On the Relation between Face and Object Recognition in Developmental Prosopagnosia: No Dissociation but a Systematic Association

Christian Gerlach[1]*, Solja K. Klargaard[1], Randi Starrfelt[2]

1 Department of Psychology, University of Southern Denmark, Odense, Denmark, 2 Department of Psychology, University of Copenhagen, Copenhagen, Denmark

* cgerlach@health.sdu.dk

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Competing Interests: The authors have declared that no competing interests exist.

Abstract

There is an ongoing debate about whether face recognition and object recognition constitute separate domains. Clarification of this issue can have important theoretical implications as face recognition is often used as a prime example of domain-specificity in mind and brain. An important source of input to this debate comes from studies of individuals with developmental prosopagnosia, suggesting that face recognition can be selectively impaired. We put the selectivity hypothesis to test by assessing the performance of 10 individuals with developmental prosopagnosia on demanding tests of visual object processing involving both regular and degraded drawings. None of the individuals exhibited a clear dissociation between face and object recognition, and as a group they were significantly more affected by degradation of objects than control participants. Importantly, we also find positive correlations between the severity of the face recognition impairment and the degree of impaired performance with degraded objects. This suggests that the face and object deficits are systematically related rather than coincidental. We conclude that at present, there is no strong evidence in the literature on developmental prosopagnosia supporting domain-specific accounts of face recognition.

Introduction

It is debated whether face recognition and object recognition constitute separate cognitive domains [1]. Clarification of this issue can have important theoretical implications as face recognition is often used as a prime example of domain-specificity in mind and brain [2]. Domain-specificity entails the proposition that specialized cognitive functions (and brain areas) can and have evolved to handle very specific types of information; e.g., processing of faces and faces only. As an alternative to domain-specific accounts are theories assuming the existence of more multi- or general-purpose mechanisms which can handle information from several domains; e.g., faces but also other types of objects. According to domain-specific

accounts we should expect to find relatively clear-cut dissociations between domains, at least in some cases. In comparison, multi-purpose accounts would expect that information processing from different but related domains will differ in degree rather than kind.

The question of whether face and object recognition can dissociate is distinct from the question of what could underlie such a dissociation [3]; and there are several possibilities. To mention two, a dissociation could reflect that evolutionary constraints have led to selective tuning to significant stimulus classes such as faces [4], or that holistic processing might be important for face but not object recognition [5]. Likewise, there are several ways in which graded but non-selective differences in face and object processing can be explained in multi-purpose accounts. As examples, faces might constitute a stimulus class with which we have more expertise than other stimulus classes [6], or face recognition might generally require more fine grained perceptual discrimination than object recognition [7]. The purpose of the present study is to examine the general question of whether face and object recognition can dissociate; not how dissociations or lack hereof can best be accounted for.

Studies of impaired face recognition (prosopagnosia) have contributed significantly to the debate concerning selectivity of face processing. Until recently, most of these studies concerned patients with deficits in face processing following brain injury; *acquired prosopagnosia* (e.g., Barton, 2003), but now an increasing number of studies report data from individuals who have experienced face recognition problems their whole life, and where there is no (known) brain damage; a disorder known as *developmental* or *congenital prosopagnosia* [8]. If face processing is the product of domain-specific operations–which are usually considered innate [9]–we would expect that these operations can be the target of selective maldevelopment, as has been argued by for example Duchaine, Yovel, Butterworth and Nakayama [10]. If not, we would expect that object recognition is also affected in developmental prosopagnosia (DP), as has been argued by for example Behrmann, Avidan, Marotta, and Kimchi [11].

Even though several studies of DP have been published since the seminal report by McConachie in 1976 [12], visual object processing has typically not been assessed in great detail in these studies. Moreover, most studies have measured performance in terms of accuracy only. This may be problematic because accuracy within the normal range can be achieved by means of alternative strategies that may be extraordinary time consuming and hence not normal; a speed-accuracy trade-off (see e.g., [13, 14]).

Another issue concerns the way dissociations are defined. Typically, a dissociation is claimed if a person performs abnormally on task *X* but within the normal range on task *Y*, with 'normality' being anchored by whether performance fall below or beyond 2 SDs of the control mean. There are two problems with this approach. First, it seems to be based on the (implicit) assumption that failure to reject the null-hypothesis regarding performance on task *Y* is proof of normality, which is a dubious assumption [15]. Secondly, it fails to take into consideration the difference in performance between task *X* and task *Y*. In the extreme case, one could claim a dissociation if a person's score amounted to -2.01 SD on task *X* and -1.99 SD on task *Y*; a trivial difference of .02 SD. These problems are not specific to studies of DP but have been discussed in other contexts as well (see e.g., [16]). To avoid them Crawford et al. [15] have suggested that a performance pattern must fulfil two criteria in order to count as a dissociation: (i) the person's performance on task *X* must differ significantly from that of the normal population, and (ii) the *difference* in performance of that person on tasks *X* and *Y* must differ significantly from the difference-scores of the normal population on tasks *X* and *Y*. To our knowledge such a "qualified" dissociation has not yet been reported in DP.

In summary; while there have been many reports of dissociations between face and object recognition performance in individuals with DP, no study has yet demonstrated such a

dissociation in terms of both accuracy and reaction time while also adopting the stringent criteria for a dissociation advocated by Crawford et al. [15].

In the present study we tested 10 individuals with DP to examine if any of them would fulfil the criteria for a dissociation as described above, that is: (i) performing worse with faces than with objects in terms of both accuracy and reaction time, and (ii) satisfying the criteria for a dissociation as defined by Crawford et al. [15]. We examined this by means of three different types of visual object processing tasks. The first type of task was *object decision*, where participants must decide whether a displayed object represents a real object or a non(sense)-object. All non-objects were chimeric combinations of real objects, for example the front part of a horse and the back part of a dog. The use of chimeric non-objects makes the task more demanding in terms of perceptual differentiation than for example object naming tasks [14, 17]. To further increase the sensitivity of the task, we presented the stimuli as silhouettes and fragmented forms in addition to presenting them as regular line drawings. The use of degraded stimuli (e.g., fragmented forms) has previously proven successful in revealing even subtle deficits in visual object processing which may otherwise go unnoticed [18]. The second type of task involved within-class recognition of objects (cars). In this task, the participant must be able to keep a stimulus in memory for a short duration and compare it with an array of objects that, just like faces, are very similar to each other and the target. The third type of task required perceptual matching of two simultaneously presented stimuli; either two faces or two houses. Hence, as opposed to the two other types of task, this task does not rely on memory to any great extent. Stimulus-pairs could differ in either 1st order relations among the elements (e.g., the nose placed below the mouth in one face but at its normal position in the other), 2nd order relations among elements (e.g., difference in the spacing between the eyes), or in their features (e.g., different shape of the nose); for a discussion of these dimensions in face processing see [19]. Furthermore, the number of differences between the stimulus-pairs varied parametrically from one to four differences along the three dimensions. Hence, with this design we could examine if the DPs were impaired in perceiving particular sorts of information (1st order, 2nd order and/or features), if their problems were modulated by visual similarity (degree of difference), and if potential difficulties were specific to faces.

If face recognition can be selectively impaired in DP, we expect our sample, or at least some of the individuals with DP, to perform within the range of controls on these tasks, and further, that the discrepancy in their performance with faces compared to objects be significantly larger than that observed in control participants (cf. [15]).

Assessment of Face Recognition Ability

Method & Participants

Ten individuals with developmental prosopagnosia (DP) and 20 typically developing controls were included in this study. The DPs and the control participants performed the same tasks, in the same order, and it was the same control participants that served in all tasks. All had normal (or corrected to normal) vision, no learning disability, and no known history of neurological damage or psychiatric illness. All participants provided written informed consent according to the Helsinki declaration. The Regional Committee for Health Research Ethics of Southern Denmark has assessed the project, and ruled that it did not need formal registration.

Participants with developmental prosopagnosia. Following some appearances in Danish media, where we have informed about developmental prosopagnosia, we have been contacted by a number of people complaining of face recognition problems. They all report difficulties recognizing friends, colleagues, and sometimes even close family members and themselves by their faces, and that these problems have been present throughout their life.

As a first screening for DP, we used the *Cambridge Face Memory Test* (CFMT) [20] and the *Cambridge Face Perception Test* (CFPT) [21], both kindly provided by Brad Duchaine and translated into Danish.

In the CFMT the participant is introduced to six target stimuli, and then tested with forced choice items consisting of three stimuli, one of which is the target. The test comprises a total of 72 trials distributed over 3 phases: (a) an intro-phase with 18 trials where the study stimulus and the target stimulus are identical, (b) a novel-phase with 30 trials where the target differs from the study stimulus in pose and/or lighting, and (c) a novel+noise phase with 24 trials where the target differs from the study stimulus in pose and/or lighting and where Gaussian noise is added to the target. The depended measure is number of correct trials. The maximum score is thus 72; chance-level is 24.

