lip thickness
Original

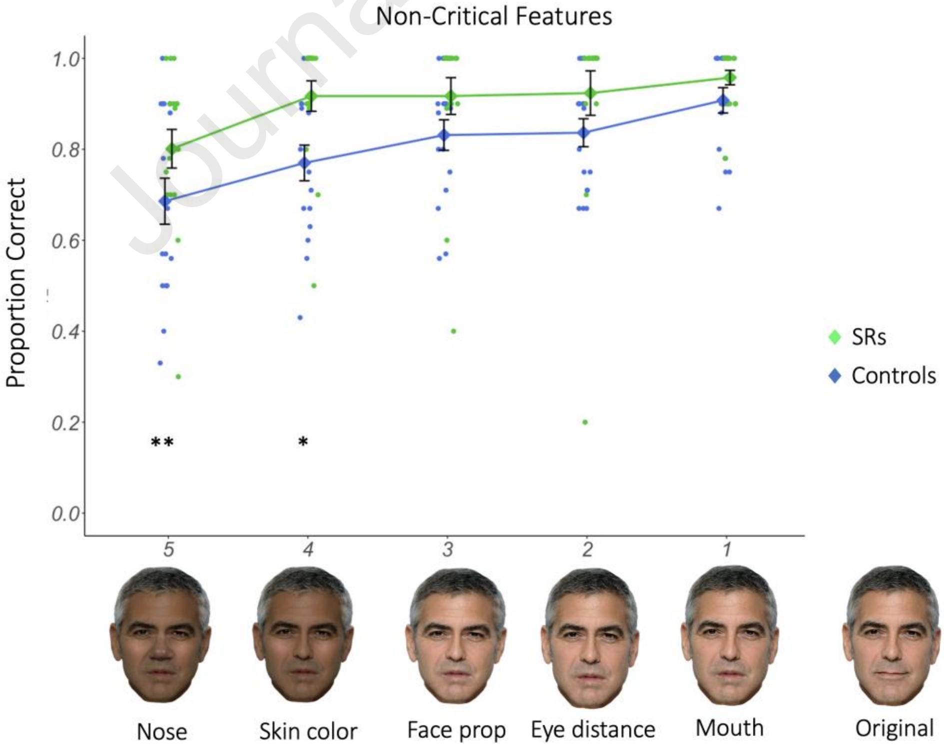

Non-Critical Features
Proportion Correct
1.0
0.8
0.6
0.4
0.2
0.0
**
*
5
4
3
2
1
SRs
Controls
Nose
Skin color
Face prop
Eye distance
Mouth
Original

Figure 2: Proportion correct recognition of famous faces that were manipulated by either changing their critical features (Top) or their non-critical features (Bottom) in super recognizers (SRs) and controls. The values on the x-axis indicate the number of feature changes such that each feature change was added to the previous one. Faces are shown from left to right in the order of presentation in the experiment starting with five feature changes to the original face. Error bars indicate the standard error of the mean. * < ,01, ** < .001

Discussion

The current study examined SRs' sensitivity to critical and non-critical features. Overall, we found that similar to controls, when changing four or five critical facial features, SRs were unable to recognize well-known celebrity faces that they could easily recognize in their original, non-manipulated form. Thus, the critical features that are used for face recognition in individuals with normal face recognition abilities are also used by SRs. However, there were two differences between SRs and controls that are noteworthy. First, changing two or three critical features had a smaller effect on recognition in SRs relative to controls, indicating that they can better rely on the remaining critical features (Figure 2, Top). Second, SRs showed much better recognition for faces that were manipulated by changing non-critical features (Figure 2, Bottom). Whereas recognition of celebrity faces is lower when four or five non-critical features are changed in controls, it remains high in SRs. Importantly, however, changing four and five critical features had similar effects on SRs and controls, which indicates the importance of these features for face recognition.

