Familiarity Facilitates

Feature-based Face Processing

Matteo Visconti di Oleggio Castello[1], Kelsey G. Wheeler[1], Carlo Cipolli[2], M. Ida Gobbini[1,2,*]

1: Department of Psychological and Brain Sciences, Dartmouth College, Hanover NH 03755, USA

2: Dipartimento di Medicina Specialistica, Diagnostica e Sperimentale (DIMES), Medical School, University of Bologna, 40126 Bologna, Italy

*: Corresponding author

mariaida.gobbini@unibo.it; maria.i.gobbini@dartmouth.edu

Abstract

Recognition of personally familiar faces is remarkably efficient, effortless and robust. We asked if feature-based face processing drives facilitated detection of familiar faces by testing the effect of face inversion on a visual search task for familiar and unfamiliar faces. Because face inversion disrupts configural and holistic face processing, we hypothesized that inversion would diminish the familiarity advantage to the extent that it is mediated by such processes. Subjects detected personally familiar and stranger target faces in arrays of two, four, or six face images. Subjects showed significant facilitation of personally familiar face detection for both upright and inverted faces, in terms of reaction times for target present trials and both reaction times and search rate for target absent trials. The effect of familiarity even on target absent trials suggests that familiarity facilitates rejection of unfamiliar distractors as well as detection of familiar targets. The preserved familiarity effect for inverted faces suggests that facilitation of face detection afforded by familiarity reflects mostly feature-based processes.

Keywords: personally familiar faces; visual search; configural processing; feature-based processing; face inversion.

Introduction

Our social life revolves around familiar faces. Recognizing the faces of friends, coworkers, and acquaintances signals the social context in which we find ourselves and is a necessary precondition for appropriate social interactions. Despite the vital importance of familiar face recognition, we know little about the algorithms that our brain uses to generate and activate representations that are surprisingly robust to noise (Jenkins & Burton, 2011). In this manuscript, we present new data from a visual search task and discuss possible mechanisms allowing efficient processing of familiar faces.

Recently we showed that personally familiar faces can be detected with reduced attentional resources, and can be processed also without conscious awareness (Gobbini et al., 2013). With a saccadic reaction paradigm, we found that participants were able to detect familiar faces in 180 ms (Visconti di Oleggio Castello & Gobbini, 2015), a latency shorter than the known evoked potentials that differentiate familiar from stranger faces. Moreover, social cues such as eye-gaze or head orientation are also processed faster when conveyed by familiar faces (Visconti di Oleggio Castello, Guntupalli, Yang, & Gobbini, 2014). These results point to a facilitation of familiar face processing that precedes the activation of a conscious and complete representation, and that extends to the local features of a familiar face.

To test the general hypothesis that the efficient detection of familiar faces also relies primarily on feature-based processing, we assessed here whether individual-specific facial features can be processed in an optimized way and thus afford faster detection and recognition of personally familiar faces. We investigate the effect of face inversion on visual search for personally familiar faces and strangers' faces.

The rationale of this experiment stems from the line of evidence that face inversion disrupts holistic (and/or configural) processing, and favors feature-based processing (Farah, Tanaka, & Drain, 1995; Freire, Lee, & Symons, 2000; see also Valentine, 1988; McKone & Yovel, 2009; Rossion, 2009; Xu & Tanaka, 2013). We hypothesized that if familiar face recognition uses primarily idiosyncratic local facial features, then the advantage for personally familiar faces should be maintained with inverted faces; on the contrary, if familiar face recognition relies overall on holistic processes, face inversion should eliminate the familiarity advantage.

Methods

Participants

19 subjects (12 male, mean age: 24.79, SD 3.71) from three groups of friends participated in the experiment. We chose friends that had extensive daily interaction with each other occurring for at least one year prior to the experiment. They were recruited from the Dartmouth College community. All had normal or corrected-to-normal vision. Subjects were reimbursed for their participation; all gave written informed consent to use their pictures for research and to participate in the experiment. The local IRB committee approved the experiment (Protocol 21200).

Stimuli

Prior to the experiment, subjects and their friends had their pictures taken to be used as stimuli in the experiment. To ensure that all stimuli were of equal image quality, pictures were taken in a photo studio with standardized lighting, camera placement and camera settings. Images were cropped and converted to gray-scale using custom code written in Python on Mac OS X 10.9.5. The average pixel intensity of each image (ranging from 0 to 255) was set to 128 with a standard deviation of 40 using the SHINE toolbox (function *lumMatch*) (Willenbockel et al., 2010) in MATLAB (R2014a).