In the CFPT the participant has to arrange six facial images according to their similarity to a target face. The images were created by morphing six different individuals with the target face. The proportion of the morph coming from the target face is varied in each image (88%, 76%, 64%, 52%, 40%, and 28%). The test comprises 16 trials, half with upright and half with inverted faces. Scores for each item are computed by summing the deviations from the correct position for each face. Scores for the 8 trials are then added to determine the total number of respectively upright and inverted errors. Hence, the depended measure is a deviation-score; the higher the score the poorer the performance, with chance-level at each orientation being a deviation-score of 93.3. For the present purpose we report only the deviation-score of the upright faces.

Likely candidates for DP were initially identified by performance falling below 2 SDs on the CFMT or the CFPT (upright faces) compared to the age and gender adjusted norms provided by Bowles et al. [22]. However, the final inclusion criteria for DP were abnormal performance on the CFMT, and the first part of the Faces and Emotion Questionnaire (FEQ) [23] compared to the matched Danish control sample. The first part of the FEQ comprises 29 statements concerning everyday face recognition such as "I rarely confuse characters in TV programs", and "I usually recognize my friends in old photographs". The statements are rated by the proband according to how strongly the proband agrees or disagrees with them using a four point Likert-scale. 10 individuals satisfied the final inclusion criteria for DP.

The DPs did not receive remuneration for their participation in this study. For the DPs to be anonymous, and yet recognizable across publications, we have kept their project case-numbers in text and tables.

Control participants. Two controls were matched for each DP on age, gender and educational level; making the groups comparable in terms of age (DP M = 37.3, SD = 13, range = 16–57; Control M = 37.3, SD = 12, range = 16–56) and years of education (DP M = 15.5, SD = 2.1, range = 11–17; Control M = 15.2, SD = 1.8, range = 10–17). All controls performed within the normal range on the CFPT and the CFMT, evaluated by the Bowles et al. norms [15]. Controls received gift certificates of ~120 DKK (~20 USD) per hour for their participation.

In all comparisons of the performance of an individual with DP with that of the small control sample (single case statistics) we used the method developed by Crawford, Garthwaite, and Porter [24] (Bayesian test for a deficit; implemented in the program SingleBayes_ES). When age and education was found to correlate reasonably ($r > .3$) with performance in the control sample, we controlled for these variables by means of the method developed by Crawford, Garthwaite, and Ryan [25] (Bayesian test for a deficit with covariates; implemented in the program BTD_Cov). A score was considered abnormal if the one-tailed probability that the score could be an observation from the control population was less than .05.

As seen in Table 1, the performance of 6/10 DPs also fell significantly below the control mean on the CFPT. All DPs performed within the normal range (score of 32 or more) on The

Table 1. Background data and test performance of the developmental prosopagnosics.

Case	Age	Gender	Handedness	CFMT	CFPT	FEQ	CCMT
PP04	57	Male	Right	**37**	**86**	**71**	42
PP07	40	Female	Right	**41**	60	**66**	55
PP09	40	Female	Left	**43**	**70**	**52**	47
PP10	34	Female	Right	**33**	58	**62**	45
PP13	51	Male	Right	**35**	42	**64**	39
PP16	23	Female	Left	**39**	**64**	**54**	—
PP17	49	Female	Right	**35**	**88**	**56**	**33**
PP18	38	Female	Left	**30**	**78**	**69**	41
PP19	16	Male	Right	**33**	48	**53**	**40**
PP27	25	Male	Right	**42**	66	**59**	45
Control mean				59.1	41.3	22.4	51
Control SD				7.9	11.4	11.4	7.2

Age, gender and performance on the Cambridge Face Memory Test (CFMT), the Cambridge Face Perception Test (CFPT: upright faces), the Face Questionnaire (FEQ), and the Cambridge Car Memory Test (CCMT) for the 10 individuals with developmental prosopagnosia. The mean performance and SD for the controls' scores are also listed. Values in boldface designate performance deviating significantly from the mean of the matched control group as assessed by means of Crawford and colleagues methods [24, 25]. In the CFMT a low score indicates a deficit, while in the CFPT and in the FQ a high score indicates a deficit.

doi:10.1371/journal.pone.0165561.t001

Autism-Spectrum Quotient (AQ) questionnaire [26]; see Table 1 for an overview of age, gender, and test scores.

Assessment of Visual Object Recognition Ability

Object decision

DPs and controls were tested on three object decision tasks which varied in stimulus type (regular line drawings, silhouettes, and fragmented forms). The demand placed on perceptual differentiation in object decision can be controlled by manipulating the type of nonobjects used. If the nonobjects are novel, that is, completely unknown to the participants, the task is relatively easy, as the only judgement needed is one of familiarity. In the present case the nonobjects were partly familiar because they were chimeric nonobjects composed by exchanging single parts belonging to different real objects. As demonstrated in previous studies, it is harder to reject chimeric nonobjects as being real objects than to reject novel nonobjects as being real objects [27, 28]. This effect of task difficulty also affects the processing of the real objects presented: It is harder to recognize a real object as a real object when it is presented in the context of chimeric nonobjects than when it is presented in the context of novel nonobjects. This can be explained in the following way: Given that the nonobjects in the easy tasks are novel it might be sufficient to identify just a few recognizable parts of an object to judge it as a real object. This strategy will not suffice in the difficult task because the nonobjects are composed of parts of real objects. Accordingly, when discriminating between real objects and chimeric nonobjects, it is necessary to keep on processing until a particular representation in visual long term-memory wins the competition and a complete match is found, i.e. the object is recognized. Otherwise, one will risk judging a nonobject as a real object [17].

Design. The DPs and the 20 control participants performed all three object decision tasks in the same order (fragmented forms, full line drawings, and silhouettes), except for PP16 who did not perform the version with silhouettes.

In each task, the participants were instructed to press '1', on a serial response box, if the picture represented a real object and '2', if it represented a nonobject. Participants were encouraged to respond as fast and as accurately as possible. Prior to each of the three tasks, the participants performed a practice version of the upcoming task. Stimuli used in these practice versions were not used in the actual experimental conditions.

The DPs and the control participants performed the fragmented version and the full line drawing version on the same day. The silhouette version was performed on another occasion separated by at least two weeks (mean = 211 days, range 20–274 days) for the DPs and one week for the control participants (mean = 109 days, range 8–199 days). Although this difference in interval was significant (t (1, 26) = 3.91, $p < .001$), there was no evidence that a shorter interval had a positive impact on performance with the silhouette version. Hence, we failed to find any significant negative correlation between interval and discrimination sensitivity (A), or a significant positive correlation between interval and RT for neither the DPs nor the Controls (all p's $> .4$ one-tailed). Accordingly, in the case that the group of DPs perform worse than the control group with the silhouette version, this performance difference cannot be attributed to a difference in test interval.

Stimuli.   160 pictures were presented in each task: 80 real objects and 80 chimeric nonobjects. The full line drawings of real objects were taken from the set of Snodgrass and Vanderwart [29]. The 80 chimeric drawings of nonobjects were selected mainly from the set made by Lloyd-Jones and Humphreys [30]. These nonobjects are line-drawings of closed figures constructed by exchanging single parts belonging to objects from the same category (see Fig 1).

The fragmented versions of the regular line drawings were made by imposing a mask as a semi-transparent layer on the regular line drawings. This mask consisted of blobs of different sizes and shapes. The regular line drawing and the mask were subsequently merged into a single layer yielding a fragmented version of the regular line drawing (see Fig 1). The same mask was used for the generation of all fragmented stimuli. The silhouette versions of the regular line drawings were made by replacing the colour of each pixel within the interiors of the regular
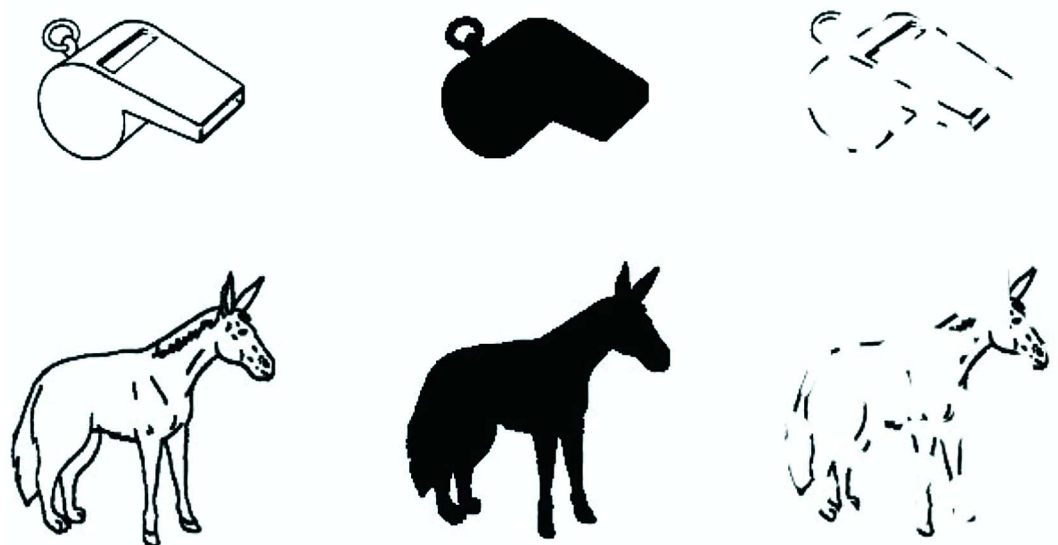

Fig 1. Examples of the stimuli used in the Object Decision Tasks. Upper panel: three versions (full drawing, silhouette, and fragmented) of a real object. Lower panel: three versions of a chimeric nonobject (half a wolf and half a donkey).

doi:10.1371/journal.pone.0165561.g001

line drawings with the colour black (see Fig 1). The order of pictures was randomized within each task.