The current study presented the feature changes in a particular order. Thus, we could not dissociate the effect of specific features for face recognition from the order in which they were changed. The effect of the order of feature changes was assessed by Abudarham et al (2019) who found similar findings when the feature changes was performed in a reversed order. We therefore believe that the results reported here can be generalized to other sequences of feature changes. Nevertheless, given that we could not run another sample of SRs with a different order of feature changes, we cannot rule out that the effect we revealed results from the specific features (eye shape and eyebrow thickness) rather the number of features that were changed. It is noteworthy that by replacing features in different orders in our previous study (Abudarham et al., 2019), we did reveal that the feature replacement that had the most significant effect on recognition was the hair. When hair was replaced last, recognition declined more slowly than when the hair was replaced earlier in the sequence. These findings highlight the important role of the hair even in recognition of familiar people and may be inconsistent with the findings that familiar face recognition rely more on internal than external features.

Interestingly, here we find that SRs were less affected by hair changes and can still recognize famous faces better than controls when they are presented with a different hair.

Our previous studies showed that these critical features are invariant across different head views and are therefore useful for view-invariant face recognition. In contrast, non-critical features vary across different appearances. For example, eye distance and face proportion vary across head view, and skin color varies in different illuminations, and are therefore not useful for face identification across different appearances (see Figure 1 Bottom). The lower sensitivity of SRs' to changes in non-critical features may enable them to generalize better across different appearances of the same identity. This ability is evident in previous demonstrations that SRs excel in tests requiring recognition of celebrities at ages before they were well known (e.g., the Before-They-Were-Famous Test) (Russell et al., 2009). In that test, SRs can recognize famous people in images taken before they became famous (images that they have not seen before) - a task that is very challenging for individuals with typical face recognition abilities. This was nicely described by the SR, Jennifer Jarett, in an interview on the CBS program 60 Minutes explaining how she was able to recognize the late American journalist Mike Wallace as a 6-year-old: "as people age…the aging process somehow in my brain seems sort of very superficial… as if someone gets a haircut you can still recognize them, still the same face to me". Although we cannot tell whether the features that we manipulated in the current study can account for recognition of faces across different ages, they are consistent with the ability of SRs to recognize familiar faces across radical changes in appearance.

On the other side of the distribution of face recognition abilities are DPs. DPs have difficulty recognizing familiar faces, and many DPs have difficulties perceptually matching unfamiliar faces, particularly if they are shown across different appearances or in difficult visual conditions (Bate, Bennetts, Gregory, et al., 2019; Bate, Bennetts, Tree, et al., 2019; Biotti, Gray, & Cook, 2019; Dalrymple, Garrido, & Duchaine, 2014). We therefore examined whether DPs would also show different sensitivity to critical and non-critical features relative to SRs and controls.

Study 2 – Critical features for face perception in SRs and DPs

In study 2, we tested individuals who suffer from face recognition difficulties, and examined their sensitivity to critical and non-critical features. Because DPs show poor performance on familiar face recognition tasks, we tested their sensitivity to critical and non-critical features with an identity matching task. Our previous studies showed that, consistent with results of the face recognition task (Study 1), pairs of faces that differ in critical features are rated as different identity faces, whereas faces that differ in non-critical features are rated as same identity. Furthermore, because DPs are known to heavily rely on the hair for face recognition (Adams, Hills, Bennetts, & Bate, 2019), we tested their sensitivity to the other four critical features. We therefore presented faces with four critical feature changes not including the hair and then compared that performance to performance with pairs that differ in four non-critical features (Figure 3). Finally, to directly compare between DPs and SRs, we also tested another group of SRs, who did not participate in Study 1, on the same face matching task and compared both groups to an age-matched control group of participants.

Methods

Participants

Nineteen DPs participated in the study (average age: 43 (26-65), 13 females). All reported severe difficulties with face recognition in daily life. Their face recognition was assessed with three face memory tests: the CFMT (Duchaine & Nakayama, 2006), an old-new face recognition test (Duchaine & Nakayama, 2005), and a famous face test (Jiahui, Yang, & Duchaine, 2018). DPs were selected based on performance that was two standard deviations or more below the mean of control participants on at least two of the three face recognition tasks. This group of DPs also performed the Cambridge Face Perception Test and 11 of them showed poor performance (>2SD below the mean of controls). DPs were selected from the faceblind.org database (n ≈ 14,000) based on the reported severity of their difficulties with faces.