For each subject we created three sets of images: target familiar faces (two identities: one male, one female), target stranger faces (two identities: one male, one female), and distractor stranger faces (twelve identities: 6 male, 6 female). For each identity we used two different pictures taken in the same session to reduce image-specific learning. The familiar targets were chosen pseudo-randomly from the subject's friends. The pictures of the 14 stranger individuals (12 distractor identities and 2 target identities) were taken at the University of Vermont with the same lighting, camera placement and settings as used for subjects recruited at Dartmouth College. Inverted stimuli were created by rotating the images 180º.

Stimuli for visual search trials consisted of two, four, or six face images positioned on the vertices of a regular hexagon centered on the fixation point, such that the center of each image was 7° of visual angle from the fixation point. Each image subtended 4° x 4° of visual angle. The position of the stimuli always created a shape symmetrical with respect to the fixation point (see Figure 1). All face images for each block were either upright or inverted.

Experimental setup

The experiment was run on a GNU/Linux workstation (Xubuntu 14.04 with low-latency kernel 3.13, CPU AMD FX-4350 quad-core 4.2 GHz, 8GB RAM, AMD Radeon R9 270 video card with radeon drivers) and a DELL 2000FP screen, set at a resolution of 1600x1200 pixels with a 60hz refresh rate, using Psychtoolbox (version 3.0.12) in MATLAB (R2014b).

Subjects sat at a distance of approximately 50 cm from the screen (eyes to screen) in a dimly lit room.

Task

Subjects were briefly familiarized with the images used in the visual search task before starting the experiment. Images (both upright and inverted) were presented in random order. Each image was presented for two seconds. After the image disappeared, subjects were required to press a key to continue to the next image. They were instructed to carefully observe each face for the entire presentation and to continue at their own pace.

Then, the visual search session consisted of eight blocks, with a short break after the first four blocks. In each block, subjects were instructed to search for one of the four target identities, with one upright and one inverted block for each identity. Within each block, all distractor faces were of the same sex and in the same orientation as the target images. Subjects responded as quickly and accurately as possible by pressing either the left-arrow key (target present) or the right arrow-key (target absent). They received feedback (a beep) if they responded incorrectly or did not respond within three seconds. Eye movements were explicitly allowed.

The order of blocks was counterbalanced for familiarity and face orientation within each subject. Familiarity always changed from one block to the next, while inversion changed every two blocks. Because of software error, the sex of the targets wasn't counterbalanced across subjects: 12/19 subjects had male targets in the first half of the experiment and female targets in the second half (and the converse for the remaining 7/19 subjects).

Each block started with 24 practice trials followed immediately by 120 test trials. At the beginning of each block, subjects were shown the target identity (upright or inverted) and pressed a key to start the block. On each trial, a central fixation cross appeared for a jittered period between 800-1000 ms, followed by a visual search array of two, four, or six faces displayed for a maximum of three seconds.

Target images appeared in half of the trials. The target was equally likely to appear in the left or right hemifield to avoid possible lateralization biases. Distractor faces were randomly chosen from the set of six distractor identities, and all distractors were different from each other. Stranger target identities never appeared as distractors. Each trial type was repeated 10 times in each block (with distractors randomly sampled every time). Each block thus had 120 trials: 3 (Set Size) x 2 (Target Presence) x 20 (2 different target images x 10 repetitions). The order of the trials within each block was randomized.