Procedure. All stimuli were presented centrally in black on a white background on a PC-monitor and subtended 3-5° of visual angle. The stimuli were displayed until the participant made a response. The interval between response and presentation of the next object was 1 s. RTs were recorded by use of a serial response box.

Statistical analyses. We used A as a measure of discriminability. This is a bias-free measure of sensitivity similar to A' and A'' but based on a corrected formula by Zhang and Mueller [31]. The measure varies between 0.5 and 1.0 with higher scores indicating better discrimination between objects and nonobjects. Prior to analyses of A-scores we screened for individuals exhibiting extreme hit/false alarm rates (rates approaching 1), as such cases may yield unreliable estimates of discriminability. One DP (PP04) had such a performance on the object decision task with silhouettes where his hit rate was .99 while his false alarm rate was .94 (indicating that he pressed 1 for almost all trials). Even though this individual clearly had difficulties with this task–indeed his performance was on chance level (53% correct responses)–we decided to exclude him, and his matched controls, from the analyses of the results from this task as his A-score and RTs could not be clearly interpreted. None other was excluded on this account.

All analyses of RTs presented below are based on correct responses to real objects only, as the nonobjects served no other purpose in the present study than to ensure detailed shape processing of the real objects. Prior to RT-analysis the data were trimmed excluding trials from a particular participant if the RT of that trial fell above or below 2.5 SD of the mean of the participant's RT. For no individual in any of the three tasks did trimming result in discard of more than 5% of the trials. Comparisons between the DPs and the control participants were based on non-parametric tests as the variables departed from normality.

Results. The comparisons between the scores of the DPs and the control group were based on Mann–Whitney U Exact test. As can be seen in Table 2, the DPs performed within the normal range with regular drawings. In comparison, the DPs performed significantly worse than the control group with silhouettes in terms of A, and marginally so in terms of RT. In the condition with fragmented drawings they were marginally slower than the controls but did not differ significantly from the controls in terms of A.

Single case analysis. To examine the performance of the DPs individually, we compared the scores (A and mean RT) of each DP with the mean score of the control participants using the methods developed by Crawford and colleagues [24, 25]. This was done for all three object conditions. As can be seen in Table 3, where the results of these comparisons are summarized,

Table 2. Group data and statistics from the object decision tasks.

	DPs		Controls			
Condition	*Mdn*	*IRQ*	*Mdn*	*IRQ*	*z*	*p-value*
Regular drawings *A*:	.945	.025	.958	.026	−0.84	.42
Regular drawings RT (ms):	860	263	707	226	−1.23	.23
Silhouette drawings *A*:	.881	.068	.926	.05	−2.33	.02
Silhouettes drawings RT:	1189	590	853	277	−1.77	.08
Fragmented drawings *A*:	.784	.086	.800	.073	−0.4	.71
Fragmented drawings RT:	1135	431	922	384	−1.94	.06

Statistics (Median, interquartile range (IRQ), *z*-score, and *p*-value) associated with the comparisons of the DP group and the control group in the three object decision tasks. The comparisons are based on Mann–Whitney *U* Exact test.

doi:10.1371/journal.pone.0165561.t002

Table 3. Single case data from the object decision tasks.

Case	Regular drawings		Silhouettes		Fragmented	
	A	RT	*A*	RT	*A*	RT
PP04	.965	1025	------	------	.782	1303
PP07	.941	584	.894	638	.786	730
PP09	.973	706	.905	**1251**	.882	947
PP10	.945	895	**.833**	**1450**	**.703**	1426
PP13	.955	895	.868	**1260**	.773	953
PP16	.963	719	------	------	.853	1368
PP17	**.865**	672	.903	728	.767	1014
PP18	.945	**2044**	**.805**	**2230**	**.694**	**2475**
PP19	.936	939	**.847**	**1127**	.820	1234
PP27	.940	825	.930	1065	.831	1037
Control mean	.951	794	.914	899	.795	1002
Control SD	.024	230	.03	184	.057	292

A-scores and RT (ms) for each of the participants with developmental prosopagnosia on the three object decision tasks. Values in boldface designate performance deviating significantly from the mean of the matched control group as assessed by means of Crawford and colleagues methods [24, 25].

three of the DPs (PP07, PP16 & PP27) scored within the normal range in terms of both *A* and RT in all conditions. Especially PP07 and PP27 seemed to perform quite well. In the section below entitled "Testing for dissociations", we examine the question of whether any of these normally performing DPs exhibit a dissociation in accordance with criteria suggested by Crawford et al. [15].

Within-class object recognition

To examine within-class object recognition performance, nine of the 10 DPs were tested on the Cambridge Car Memory Test [32] (PP16 did not take this test, and hence her matched control participants were also not included in the analysis), which is equivalent to the Cambridge Face Memory Test just using cars instead of faces as stimuli. Hence, the maximum score in this test is also 72, and chance-level is at 24.

Results. The median score of the DPs was 42 (IQR = 6.5). In comparison the median score of the control participants was 51 (IQR = 10.5). This difference was significant (Mann–Whitney *U* Exact test, $z = -2.6$, $p = .009$). Analysis at the single case level, by means of the methods developed by Crawford and colleagues [24, 25], revealed that only PP17 & PP19 scored outside the range of the control participants (see Table 1).

Discussion

The results presented above suggest that our sample of DPs performed within the normal range on a quite demanding task of visual object recognition with regular line drawings. When the same task involved degraded stimuli, subtle deficits were nevertheless revealed. Hence, with silhouette drawings the group of DPs were significantly impaired relative to the control participants in terms of discriminability and also performed more slowly although this difference in latency was only marginally significant. With fragmented forms the DPs also exhibited prolonged RTs; again an effect that was only marginally significant. When we consider these deficits subtle, it is because three out of eight of the DPs performed within the normal range with silhouettes in terms of both discriminability and RT, whereas eight out of 10 performed within the control range with fragmented forms. Indeed two of the DPs fell within the performance

range of the control sample in all three tasks in terms of both RT and discriminability. A similar pattern was seen for performance on the CCMT. Here the DP group scored significantly below the control group even though only two DPs fell outside the range of the control participants based on single case statistics.

Relating Face and Object Recognition Abilities

So far we have established that the group of DPs performed abnormally with object recognition of degraded stimuli and within-class recognition of objects (cars). This, however, does not indicate that these deficits are directly related to the face processing deficits observed; it could reflect associated deficits. Moreover, one may even question the validity of using impoverished stimuli in the first place.

In many instances where the demand on perceptual differentiation is low, successful classification of an object can be based on a limited amount of information, e.g., a couple of features [14, 17]. This is the reason why we chose to use a difficult object decision tasks. The decision to also use impoverished stimuli is a logical extension of this line of reasoning because subtle deficits may only become apparent when the recognition system is challenged [18]. We note that similar assumptions lie behind the construction of the Cambridge Tests in that they also make use of impoverished stimuli (faces and cars with added noise). It is also the case that the object recognition system often deals with stimuli that are impoverished for natural reasons. Objects may be partly occluded by other objects (causing them to appear fragmented), or they may be viewed in dim light or strong backlight (causing them to appear as silhouettes). Hence, the use of fragmented forms and silhouettes per se does not necessarily make the recognition situation artificial (or un-ecological).

While the use of impoverished stimuli can be justified, and even commendable, the observed association between face and object recognition performance can still be considered coincidental. To show that the connection is more than coincidental, a systematic relationship between variation in object recognition performance and variation in face recognition performance must be established. We examine this possibility below.