Twenty SRs who did not participate in Study 1 performed the same identity matching task (average age: 42 (23-60), 12 females). They all scored 1.96 SDs above the mean on the CFMT+ and at least one of the two additional tests, PMT and MMT.

Twenty-eight age-matched control participants (average age: 43 (37-48), 11 females) were recruited through Mechanical Turk to perform the same task. One participant who pressed the same key for all the trials throughout the experiment was excluded from the analysis. The study was approved by the ethics committee of Tel Aviv University.

Stimuli

We used the same 10 celebrity faces used in Study 1. Because DPs are known to rely on the hair for face recognition (Adams et al., 2019; Duchaine, 2011), we tested their sensitivity to the four critical features that did not include the hair (namely eyebrow-thickness, eye-shape, eye-color, lip-thickness) and compared them to changes in four non-critical features (face-proportion, eye-distance, mouth-size and skin-color) (see Figure 3, bottom).

Procedure

The perceptual matching task included four types of face pairs: *Same*, which presented two different images of the same identity, *Non-Critical*, in which non-critical features were modified; *Critical*, in which critical features were modified, and *Different*, in which two faces of different identities were presented (Figure 3). The pairs of faces were presented on the screen side by side, and participants were asked to indicate whether the faces belong to the same identity or to different identities, on a scale from 1-6. '1' "definitely the same person"; '2' "same person"; '3' "possibly the same person"; '4' "possibly different people"; '5' "different people" and '6' "definitely different people". The two faces were presented on the screen until response, after which the next pair of faces were presented. The order of the pairs, as well as the right-left positions of the images within each pair, were randomized across participants.

In order to not repeat the same identity in the critical and non-critical feature conditions, we created two versions of the task. In one version five identities were modified by changing their critical features and the remaining five by changing their non-critical features, and vice versa on the other version. Thus, each version included ten same pairs, five non-critical feature pairs, five critical features pairs and ten different pairs. Analysis was performed across the two versions and therefore included 10 identities for each of the four conditions.

Results

A mixed ANOVA with Group (SR, DP, Controls) as a between-subject factor and Pair Type (Same, Non-Critical, Critical, Different) as repeated measures revealed a main effect of Pair Type, $F(3,189) = 373.02$, $p < .001$, $\eta^2_p = 0.86$) and a main effect of Group, $F(2,63) = 4.46$, $p < .01$, $\eta^2_p = 0.12$). As seen in Figure 3, the overall pattern of results was similar in all three groups, with lower scores for Same pairs, then increasing for Non-Critical Feature pairs, Critical Feature pairs, and Different pairs. The main effect of Group indicates overall higher scores in DPs than SRs ($t(37) = 2.98$, $p < .015$, corrected for three comparisons, Cohen's $d = 0.37$). The control group did not differ significantly from SRs or DPs.

We also found a significant interaction between the Group and Pair Type, $F(6,189) = 2.78$, $p = .013$, $\eta^2_p = 0.08$). Post hoc comparisons (Bonferroni corrected for multiple comparisons) indicate that this interaction is due to lower confidence in rating same-identity pairs as same in DPs than SRs ($p = .004$) but no other significant group differences for any of the other comparisons.