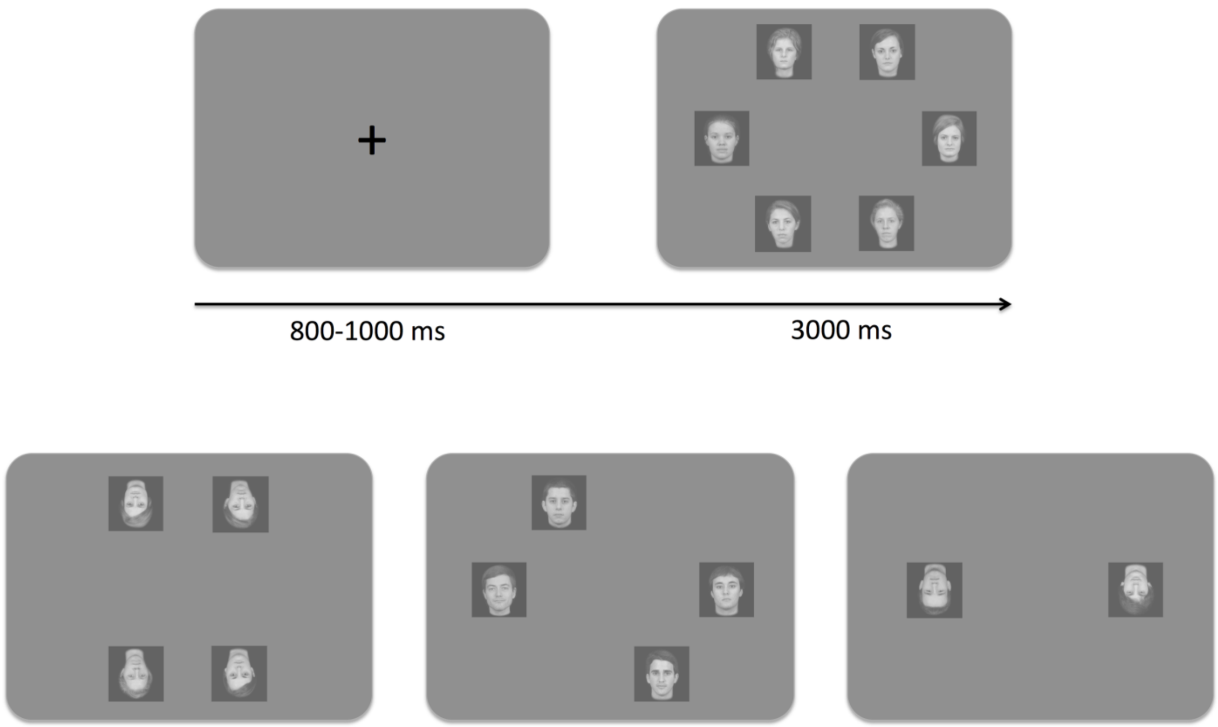

Figure 1. Experimental paradigm and example of the stimuli.

Statistical Analyses

All analyses were run in R (version 3.2.3). To assess statistical significance we fitted Generalized Mixed Models using the package *lme4* (version 1.1-11).

We analyzed subjects' accuracies using Logit Mixed Models (Jaeger, 2008), and reaction times with Linear Mixed Models, separately for target present and target absent trials. For each model, we entered Set Size, Familiarity, and Target Orientation as main effects with all their interactions. The initial random-effect structure contained both subjects and items terms. For the latter term we entered the combination of stimuli appearing on the screen regardless of their position. This allowed us to model the variance due to subject and item (specific images) differences. We also added an extra regressor that indicated the sex of the target, and added random slopes with respect to this term for both subjects and items. We considered this term as a covariate, and thus we didn't analyze it further.

The initial random-effect structure was tested using a log-likelihood ratio test against reduced models (created by removing random slopes first). For the linear models on reaction times, the final structure contained subjects with random slopes and intercepts, and items with random intercepts—the model with random slopes for items failed to converge, thus we used a less complex model. The logit models on accuracies had subjects with random slopes and intercepts only, because the items term did not improve the model fit significantly.

After fitting the models with zero-sum contrasts for the regressors, we tested statistical significance of the fixed-effect terms using a Type 3 analysis of deviance (Wald's χ^2 test), as implemented in the package *car* ((Fox & Weisberg, 2010), version 2.1-1).

For the models on reaction times we log-transformed the independent variable to account for the skewness of the distribution of reaction times; visually inspecting the predicted vs. residual plot confirmed that such a transformation provided a better fit for the model. Finally, the linear model was fitted using restricted maximum likelihood estimation (REML) after determining the best random effect structure using maximum likelihood estimation.

We used a bootstrapping procedure (Efron & Tibshirani, 1994) to investigate the direction of the significant effects found by the models. Trials were always bootstrapped maintaining the structure of the original dataset. For example, for any bootstrap sample the number of trials within each subject and condition (Set Size, Target Presence, Target Orientation, Familiarity, and Target Sex) was preserved, and trials were sampled with replacement only within the appropriate subject and condition. For the next sections, numbers in square brackets represent 95% basic bootstrapped confidence intervals (CI) after 10,000 replications.