Procedure

To examine whether variation in object recognition performance is systematically related to face recognition performance, we compared the scores obtained from the CFMT with scores from the CCMT and the discriminability scores obtained from the two object decision conditions where there were signs of impairment; the one with silhouettes and the one with fragmented forms. This was done by means of correlation analyses. We did not perform the same analyses with the RTs from the object decision tasks with degraded stimuli because the CFMT is based on accuracy only. We report statistics derived by means of Pearson's correlation coefficient (r). These analyses were performed separately for the DPs and control participants.

Results

There was a significant positive correlation between performance on the CFMT and A-scores with silhouettes ($r = .87$, $p = .005$, 95% CI = [.55, 1] based on bootstrap analysis with1000 samples) for the DP group but not for the control group ($r = .07$, $p = .8$). There was also a significant positive correlation between performance on the CFMT and A-scores with fragmented forms ($r = .78$, $p = .007$, 95% CI = [.26, .98] based on bootstrap analysis with1000 samples) for the DP group but again not for the control group ($r = -.04$, $p = .87$). The correlation between the CFMT and the CCMT failed to reach significance in both the DP ($r = .58$, $p = .1$) and the control group ($r = .08$, $p = .76$). For a graphical illustration of the significant findings see Fig 2.
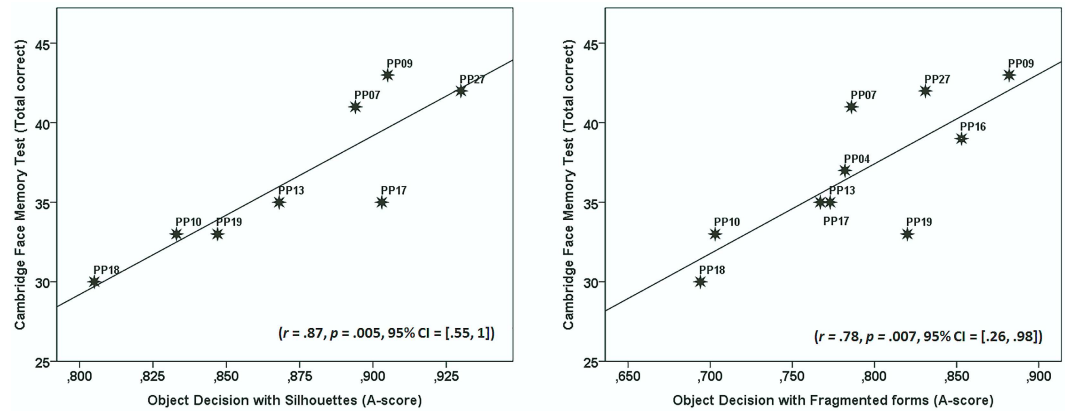
Fig 2. Scatterplots showing the relationship between performance on the Cambridge Face Memory Task and the object decision tasks with silhouettes (left panel) and fragmented forms (right panel) in the DPs. Also shown are the regression lines, the Pearson correlation coefficients (*r*) and their associated *p*-value.

Discussion

For the DP-group, performance on the CFMT correlated significantly with performance on both the object decision task with silhouettes and the object decision task with fragmented forms. While caution should be exercised in interpreting correlations based on small samples (n = 8 for silhouettes and N = 10 for fragmented forms), we note that the correlations were quite reliable as reflected by the fact that the 95% CIs did not include 0. Accordingly, while there is some uncertainty associated with using the correlations as estimates of correlations in the general population of DPs, they do seem robust in that the lower bound was r = .55 for object decision with silhouettes and r = .26 for fragmented forms. These findings suggest that the face and object recognition deficits observed in the DP group are systematically related and thus unlikely to reflect two associated deficits.

Perceptual Matching

Participants

All DPs except for PP16 participated (n = 9). To keep the DP group and the control group matched, we excluded the two controls of PP16 yielding a total of 18 individuals in the control group.

Design

The participants were presented with two stimuli at a time, either two faces or two houses, and had to decide whether the stimuli were identical or differed. They were instructed to press the 'same'-key on a serial response-box (index finger) if the stimuli were identical and the 'different'-key (middle finger) if the stimuli differed in any respect. The stimuli could differ in either 1st order relations (e.g., the nose placed below the mouth in one face but at its normal position in the other), 2nd order relations (e.g., difference in the spacing between the eyes), or in their consistent features (e.g., different shape of the nose). Furthermore, the differences along these three dimensions varied parametrically between the stimulus-pairs as illustrated in Fig 3.

In the task instructions the participants were made aware that there was an equal amount of identical and different stimulus-pairs. They were also informed that the stimuli could differ in
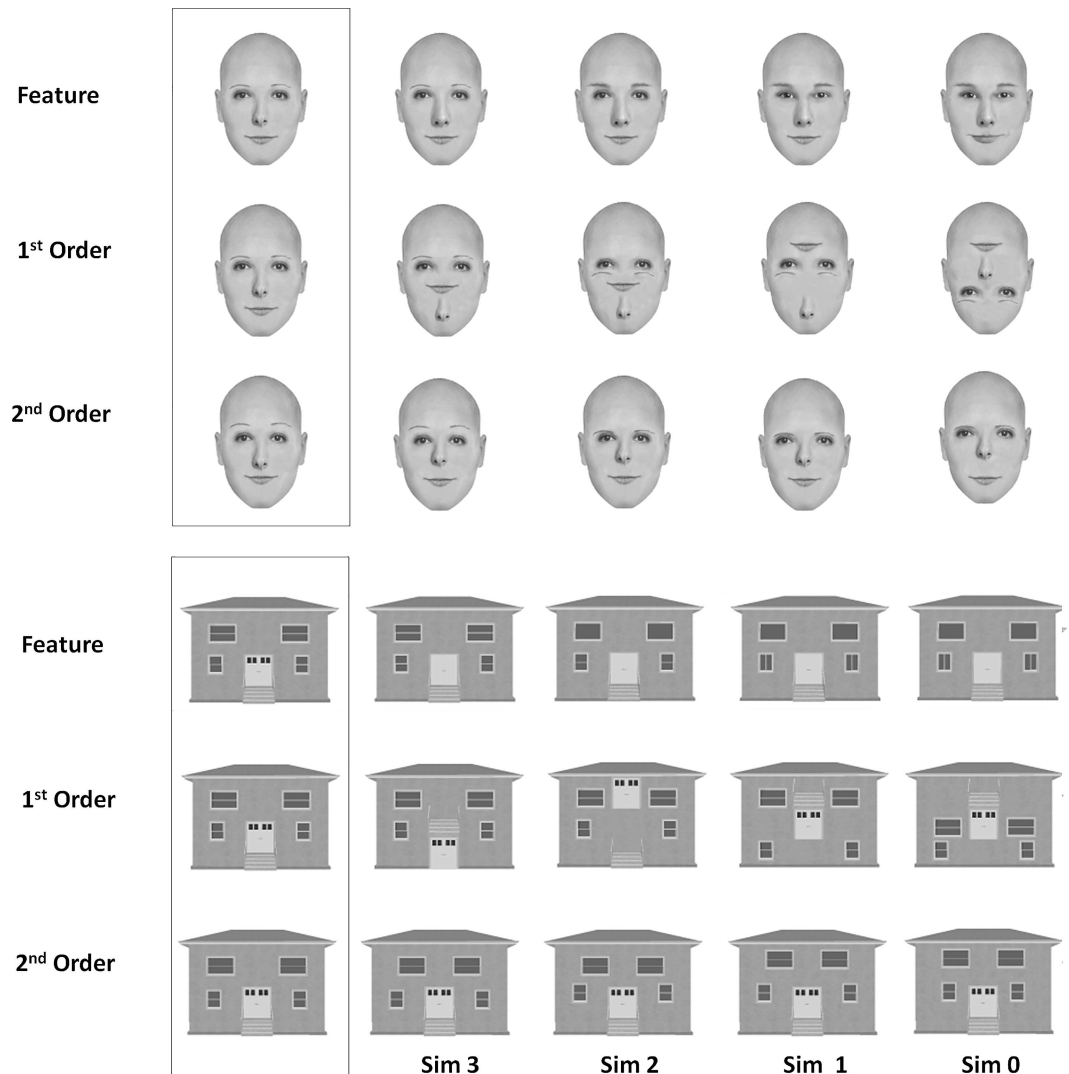
Fig 3. Examples of the face and house stimuli used in the perceptual matching task. 'Sim' designates similarity level with each stimulus differing from the one presented in the boxed area by 1 difference (Sim 3), 2 differences (Sim 2), 3 differences (Sim 1), and four differences (Sim 0).

either 1st order relations, 2nd order relations or in their constituent features. These differences were also illustrated by example stimulus-pairs which were not used in the actual experiment. Finally they were told that the stimulus-pairs could differ in how similar they were; as an example, some pairs would differ in only one feature whereas others would differ in several of their constituent features. The participants were encouraged to respond as fast and as accurately as possible. Prior to the actual experiment the participants performed 48 practice trials with all combinations of differences (1st order, 2nd order and featural) and similarity levels. Stimuli used in practice trials were not used in the actual experimental conditions. Feedback was provided during practice but not during the actual experiment. A similar experimental paradigm, using the same stimuli as the present, has previously been used by Collins, Zhu, Bhatt, Clark and Joseph [33] and by Joseph, DiBartolo and Bhatt [34]. The stimuli used in the present experiment were kindly made available by Jane E. Joseph. Hence, for specification of stimulus parameters, beyond what is given below, we refer to these studies.