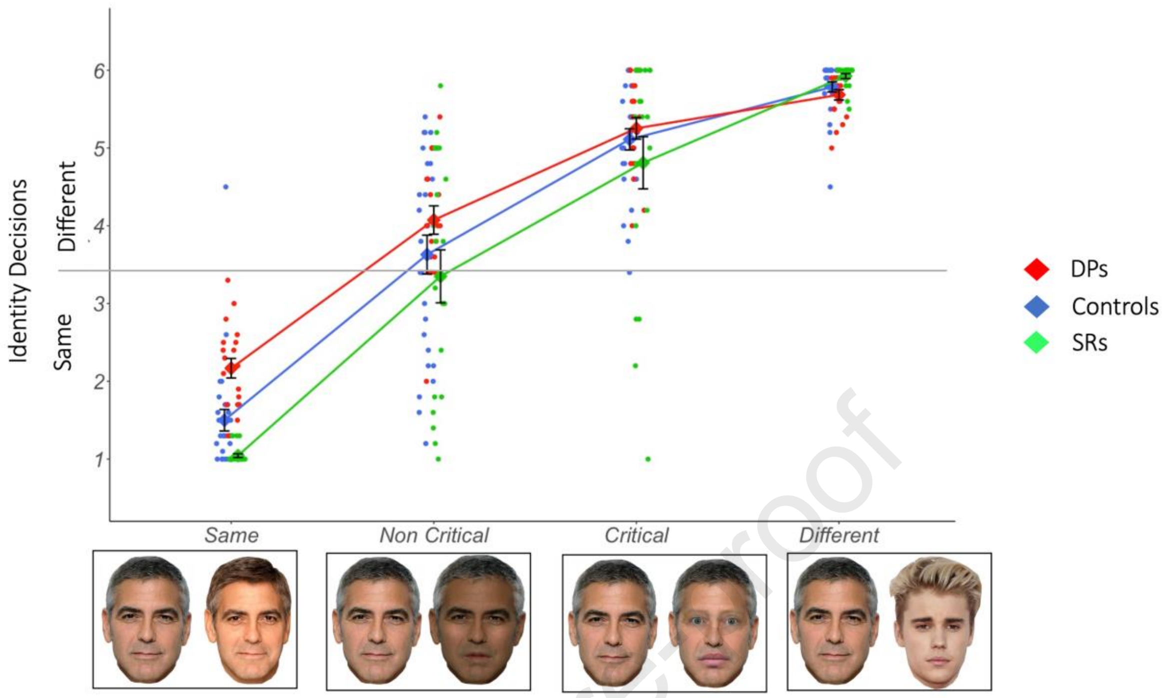


Figure 3: Identity decision ratings to the four face pairs presented simultaneously in SRs, DPs and controls. The horizontal line indicates the border between same and different identity decisions. Error bars indicate the standard error of the mean.

Discussion

Performance of SRs and DPs on the identity matching task was generally similar to controls. Pairs of faces that differed in critical features were rated as more likely to be different identities than faces that differed in non-critical features. The only difference between groups was that DPs were less confident than SRs that the same identity faces belong to the same identity. This may be related to the tendency of DPs to classify faces as different, which was reported in a previous study (White et al., 2017).

Notably, the critical features did not include the hair, which has a major effect on determining the identity of a face both in controls (Abudarham et al., 2019) and DPs (Adams et al., 2019; Duchaine, 2011; Murray, Hills, Bennetts, & Bate, 2018). In our previous study we showed that when changes in four critical features do include the hair, control participants showed poor recognition of these faces in a face recognition task (Abudarham et al., 2019). The current findings show that even when hair is not changed, changes in critical features significantly change the identity of a face.

Performance of SRs on the identity matching task was similar to controls. When presented with the original and the faces that differ in four non-critical features side by side in an identity matching task, SRs indicated that they belonged to different identities. These results are somewhat different from SRs performance on face recognition in Study 1, where they were able to recognize faces that differed in four non-critical features significantly better than controls (Figure 2, bottom). This implies that the representation that SRs generate for familiar faces in memory is more tolerant to feature manipulations.

General Discussion

The current study asked whether SRs and DPs use the same features to determine the identity of a face as people with normal face recognition do. Overall, our findings show that similar to control participants, SRs and DPs are more sensitive to critical than non-critical features. These results indicate that individuals across the spectrum of face recognition ability rely on the same facial information for face identification. Nevertheless, there are also some differences between the groups that emerge from these tasks. SRs were tested using a familiar face recognition task that enabled us to assess their representation of familiar faces in memory. Overall, SRs were better than controls in recognition of familiar faces that differ in two or three critical features (Figure 2, top), and were also better than controls in recognizing familiar faces following changes in four or five non-critical features (Figure 2, bottom). This performance may be consistent with their exceptional ability to generalize across different appearances of the same identity (Russell et al., 2009). Interestingly, this better tolerance to feature manipulation in face recognition of SRs, was not evident in the face matching task, where performance of SRs was similar to controls. These different outcomes may indicate different decision criteria in SRs for recognition vs. perceptual matching.