We also estimated Set Size 1 intercept and search slopes—which provide information about target-recognition and distractor-rejection processes (Tong & Nakayama 1999)—by fitting a regression line for each subject and condition separately. To obtain 95% confidence intervals we bootstrapped the trials (in a stratified fashion, i.e., maintaining the factorial design of the conditions) and ran the regression model again, repeating this process 10,000 times.

Data Availability

Raw data are available on figshare at the URL

Results

Accuracy

Target Present Trials

Subject responses were overall highly accurate, with average accuracy in target present trials of 93.47% CI: [92.99, 93.95]. We found a significant main effect of set size (χ^2 (2) = 66.95, p < .001) and of target orientation (χ^2 (1) = 17.27, p < .001). Subjects were more accurate when fewer distractors appeared on the screen (one distractor 96.08% [95.43, 96.74]; three distractors 93.71% [92.85, 94.53]; and five distractors 90.60% [89.64, 91.60]), and when faces were presented upright (upright 94.75% [94.12, 95.37]; inverted 92.18% [91.45, 92.93]). The other possible main and interaction effects were not significant (see Supplementary Table 1).

Target Absent Trials

Subject responses were also highly accurate on target absent trials, with average accuracy of 97.69% [97.40, 97.98]. We found a significant main effect of familiarity (χ^2 (1) = 6.60, p < .05) and of target orientation (χ^2 (1) = 15.93, p < .001), but no other significant main or interaction effects (see Supplementary Table 2). Subjects were more accurate at saying the target was absent when looking for a familiar face (familiar 98.10% [97.70, 98.48]; stranger 97.28% [96.84, 97.72]) and when faces were presented upright (upright 98.32% [97.95, 98.68]; inverted 97.06% [96.60, 97.50]), regardless of the number of distractors (two distractors 97.93% [97.47, 98.39]; four distractors 97.92% [97.42, 98.38]; and six distractors 97.22% [96.65, 97.76]).

Reaction Times

Target Present Trials

All main effects of interest were statistically significant: Set Size (χ^2 (2) = 1318.93, $p < .001$), Familiarity (χ^2 (1) = 169.61, $p < .001$), and Target Orientation (χ^2 (1) = 400.49, $p < .001$). We found significant interactions of Set Size x Familiarity (χ^2 (2) = 8.59, $p < .05$) reflecting faster reaction times for familiar face trials; of Familiarity x Target Orientation (χ^2(1) = 9.16, $p < .001$) reflecting a larger familiarity effect for upright faces, and Set Size x Familiarity x Target Orientation (χ^2(2) = 11.17, $p < .001$) reflecting mostly a difference in the effect of familiarity on slopes for upright versus inverted faces (see Supplementary Table 3).

Subjects were overall faster when searching for a familiar face than a stranger face, and they were faster with upright faces than inverted faces (see Figure 2). The advantage for familiar faces was of 113.64 ms [96.97, 130.48] in the upright condition, and of 74.67 ms [54.26, 95.06] in the inverted condition, with a difference of 38.98 ms [12.83, 65.56]). Figure 3 shows the effect size of Familiarity at each set size.

These differences were further analyzed by looking at the estimates of Set Size 1 and slopes. With upright faces, the Set Size 1 estimates were 631.92 ms [615.05, 649.47] for familiar faces, and 682.64 ms [663.00, 701.95] for stranger faces. With inverted faces, they were 699.26 ms [677.14, 722.73] for familiar and 782.75 ms [759.60, 806.17] for stranger faces. We found a trend towards a greater effect of familiarity for inverted faces: 50.72 ms [24.37, 76.13] for upright faces, and 83.49 ms [50.50, 116.24] for inverted faces, but the difference of the effect sizes was not significant (32.76 ms [-8.73, 73.06]).

The significant interaction terms in the linear mixed-effect model reflected differences in the search slopes. Search slope estimates were significantly lower for familiar faces in the upright condition: 86.78 ms/item [79.55, 93.85] vs. 108.31 ms/item [100.67, 116.19] for stranger faces (difference of 21.53 ms/item [11.14, 32.10]). The search slopes for inverted faces were steeper than those for upright faces, and they did not differ across familiarity (familiar faces 120.68 ms/item [111.87, 129.30]; stranger faces 116.32 ms/item [107.18, 125.43]; difference -4.36 ms/item [-17.13, 8.35]).