Stimuli

A total of 560 stimulus-pairs were presented. 176 were 1st order stimulus-pairs [11 different face-pairs by 4 different similarity levels + 11 different House-pairs by 4 different similarity levels + 11 identical face-pairs by 4 different combinations + 11 identical house-pairs by 4 different combinations], 192 were 2nd order stimulus-pairs [12 different face-pairs by 4 different similarity levels + 12 different House-pairs by 4 different similarity levels + 12 identical face-pairs by 4 different combinations + 12 identical house-pairs by 4 different combinations], and 192 were featural stimulus-pairs [12 different face-pairs by 4 different similarity levels + 12 different House-pairs by 4 different similarity levels + 12 identical face-pairs by 4 different combinations + 12 identical house-pairs by 4 different combinations]. The stimulus-pairs were arranged such that the two faces/houses were presented one above the other. Each stimulus in a pair subtended 2.5–5.3° of visual angle.

Procedure

Each stimulus-pair was shown until the participant made a response or for a maximum duration of 10s. If no response was made within 10s the trail was terminated and counted as an error. The interval between response (or termination) and presentation of the next stimulus-pair was 1.5s. The presentation order of the 560 stimulus-pairs was randomised with a rest period inserted following a row of 56 trials. The rest period was terminated by the participant when (s)he was ready to proceed.

Statistical analyses

We measured overall discriminability between same and different trials in terms of A. Prior to analyses of A-scores we screened for individuals exhibiting extreme hit/false alarm rates (rates approaching 1). No such cases were observed. While A can be computed for faces and houses separately over all conditions, it cannot also be computed separately for the 1st order, 2nd order or featural condition because stimulus presentation was randomised such that stimulus-pairs with identical faces/houses (same-responses) could not be assigned to a particular condition (1st order, 2nd order or featural). Furthermore, as accuracy was quite low, and for some individuals at chance-level in some of the conditions (see below), we decided not to perform a full analysis of the RT-data. Finally, given that we were interested in effects of visual similarity, a dimension that only varied for different trials, analyses of accuracy data were limited to different trials only (number of correct responses). Comparisons between the DPs and the control participants were based on non-parametric tests as several of the variables of interest departed from normality.

Results

Discriminability. The median A-score of the DPs for faces was .932 (IQR = .129) and .978 (IQR = .029) for the control group. This difference was significant (Mann–Whitney U Exact test, $z = -2.5$, $p = .011$). The median A-score of the DPs for houses was .939 (IQR = .1) and .965 (IQR = .02) for the control group. This difference was also significant (Mann–Whitney U Exact test, $z = -2.8$, $p = .004$).

Accuracy. To examine whether the DP group generally performed differently than the control group on 1st order relations, 2nd order relations and featural differences for faces and houses, we first computed the mean number of correct responses to different trials averaged over the four similarity levels for each participant for each of the six conditions. We next compared these median scores of the DP group and the control group for the six conditions by

Table 4. Group data and statistics from the perceptual matching task.

Condition	DP group	Control group	z	p-value
1st order faces	44 (13–44)	44 (33–44)	-.84	.41
1st order houses	43 (1–44)	44 (41–43)	-2.0	.05
2nd order faces	32 (1–44)	45 (16–48)	-2.96	.002
2nd order houses	26 (0–42)	39 (13–47)	-2.63	.007
Fetural faces	39 (10–46)	46 (37–48)	-2.56	.009
Featural houses	38 (11–43)	42 (31–47)	-2.82	.004

Comparison of the median number of correct responses to different trials (range in brackets), averaged over the four similarity levels, for the DP group and the control group. Differences between groups were examined by means of Mann-Whitney U Exact tests.

doi:10.1371/journal.pone.0165561.t004

means of Mann-Whitney U Exact test. As can be seen from Table 4, where these results are summarized, the DP group differed from the control group in all conditions except for 1st order differences in faces. Although the range also differed between the groups, the variability between them did not differ significantly (Moses Test of Extreme Reaction; all p's > .05). This suggests that the differences observed reflect differences in the medians rather than in the distributions of scores.

Effects of visual similarity. To examine effects of visual similarity in the six conditions, we first computed the mean percentage of correct responses for each similarity level across the individuals within each group. We then computed the correlation between similarity level and the mean percentage of correct responses to different trials for each of the six conditions for each of the two groups by means of Pearson's correlation coefficient. Note that these analyses ignore intersubject variability because they are based on the grand mean of all individuals across a particular similarity level and not on the means of the individual participants for a particular similarity level. Accordingly, the degrees of freedom for each analysis were 4. The fact that the correlation analyses are based on averaged data also means that any effect revealed by them could in principle reflect the performance of a few individuals. We nevertheless present them here as they give a general impression of the data. In the instances where the analyses indicated a potential difference between the groups, we thus tested these potential group differences more formally (see below). The correlation analyses were based on one-tailed statistics as we had a directional hypothesis: Accuracy will decrease as similarity increase. As can be seen from Fig 4, where the results of these analyses are summarised, accuracy generally did decrease as similarity increased. However, in 5 out of 12 cases this linear effect was not significant: 1st order differences in faces (Controls); 1st order differences in houses (DPs and Controls), 2nd order differences in houses (DPs and Controls).

As can be seen in Fig 4, the regression lines for the DPs and the control participants are rather parallel for all conditions with houses suggesting that the effect of visual similarity did not impact differently on the DP group and the control group. In comparison, the slopes seem steeper for the DP group than for the control group in the conditions with faces, suggesting that increasing visual similarity impacted more on the performance of the DPs than the control participants' during processing of faces. To examine these potential group differences more formally we first computed individual difference-scores by subtracting the accuracy on similarity level 3 (high similarity) from the accuracy on similarity level 0 (low similarity); the larger the difference-score, the greater the impact of visual similarity. This was done for every individual and for each condition (1st order, 2nd order and the featural condition). We next examined whether the difference-scores from each condition differed significantly between the DP and the control group by means of Mann–Whitney U Exact tests. Hence, what is basically tested is
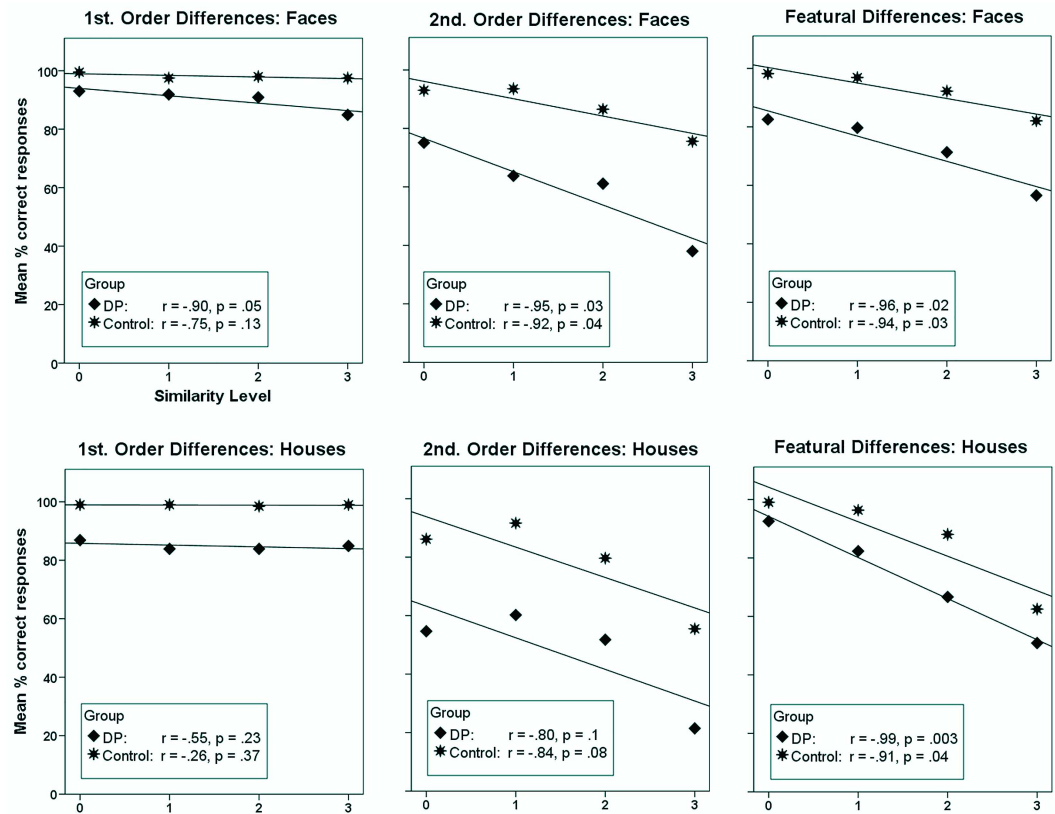
Fig 4. Scatterplots showing the relationship between similarity level and mean % correct responses to different trials in the six conditions of the perceptual matching experiment for the developmental prosopagnosics (DP) and the control group. Similarity increases parametrically from level 0 to level 3, with 0 indicating minimum similarity with four differences and 3 indicating maximum similarity with only one difference. Also shown is the Pearson correlation coefficient (*r*) and its associated (one-tailed) *p*-value.