Unlike SRs who completed both a face recognition and an identity matching task, DPs were tested only with the identity matching task because they struggle with familiar face recognition. These two types of tasks have shown comparable findings with respect to differences between critical and non-critical features in normal participants (Abudarham et al., 2019). In the current study, we found that similar to controls, DPs judge faces that differ in critical features as more different than faces that differ in non-critical features. The only difference that was found between DPs and SRs was DPs' lower confidence in rating same identity faces as the same identity. These findings suggest that DPs are more likely to misidentify a familiar person when they see them in a different appearance rather than falsely identify two different identities as the same person. The difficulty of "telling faces together" more than "telling faces apart" has been reported also in normal participants (Andrews, Jenkins, Cursiter, & Burton, 2015; Jenkins, White, Van Montfort, & Burton, 2011). In these studies, participants are asked to sort different images of the same identities of two or more different identities. Results typically show that participants make generalization errors but not discrimination errors. Our findings show that this bias is stronger in DPs. Similar findings were also reported by White and colleagues (2017) who showed that DPs tend to make more errors relative to controls on same identity than different identity face pairs (White, Rivolta, Burton, Al-Janabi, & Palermo, 2017).

Taken together these findings indicate that the ability to tune to critical features and overlook non-critical features is important for face identification. The relationship between sensitivity to critical features and performance on a face perception task has been recently further supported in a study that used a face recognition deep convolutional neural network (DCNN) to model human face recognition (Abudarham & Yovel, 2020). Using the same stimuli used in the current study, this study examined the sensitivity of DCNNs to critical features across its different layers. Results showed that sensitivity to critical over non-critical features was found in higher layers of a face recognition DCNN. In contrast, there was no difference between critical features than non-critical features in low-layers of DCNNs, indicating that these feature manipulations are apparent for high-level but not low-level representations. Moreover, performance of the network on a face matching task was highly correlated with its sensitivity to critical features across its different layers.

The different performance of DPs in the identity matching task is consistent with studies that show that DPs are poor not only in face recognition but also in face perception tasks (Bate, Bennetts, Tree, et al., 2019; Biotti et al., 2019; Dalrymple et al., 2014; Ulrich et al., 2017, but see Bate, Bennetts, Gregory, et al., 2019). The poor performance on perceptual matching tasks reported in previous studies, such as the CFPT may reflect difficulty in extracting these features in the face stimuli that were typically used in these tasks. The faces used in the current study were high-quality color images, and it is possible that the added noise to the grayscale morphed face stimuli in the CFPT is more detrimental to DPs than Controls. Future studies that systematically manipulate the visibility of the critical features may be able to link between current findings with the more challenging tasks that were used in previous perceptual matching face tasks.

In summary, our results show that DPs, SRs, and Controls are all tuned to features that are critical for face identification. This has been shown on both face recognition and perceptual matching tasks indicating that these features are fundamental to the identity of a face.

References

Abudarham, N.;Yovel, G. (2020). Face recognition depends on specialized mechanisms tuned to view-invariant facial features: Insights from deep neural networks optimized for face or object recognition. *BioRxiv*.