Target Absent Trials

All main effects of interest were significant: Set Size (χ^2 (2) = 8131.39, $p < .001$), Familiarity (χ^2 (2) = 414.31, $p < .001$), and Target Orientation (χ^2 (2) = 792.64, $p < .001$). The two-way interactions were significant, but not the three-way one: Set Size x Familiarity (χ^2(2) = 6.59, $p < .05$); Set Size x Target Orientation (χ^2(2) =

6.20, $p < .05$); and Familiarity x Target Orientation ($\chi^2(1) = 6.75$ $p < .05$) (see Supplementary Table 4).

The average effect size for Familiarity was 122.43 ms [109.80, 135.17] in the upright condition, and 102.63 ms [86.94, 118.04] in the inverted condition (difference 19.79 ms [-0.08, 40.29]). Figure 3 shows the effect size of Familiarity at each set size.

The search slopes in the target absent trials were about two times those in the target present trials, consistent with a serial self-terminating search. Interestingly, search slopes were steeper when subjects were looking for stranger targets, despite the distractors presented being the same in both familiar and stranger blocks. With upright faces, the search slope was 183.53 ms/item [177.78, 189.30] for familiar targets and 209.98 ms/item [204.40, 215.44] for stranger targets (difference 26.46 ms [18.46, 34.48]). The search slopes were steeper for inverted targets, but less so in familiar than stranger blocks (familiar: 210.06 ms/item [203.58, 216.63]; stranger: 236.98 ms/item [230.52, 243.32]; difference 26.92 ms [17.84, 35.75]).