doi:10.1371/journal.pone.0165561.g004

whether there is an interaction between Group and Similarity level (focusing on the extreme ends of the similarity variable ignoring any difference at intermediate similarity levels 1 and 2), and, as opposed to the correlation analyses presented above, these analyses are based on individual scores and not average scores. These analyses revealed no significant difference for neither the 1st order condition ($z = -1.02$, $p = .37$) or the featural condition ($z = -1.57$, $p = .12$). However, in the 2nd order condition, the difference between the DPs and controls was significantly larger for similarity level 3 than for similarity level 0 ($z = -2.27$, $p = .02$). In summary, the impact of visual similarity did not differ reliably between DPs and the controls with the exception of perceiving differences in 2nd order relations for faces where the performance of the DP group was more affected than that of the control group as similarity increased.

Single case analysis. To examine the performance of the DPs individually, we compared the % correct responses on different trials in each of the six conditions of the perceptual matching task of each DP with the mean % correct responses of the control participants using the methods developed by Crawford and colleagues [24, 25]. As can be seen in Table 5, where the results of these comparisons are summarized, four of the DPs (PP07, PP10, PP19 & PP27) scored within the normal range in all conditions. In addition, one DP (PP04) performed within the normal range on all three conditions with houses. Four of the DPs (PP04, PP13, PP17 & PP18) exhibited significantly reduced performance with faces in one or more conditions, as did

Table 5. Single case accuracy in the perceptual matching task.

Case	1st order		2nd order		Featural	
	Faces	Houses	Faces	Houses	Faces	Houses
PP04	100	96	**54**	79	**57**	75
PP07	100	98	92	54	96	83
PP09	100	96	67	**19**	90	81
PP10	100	98	85	73	90	90
PP13	100	100	56	48	**69**	**73**
PP17	**87**	**75**	**2**	**0**	**21**	**23**
PP18	**30**	**2**	**29**	**0**	**58**	**71**
PP19	96	100	82	88	92	79
PP27	100	100	69	63	81	83
Control mean	98	99	87	78	92	87
Control SD	5.8	2.1	17.6	17.1	7.2	7.3

% correct responses to different trials in the six conditions of the perceptual matching task for each of the individual DPs. Values in boldface designate performance deviating significantly from the mean of the matched control group as assessed by means of Crawford and colleagues methods [24, 25].

doi:10.1371/journal.pone.0165561.t005

four DPs with performance with houses (PP09, PP13, PP17 & PP18). Hence, in terms of frequency, the DPs were just as impaired with houses as they were with faces.

As discussed in the introduction, good performance in terms of accuracy may be achieved at the price of prolonged RTs. With this in mind, we wanted to examine whether the four DPs, who performed within the normal range of the control group in all conditions, would also exhibit normal performance in terms of RT. Hence, we repeated the analyses presented above but now on the RTs of these DPs (PP07, PP10, PP19 & PP27) and their eight controls. We considered these analyses appropriate for this particular subsample of the DP group because they performed within the range of the control participants in terms of accuracy. We also included PP04 in these analyses when examining RT-performance with houses because PP04 had within normal-range performance in all three conditions with houses. Accordingly, 10 control participants were included in RT analyses with houses. As can be seen from Table 6 none of the five DPs were significantly impaired in terms of RT. In conclusion, PP07, PP10, PP19 & PP27 performed within the normal range with both faces and houses with respect to both accuracy and RT, and PP04 did so for houses but not for faces.

Table 6. Single case RT (ms) in the perceptual matching task.

Case	1st order		2nd order		Featural	
	Faces	Houses	Faces	Houses	Faces	Houses
PP04	-------	2426	-------	4783	-------	3677
PP07	1447	1440	2826	3344	2515	2276
PP10	1395	1607	2219	3066	2310	2454
PP19	1533	2140	3520	2853	2669	2430
PP27	1543	1596	1744	2779	1917	2142
Control mean	1294	1636	2161	2831	1961	2099
Control SD	277	426	692	913	644	605

Mean RT to different trials in the six conditions of the perceptual matching task for the subsample of DPs with accuracy within the normal range. No values deviated significantly from the mean of the matched control group as assessed by means of Crawford and colleagues methods [24, 25].

doi:10.1371/journal.pone.0165561.t006

Discussion

Considered as a group, the DPs generally exhibited poorer discriminability of both faces and houses compared with controls. This was reflected in their performance in all conditions except for 1st order differences in faces where the DP group did not differ significantly from the control group in accuracy. In general, however, the DP group was not more affected by visual similarity than the control group in the 1st order and featural conditions. Only when required to make discriminations based on differences in 2nd order relations in faces was the DP group more affected by visual similarity than the control group. It is worth noting that this finding cannot reflect that this condition just happened to be the most challenging. While it was difficult, it was not more difficult than the 2nd order condition with houses where the DP group and the control group were equally affected by increasing levels of visual similarity. Accordingly, at least this effect seems specific to faces.

Analyses of the individual performances of the DPs revealed a similar pattern as described for the group differences: Four DPs fell outside the range of the control participants in one or more conditions with faces and four did so with houses. Except for one DP (PP04), who was only impaired with faces, it was the same DPs who fell outside the normal range with faces and with houses. Likewise, the four DPs who performed within the normal range in all conditions with faces also performed within the normal range with houses. Accordingly, expect for PP04 there was little evidence of a face selective impairment in this task.

Testing for Dissociations

As described in the introduction we test dissociations by means the criteria suggested by Crawford et al. [15] which entail that: (i) a person's performance on task X must differ significantly from that of the normal population, and (ii) the *difference* in performance of that person on tasks X and Y must differ significantly from the difference-scores of the normal population on tasks X and Y.

Two of the four DPs (PP07 & PP27), who performed within the normal range with houses in the perceptual matching test, also performed within the normal range on the object decision tasks and the Cambridge Car Memory Task. Hence, these individuals fulfil one of the premises for a dissociation. What further needs to be examined is whether the difference they exhibit in their performance with faces and objects is significantly greater than what can be expected in the normal population. To test this we adopted the approach developed by Crawford and colleagues for testing for the presence of dissociations [24, 25] (implemented in the programs DissocsBayes_ES and BSDT_Cov). In addition to testing whether a case's scores on tasks X and Y differ significantly from the normal population by using the control participants' mean and SD as sample statistics (rather than as population parameters), this test estimates whether the case's standardized difference between tasks X and Y differs significantly from the standardized differences in controls taking into account the correlation between tasks X and Y in the control sample. The test also provides a Bayesian point estimate of the percentage of the control population exhibiting a more extreme discrepancy in same direction as the case and a 95% CI of this estimate.

Specifically, we compared the difference in each DP's score on the CFMT and their A sensitivity score in the object decision task with silhouettes; the condition in which the DPs as a group performed worse than the control group in terms of discriminability (A). These comparisons revealed that neither PP07 nor PP27 exhibited a dissociation between these tasks that significantly exceeded what can be expected in the normal population ($p = .27$ and $p = .07$ for PP07 and P27 respectively). For PP07 the Bayesian point estimate for the difference was 14% (95% CI = [2, 35]). For PP27 it was 4% (95% CI = [0.2, 13]).

When the same analyses were conducted on the scores from the CFMT and the CCMT, neither PP07 nor PP27 exhibited a dissociation that significantly exceeded what can be expected in the normal population ($p = .06$ and $p = .71$ for PP07 and P27 respectively). For PP07 the Bayesian point estimate for the difference was 3% (95% CI = [0.04, 13]). For PP27 it was 35% (95% CI = [7, 71]).