Abudarham, N., & Yovel, G. (2016). Reverse engineering the face space: Discovering the critical features for face identification. *Journal of Vision, 16*(3). https://doi.org/10.1167/16.3.40

Abudarham, Naphtali, Shkiller, L., & Yovel, G. (2019). Critical features for face recognition. *Cognition*, *182*, 73–83. https://doi.org/10.1016/j.cognition.2018.09.002

Adams, A., Hills, P. J., Bennetts, R. J., & Bate, S. (2019). Coping strategies for developmental prosopagnosia. *Neuropsychological Rehabilitation*. https://doi.org/10.1080/09602011.2019.1623824

Andrews, S., Jenkins, R., Cursiter, H., & Burton, A. M. (2015). Telling faces together: Learning new faces through exposure to multiple instances. *Quarterly Journal of Experimental Psychology*. https://doi.org/10.1080/17470218.2014.1003949

Bate, S., Bennetts, R. J., Gregory, N., Tree, J. J., Murray, E., Adams, A., … Banissy, M. J. (2019). Objective patterns of face recognition deficits in 165 adults with self-reported developmental prosopagnosia. *Brain Sciences*. https://doi.org/10.3390/brainsci9060133

Bate, S., Bennetts, R. J., Tree, J. J., Adams, A., & Murray, E. (2019). The domain-specificity of face matching impairments in 40 cases of developmental prosopagnosia. *Cognition*. https://doi.org/10.1016/j.cognition.2019.104031

Bate, S., Frowd, C., Bennetts, R., Hasshim, N., Murray, E., Bobak, A. K., … Richards, S. (2018). Applied screening tests for the detection of superior face recognition. *Cognitive Research: Principles and Implications*. https://doi.org/10.1186/s41235-018-0116-5

Biotti, F., Gray, K. L. H., & Cook, R. (2019). Is developmental prosopagnosia best characterised as an apperceptive or mnemonic condition? *Neuropsychologia*. https://doi.org/10.1016/j.neuropsychologia.2018.11.014

Bobak, A. K., Hancock, P. J. B., & Bate, S. (2016). Super-recognisers in Action: Evidence from Face-matching and Face Memory Tasks. *Applied Cognitive Psychology*. https://doi.org/10.1002/acp.3170

Butler, S., Blais, C., Gosselin, F., Bub, D., & Fiset, D. (2010). Recognizing famous people. *Attention, Perception, and Psychophysics*. https://doi.org/10.3758/APP.72.6.1444

Chang, L., & Tsao, D. Y. (2017). The Code for Facial Identity in the Primate Brain. *Cell*. https://doi.org/10.1016/j.cell.2017.05.011

Dalrymple, K. A., Garrido, L., & Duchaine, B. (2014). Dissociation between face perception and face memory in adults, but not children, with developmental prosopagnosia. *Developmental Cognitive Neuroscience*. https://doi.org/10.1016/j.dcn.2014.07.003

Duchaine. (2011). Developmental prosopagnosia. In H. Calder, Rhodes & Johnson (Eds.), *Handbook of Face Perception*. Oxford University Press.

Duchaine, B. C., & Nakayama, K. (2006). Developmental prosopagnosia: a window to content-specific face processing. *Current Opinion in Neurobiology*. https://doi.org/10.1016/j.conb.2006.03.003

Duchaine, B., & Nakayama, K. (2005). Dissociations of face and object recognition in developmental prosopagnosia. *Journal of Cognitive Neuroscience*. https://doi.org/10.1162/0898929053124857

Duchaine, B., & Nakayama, K. (2006). The Cambridge Face Memory Test: Results for neurologically intact individuals and an investigation of its validity using inverted face stimuli and prosopagnosic participants. *Neuropsychologia*. https://doi.org/10.1016/j.neuropsychologia.2005.07.001

Ellis, H. D., Shepherd, J. W., & Davies, G. M. (1979). Identification of familiar and unfamiliar faces from internal and external features: Some implications for theories of face recognition. *Perception*, *8*(4), 431–439. https://doi.org/10.1068/p080431

Freiwald, W. A., & Tsao, D. Y. (2010). Functional compartmentalization and viewpoint generalization within the macaque face-processing system. *Science*. https://doi.org/10.1126/science.1194908

Gosselin, F., & Schyns, P. G. (2001). Bubbles: A technique to reveal the use of information in recognition tasks. *Vision Research*. https://doi.org/10.1016/S0042-6989(01)00097-9