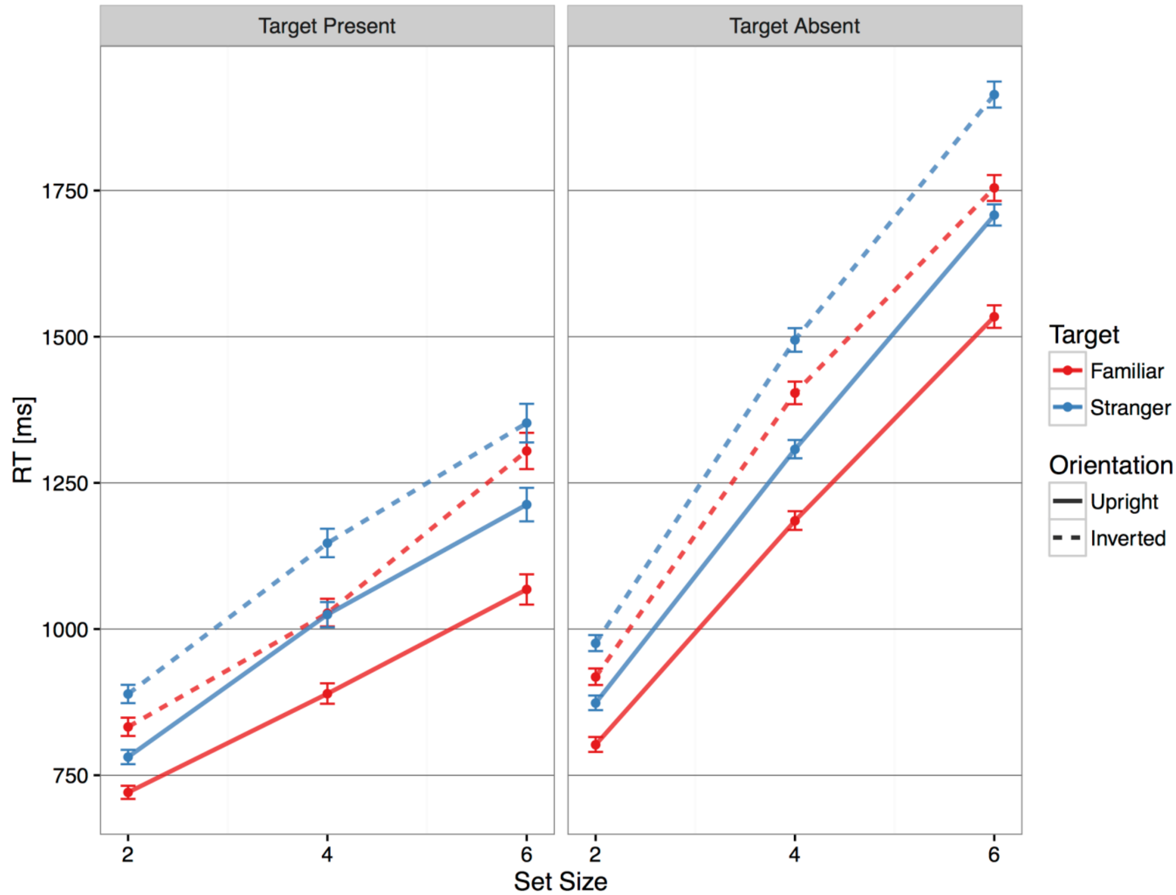

Figure 2. Average reaction times according to presence of the target. Solid lines show upright condition, dashed lines show inverted condition. Red: familiar targets; Blue: stranger targets. Error bars show 95% bootstrapped confidence intervals.

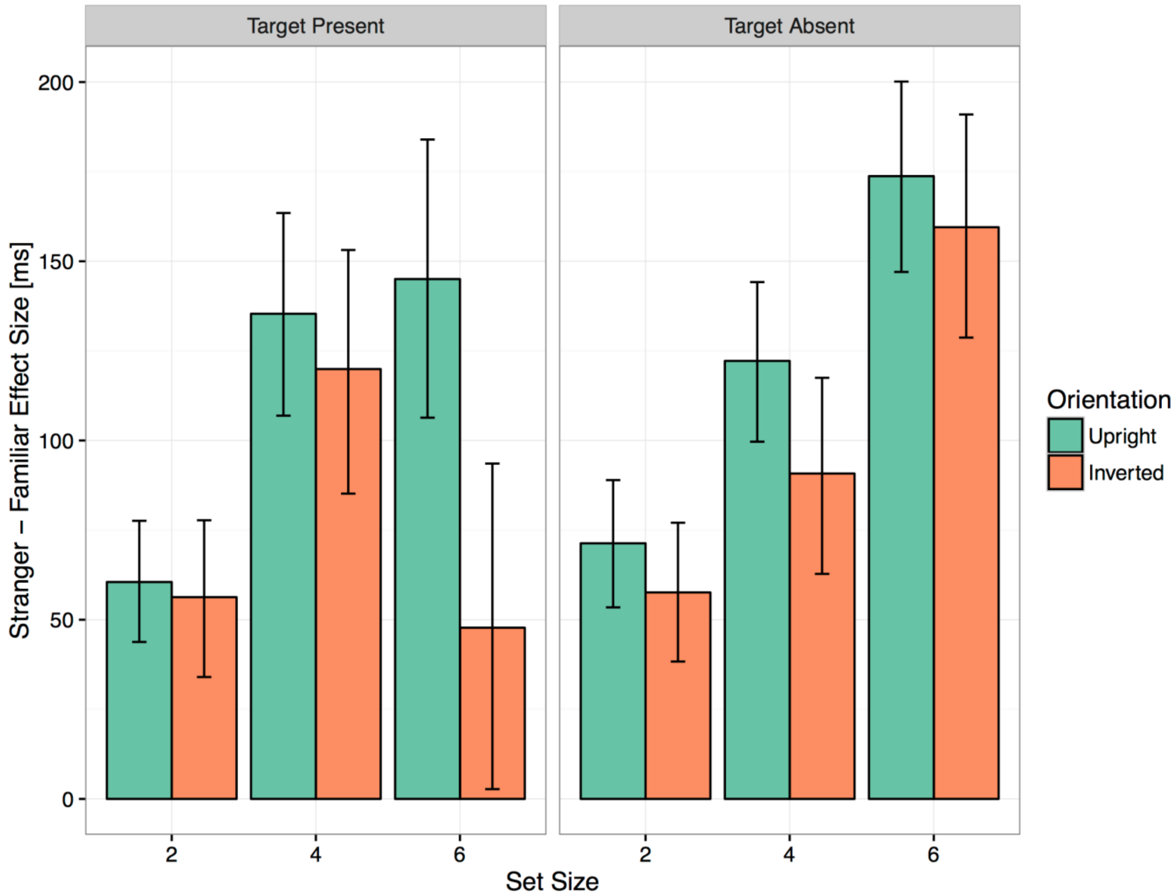

Figure 3. Average effect size of Familiarity for Upright and Inverted faces in Target Present and Absent trials. Error bars show 95% bootstrapped confidence intervals.