General Discussion

If face recognition and object recognition depend on separate domain-specific operations, we would expect to find individuals with face recognition deficits who perform within the normal range with recognition of other categories of objects, and the reverse; a double-dissociation. In the present study we examined one side of such a potential double-dissociation. This was done by testing whether a group of individuals ($N = 10$) with developmental prosopagnosia (DP) performed within the normal range on demanding visual object processing tasks. While we do find that our group of DPs performed within the normal range on a demanding object recognition task with regular drawings, we also find that they as a group are impaired relative to control participants with degraded stimuli (silhouettes and fragmented forms) and with within-class recognition of objects (cars). In addition, the DPs' object recognition performance with degraded stimuli is systematically related to the severity of their face recognition deficit. Hence, we find that their face recognition performance is highly correlated with recognition of both silhouette objects ($r = .87$) and fragmented objects ($r = .78$).

So, is there no evidence of face selectivity in our sample of DPs? If one considers the DPs individually there appears to be some evidence because many of them performed within the normal range on many of the tasks. Indeed, two of them (PP07 & PP27) scored within the normal range on *all* tasks including the ones with degraded stimuli, which is quite impressive. However, when judged on a group level it is also the case that the DPs were in fact significantly impaired in the majority of the tasks performed: Within-class recognition of objects (cars), object decision with degraded material, and perceptual matching of faces and houses (2nd order and featural differences). Hence, we are caught in an old dilemma: Should (neuropsychological) evidence be based on individual cases, which may leave something to be desired in terms of representability and statistical power, or should it rather be based on group studies which typically excel in both respects but which may yield results that are basically "averaging" artefacts [35]? And how do we approach the old dogma, in cognitive neuropsychology, that dissociations trumps associations when we can show robust correlations on a group level, while some individuals exhibit borderline significant dissociations? The most reasonable thing to do is to weigh the evidence, considering the strength and weaknesses of each approach.

In doing so we first note that even though not all DPs performed significantly outside the normal range when considered individually, especially not with fragmented forms and within-class recognition of objects, most of them did perform in the lower normal range in terms of either accuracy (discriminability) or RT. Was this not the case, the reliable group differences would not have been found. Secondly, even the two DPs who performed the best (PP07 & PP27) did not fulfil the criteria for a dissociation. Although PP07 almost did show a reliable dissociation when considering the difference in her performance between the CFMT and the CCMT ($p = .06$), the difference in her performance between the CFMT and the object decision task with silhouettes was not significantly different from what can be expected in the normal population ($p = .27$). In fact, approximately 14% of the normal population would be expected to exhibit a more extreme discrepancy in the same direction as PP07. The opposite was true of PP27. He exhibited a marginally significant dissociation in his performance on the CFMT and the object decision task with silhouettes ($p = .07$), but not in his performance on the CFMT and

the CCMT (p = .71), where approximately 35% of the normal population would be expected to exhibit a more extreme discrepancy in the same direction as PP27. Finally, the systematic relationships between performance with degraded material and face recognition performance suggests that the results of the group comparisons do not reflect averaging artefacts. In our opinion this finding is important because a correlation, as opposed to a dissociation, does not depend on the performance of a control group; a performance which may vary from sample to sample (cf. the discussion below). With these considerations in mind, we do not think that our findings yield support for face selective impairments in our sample of DPs. Rather, our data suggests that their object recognition performance is also impaired, although often in a subtle manner. Indeed, the finding that the severity of their object recognition deficit is systematically related to the severity of their face recognition deficit is difficult to account for should their face deficit be selective.

If our findings are representative of DP in general, our conclusion regarding non-selectivity contrasts with the conclusion reached in several other studies reporting a dissociation between face and object recognition ability. To our knowledge, the most compelling evidence for a dissociation involving impaired face recognition and normal object recognition has been reported by Duchaine and Nakayama [36]. They compared the performance of seven DPs on old/new recognition memory tests involving the following categories: Faces, Cars, Houses, Scenes, Horses, Tools, and Guns. Of these DPs, two cases (M1 and M2) performed within the normal range on all contrast categories when performance was assessed by means of discrimination sensitivity, and one case (F2) performed within the lower normal range on all categories. However, none of these DPs performed within the normal range in terms of RT. Cases M1 and M2 –who were the best performing of the DPs in terms of accuracy on the contrast tasks–fell more than two SDs below the mean on respectively all or four of the contrast categories, whereas case F2 fell more than two SDs below the mean on two of the contrast categories. Hence, if both accuracy and RT are considered, none of the DPs reported by Duchaine and Nakayama [36] performed within the normal range with objects. The same is true of other studies where the performance of DPs fall outside the normal range on either RT or accuracy (e.g., [10, 37]).

Even though we advocate for an association between face and object recognition impairments in DP, which might lead to the conclusion that faces are not 'special', there is no doubt in our minds that the DPs included in our study experience tremendously greater problems with face recognition than they do with object recognition in everyday life. In this sense faces are special. To make our stance clear, we find no compelling evidence in the literature on DP or in the present study suggesting that face recognition is so special that it can be selectively impaired, but it is special enough to place greater strains on the recognition system–or parts of the system–than recognition of other categories of objects typically do, causing recognition of faces to suffer significantly more. Accordingly, it is only when the recognition system is seriously challenged that recognition deficits with other objects than faces becomes apparent. We also note that our claim regarding a lack of convincing cases of selectivity (dissociations) in face recognition only applies to the field of DP. While the same may be true for acquired prosopagnosia (for a discussion of this see [38]), a detailed discussion of this evidence falls outside the scope of the present paper.

A concluding note on selectivity

But what is "selectivity" anyway? In a sense, selectivity is a result of how observations are treated. The traditional way, which we also adopted here, is to classify an individual's performance as abnormal if it falls outside the range of normal participants; typically defined as 2 SDs below the mean of the normal participants. If it does not, the individual's performance is

considered normal. To some extent this cut-off is arbitrary and its application will yield different outcomes depending on the norms it is applied to. However, if we abandon the use of cut-off scores, selectivity ceases to exist. Then it becomes a question of degree rather than kind. We are not arguing that a cut-off at 2 SDs necessarily represents an insensible way to operationalize a deficit. Our point is rather that it is 'just' an operationalization which does not (necessarily) imply that the cognitive or neural machinery which supports recognition must consist of two separate systems (modules) if it can be demonstrated that some individuals show "> 2SDs selectively" impaired face recognition, whereas others show "> 2SDs selectively" impaired object recognition.

If we look at things from this perceptive it seems productive to further investigate, for example, what it is that recognition of degraded objects has in common with face recognition. One possibility is that it is related to the derivation of global shape information, which has been suggested to play a pivotal role in recognition of both silhouettes and fragmented forms [17], as well as face recognition [39].

Supporting Information

S1 Dataset. Individual data for developmental prosopagnosics and control participants. (SAV)

Acknowledgments

Author Contributions

Conceptualization: CG RS.

Data curation: CG SKK.

Formal analysis: CG SKK.

Funding acquisition: CG RS.

Investigation: SKK.

Methodology: CG RS.

Project administration: CG RS.

Resources: CG SKK RS.

Software: CG SKK.

Supervision: CG RS.

Validation: CG SKK RS.

Visualization: CG.

Writing – original draft: CG RS SKK.

Writing – review & editing: CG SKK RS.

References

1. McKone E, Robbins R. Are Faces Special? In: Rhodes G, Calder A, Johnson M, Haxby JV, editors. Oxford Handbook of Face Perception New York: ' Oxford University Press'; 2011. p. 149–76.

2. Gazzaniga MS, Ivry RB, Mangun GR. Cognitive neuroscience: The biology of the mind. New York: W. W. Norton & Company, Inc.; 2015.

3. Moscovitch M, Winocur G, Behrmann M. What Is Special about Face Recognition? Nineteen Experiments on a Person with Visual Object Agnosia and Dyslexia but Normal Face Recognition. J Cogn Neurosci. 1997; 9(5):555–604. Epub 1997/10/01. doi: 10.1162/jocn.1997.9.5.555 PMID: 23965118.

4. Caramazza A, Mahon BZ. The organisation of conceptual knowledge in the brain: The future's past and some future directions. Cogn Neuropsychol. 2006; 23(1):13–38. Epub 2006/01/01. doi: 10.1080/02643290542000021 PMID: 21049320.

5. Tanaka JW, Gordon I. Features, configuration, and holistic face processing. In: Calder AJ, Rhodes G, Johnson MK, Haxby J, editors. The Oxford Handbook of Face Perception. Oxford: Oxford University Press; 2011. p. 177–94.

6. Diamond R, Carey S. Why faces are and are not special: an effect of expertise. J Exp Psychol Gen. 1986; 115(2):107–17. Epub 1986/06/01. PMID: 2940312.