Jenkins, R., White, D., Van Montfort, X., & Mike Burton, A. (2011). Variability in photos of the same face. *Cognition*. https://doi.org/10.1016/j.cognition.2011.08.001

Jiahui, G., Yang, H., & Duchaine, B. (2018). Developmental prosopagnosics have widespread selectivity reductions across category-selective visual cortex. *Proceedings of the National Academy of Sciences of the United States of America*. https://doi.org/10.1073/pnas.1802246115

Kramer, R. S. S., Towler, A., Reynolds, M. G., & Burton, A. M. (2018). *Familiarity and Within- Person Facial Variability : The Importance of the Internal and External Features*. https://doi.org/10.1177/0301006617725242

LE GRAND, R., MONDLOCH, C., MAURER, D., & BRENT, H. (2001). Early visual experience and face processing. *Nature (London)*.

Maurer, D., Le Grand, R., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences*. https://doi.org/10.1016/S1364-6613(02)01903-4

Mckone, E., & Yovel, G. (2009). Why does picture-plane inversion sometimes dissociate perception of features and spacing in faces, and sometimes not? toward a new theory of holistic processing. *Psychonomic Bulletin and Review*, *16*(5). https://doi.org/10.3758/PBR.16.5.778

Mondloch, C. J., Le Grand, R., & Maurer, D. (2002). Configurai face processing develops more slowly than featural face processing. *Perception*. https://doi.org/10.1068/p3339

Murray, E., Hills, P. J., Bennetts, R. J., & Bate, S. (2018). Identifying Hallmark Symptoms of Developmental Prosopagnosia for Non-Experts. *Scientific Reports*. https://doi.org/10.1038/s41598-018-20089-7

Ramon, M., Bobak, A. K., & White, D. (2019). Super-recognizers: From the lab to the world and back again. *British Journal of Psychology*, *110*(3), 461–479. https://doi.org/10.1111/bjop.12368

Riesenhuber, M., & Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nature Neuroscience*. https://doi.org/10.1038/14819

Russell, R., Duchaine, B., & Nakayama, K. (2009). Super-recognizers: People with extraordinary face recognition ability. *Psychonomic Bulletin and Review*. https://doi.org/10.3758/PBR.16.2.252

Sergent, J. (1984). An investigation into component and configural processes underlying face perception. *British Journal of Psychology*. https://doi.org/10.1111/j.2044-8295.1984.tb01895.x

Tanaka, J. W., & Sengco, J. A. (1997). Features and their configuration in face recognition. *Memory and Cognition*. https://doi.org/10.3758/BF03211301

Tardif, J., Morin Duchesne, X., Cohan, S., Royer, J., Blais, C., Fiset, D., … Gosselin, F. (2019). Use of Face Information Varies Systematically From Developmental Prosopagnosics to Super-Recognizers. *Psychological Science*, *30*(2), 300–308. https://doi.org/10.1177/0956797618811338

Ulrich, P. I. N., Wilkinson, D. T., Ferguson, H. J., Smith, L. J., Bindemann, M., Johnston, R. A., & Schmalzl, L. (2017). Perceptual and memorial contributions to developmental prosopagnosia. *Quarterly Journal of Experimental Psychology*. https://doi.org/10.1080/17470218.2016.1177101

White, D., Rivolta, D., Burton, A. M., Al-Janabi, S., & Palermo, R. (2017). Face matching impairment in developmental prosopagnosia. *Quarterly Journal of Experimental Psychology*, *70*(2), 287–297. https://doi.org/10.1080/17470218.2016.1173076

Yovel, G., & Duchaine, B. (2006). Specialized face perception mechanisms extract both part and spacing information: Evidence from developmental prosopagnosia. *Journal of Cognitive Neuroscience*, *18*(4). https://doi.org/10.1162/jocn.2006.18.4.580

Yovel, Galit, & Kanwisher, N. (2004). Face perception: Domain specific, not process specific. *Neuron*, *44*(5). https://doi.org/10.1016/j.neuron.2004.11.018