7. Damasio AR. Category-related recognition defects as a clue to the neural substrates of knowledge. Trends Neurosci. 1990; 13(3):95–8. Epub 1990/03/01. PMID: 1691878.

8. Duchaine B. Developmental prosopagnosia. In: Calder AJ, Rhodes G, Johnson MK, Haxby J, editors. The Oxford Handbook of face perception Oxford: Oxford University Press; 2011. p. 821–38.

9. Nachson I. On the modularity of face recognition: the riddle of domain specificity. Journal of clinical and experimental neuropsychology. 1995; 17(2):256–75. Epub 1995/04/01. doi: 10.1080/01688639508405122 PMID: 7629271.

10. Duchaine BC, Yovel G, Butterworth EJ, Nakayama K. Prosopagnosia as an impairment to face-specific mechanisms: Elimination of the alternative hypotheses in a developmental case. Cogn Neuropsychol. 2006; 23(5):714–47. Epub 2006/07/01. doi: 10.1080/02643290500441296 PMID: 21049351.

11. Behrmann M, Avidan G, Marotta JJ, Kimchi R. Detailed exploration of face-related processing in congenital prosopagnosia: 1. Behavioral findings. J Cogn Neurosci. 2005; 17(7):1130–49. Epub 2005/08/17. doi: 10.1162/0898929054475154 PMID: 16102241.

12. McConachie HR. Developmental prosopagnosia. A single case report. Cortex; a journal devoted to the study of the nervous system and behavior. 1976; 12(1):76–82. Epub 1976/03/01. PMID: 1261287.

13. Gauthier I, Behrmann M, Tarr MJ. Can face recognition really be dissociated from object recognition? J Cogn Neurosci. 1999; 11(4):349–70. Epub 1999/09/03. PMID: 10471845.

14. Gerlach C, Marstrand L, Habekost T, Gade A. A case of impaired shape integration: Implications for models of visual object processing. Visual Cognition. 2005; 12(8):1409–43.

15. Crawford JR, Garthwaite PH, Gray CD. Wanted: fully operational definitions of dissociations in single-case studies. Cortex; a journal devoted to the study of the nervous system and behavior. 2003; 39(2):357–70. Epub 2003/06/06. PMID: 12784893.

16. Laws KR. "Illusions of normality": a methodological critique of category-specific naming. Cortex; a journal devoted to the study of the nervous system and behavior. 2005; 41(6):842–51; discussion 52–3. Epub 2005/12/15. PMID: 16350664.

17. Gerlach C, Toft KO. Now you see it, now you don't: The context dependent nature of category-effects in visual object recognition. Visual Cognition. 2011; 19(10):1262–97. doi: 10.1080/13506285.2011.630044

18. Starrfelt R, Habekost T, Gerlach C. Visual processing in pure alexia: a case study. Cortex; a journal devoted to the study of the nervous system and behavior. 2010; 46(2):242–55. Epub 2009/05/19. doi: 10.1016/j.cortex.2009.03.013 PMID: 19446802.

19. Maurer D, Grand RL, Mondloch CJ. The many faces of configural processing. Trends Cogn Sci. 2002; 6(6):255–60. PMID: 12039607.

20. Duchaine B, Nakayama K. The Cambridge Face Memory Test: results for neurologically intact individuals and an investigation of its validity using inverted face stimuli and prosopagnosic participants. Neuropsychologia. 2006; 44(4):576–85. Epub 2005/09/20. doi: 10.1016/j.neuropsychologia.2005.07.001 PMID: 16169565.

21. Duchaine B, Germine L, Nakayama K. Family resemblance: ten family members with prosopagnosia and within-class object agnosia. Cogn Neuropsychol. 2007; 24(4):419–30. doi: 10.1080/02643290701380491 PMID: 18416499

22. Bowles DC, McKone E, Dawel A, Duchaine B, Palermo R, Schmalzl L, et al. Diagnosing prosopagnosia: effects of ageing, sex, and participant-stimulus ethnic match on the Cambridge Face Memory Test and Cambridge Face Perception Test. Cogn Neuropsychol. 2009; 26(5):423–55. Epub 2009/11/19. doi: 10.1080/02643290903343149 PMID: 19921582.

23. Freeman P, Palermo R, Brock J. Faces and Emotion Questionnaire.

24. Crawford JR, Garthwaite PH, Porter S. Point and interval estimates of effect sizes for the case-controls design in neuropsychology: rationale, methods, implementations, and proposed reporting standards. Cogn Neuropsychol. 2010; 27(3):245–60. Epub 2010/10/12. doi: 10.1080/02643294.2010.513967 PMID: 20936548.

25. Crawford JR, Garthwaite PH, Ryan K. Comparing a single case to a control sample: testing for neuro-psychological deficits and dissociations in the presence of covariates. Cortex; a journal devoted to the study of the nervous system and behavior. 2011; 47(10):1166–78. Epub 2011/04/05. doi: 10.1016/j.cortex.2011.02.017 PMID: 21458788.

26. Baron-Cohen S, Wheelwright S, Skinner R, Martin J, Clubley E. The autism-spectrum quotient (AQ): evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathe-maticians. Journal of autism and developmental disorders. 2001; 31(1):5–17. Epub 2001/07/07. PMID: 11439754.

27. Gerlach C. Structural similarity causes different category-effects depending on task characteristics. Neuropsychologia. 2001; 39(9):895–900. Epub 2001/08/23. PMID: 11516442.

28. Gerlach C, Law I, Gade A, Paulson OB. Perceptual differentiation and category effects in normal object recognition: a PET study. Brain: a journal of neurology. 1999; 122 (Pt 11):2159–70. Epub 1999/11/05. PMID: 10545400.

29. Snodgrass JG, Vanderwart M. A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. Journal of experimental psychology Human learning and memory. 1980; 6(2):174–215. Epub 1980/03/01. PMID: 7373248.

30. Lloyd-Jones TJ, Humphreys GW. Perceptual differentiation as a source of category effects in object processing: evidence from naming and object decision. Memory & cognition. 1997; 25(1):18–35. Epub 1997/01/01. PMID: 9046867.

31. Zhang J, Mueller ST. A note on ROC analysis and non-parametric estimate of sensitivity. Psychome-trika. 2005; 70(1):203–12. doi: 10.1007/s11336-003-1119-8

32. Dennett HW, McKone E, Tavashmi R, Hall A, Pidcock M, Edwards M, et al. The Cambridge Car Mem-ory Test: a task matched in format to the Cambridge Face Memory Test, with norms, reliability, sex dif-ferences, dissociations from face memory, and expertise effects. Behav Res Methods. 2012; 44 (2):587–605. Epub 2011/10/21. doi: 10.3758/s13428-011-0160-2 PMID: 22012343.

33. Collins HR, Zhu X, Bhatt RS, Clark JD, Joseph JE. Process and domain specificity in regions engaged for face processing: an fMRI study of perceptual differentiation. J Cogn Neurosci. 2012; 24(12):2428–44. Epub 2012/08/02. doi: 10.1162/jocn_a_00273 PMID: 22849402; PubMed Central PMCID: PMCPmc4527533.

34. Joseph JE, DiBartolo MD, Bhatt RS. Developmental changes in analytic and holistic processes in face perception. Front Psychol. 2015; 6:1165. Epub 2015/08/25. doi: 10.3389/fpsyg.2015.01165 PMID: 26300838; PubMed Central PMCID: PMCPmc4528094.

35. Caramazza A. On drawing inferences about the structure of normal cognitive systems from the analy-sis of patterns of impaired performance: the case for single-patient studies. Brain Cogn. 1986; 5 (1):41–66. Epub 1986/01/01. PMID: 3954906.

36. Duchaine B, Nakayama K. Dissociations of face and object recognition in developmental prosopagno-sia. J Cogn Neurosci. 2005; 17(2):249–61. Epub 2005/04/07. doi: 10.1162/0898929053124857 PMID: 15811237.

37. Garrido L, Duchaine B, Nakayama K. Face detection in normal and prosopagnosic individuals. J Neu-ropsychol. 2008; 2(Pt 1):119–40. Epub 2009/04/01. PMID: 19334308.

38. Busigny T, Joubert S, Felician O, Ceccaldi M, Rossion B. Holistic perception of the individual face is specific and necessary: evidence from an extensive case study of acquired prosopagnosia. Neuropsy-chologia. 2010; 48(14):4057–92. Epub 2010/09/30. doi: 10.1016/j.neuropsychologia.2010.09.017 PMID: 20875437.

39. Goffaux V, Hault B, Michel C, Vuong QC, Rossion B. The respective role of low and high spatial fre-quencies in supporting configural and featural processing of faces. Perception. 2005; 34(1):77–86. Epub 2005/03/19. PMID: 15773608.