Discussion

In this study subjects searched for friends' faces and strangers' faces in a visual search task. We found a processing advantage for personally familiar faces that was robust to face inversion. Subjects' behavior could be framed in terms of a self-terminating serial search, with target-absent search slopes about twice the target-present ones. In target present trials subjects were highly accurate both with familiar and stranger targets, showing no evidence of a speed-accuracy trade-off.

Set Size 1 estimates showed that familiar face targets were processed faster than stranger target faces when presented both upright and inverted. This result adds to the evidence that personally familiar faces enjoy facilitated processing in a variety of experimental conditions (Gobbini et al., 2013; Visconti di Oleggio Castello et al., 2014; Visconti di Oleggio Castello & Gobbini, 2015) and real-life situations (Jenkins & Burton, 2011).

Critically, the advantage of familiar face processing extended to inverted faces. Evidence suggests that turning a face upside-down reduces holistic perceptual processing and favors feature-based processing (Farah et al., 1995; Freire et al., 2000, see also McKone & Yovel, 2009; Valentine, 1988). Thus, the advantage for personally familiar faces in the inverted condition suggests that personal familiarity enhances processing of local facial features.

This finding extended the theoretical relevance of the results by Tong and Nakayama (1999)—who used subjects' own faces as familiar identities. Unlike Tong and Kakayama (1999), though, we used faces of subjects' friends—instead of subjects' own faces. In this way the experimental task was closer to everyday experience. Our stimuli were closer to those to be analyzed when we look for a friend in a crowd of stranger faces. Our results mirror those of Tong and Nakayama (1999) for visual search for one's own face. Unlike Tong and Nakayama (1999), however, who found no effect of the familiarity of the target face (self versus stranger) on distractor rejection rate, we found that stranger faces are rejected more quickly when the target face is personally familiar.

We found that the search slopes differed between familiar and stranger conditions for upright faces on target present trials and for both upright and inverted faces on target absent trials. Slopes did not differ for familiar and stranger conditions for inverted faces on target present trials. These results indicate that subjects were faster at rejecting a stranger distractor when looking for a familiar face target than when looking for a stranger face target, also in target absent trials for both upright and inverted faces. In target absent trials the stimulus arrays were equivalent for familiar target and unfamiliar target blocks, underscoring how the rejection of stranger face distractors varied as a function of the familiarity of the target against which distractors were compared. The increase of the reaction times based on the number of items in the search array is consistent with a serial self-terminating search that was faster for the familiar face targets as compared to the faces of strangers. This indicates that the internal representation of a familiar face, against which each distractor is compared, is either more robust and precise or sparser. We propose that familiarity may direct processing to specific features that are diagnostic of a familiar face's identity, whereas the representation of a stranger's face does not focus processing on similar diagnostic features.

Our hypothesis is plausible because a) changes in eye gaze, a local feature that signals social cues, are detected faster when conveyed by personally familiar faces (Visconti di Oleggio Castello et al., 2014), and b) personally familiar faces are distinguished from stranger faces in a saccadic reaction time task at a latency (180ms) that precedes evoked potentials discriminating familiar from stranger faces (Visconti di Oleggio Castello & Gobbini, 2015). This very rapid detection of familiarity, which precedes the time required to build a view-invariant representation of faces in anterior monkey face patches (Freiwald & Tsao, 2010), further corroborates the hypothesis that rapid familiarity detection is based on a simpler process, perhaps feature-based, consistent also with automatic preattentive detection (Gobbini et al., 2013).

The slightly smaller effect of familiarity on reaction times for inverted faces than for upright faces, as reflected by the significant Familiarity x Orientation interaction, may suggest that some of the features of familiar face representations that afford more rapid processing are holistic. However, the greater magnitude of the familiar advantage even for the inverted faces is consistent with the possibility that this facilitation relies mainly on local features. Additionally, related work indicates that holistic information is less important for recognition of a familiar identity (see Burton, Schweinberger, Jenkins, & Kaufmann, 2015 for a cogent argument).

To sum up, the results of our experiment add to the existing evidence that the human visual system is finely tuned for rapid detection and identification of familiar faces, much more so than of stranger faces. Participants searched for a familiar or stranger identity among distractors presented in either an upright or inverted orientation. They responded faster when searching for familiar faces even in the inverted condition. Taken together, our results suggest that robust representations for familiar faces contain information about idiosyncratic facial features that allow subjects to detect or reject identities when searching for a friend's face in a crowd of stranger faces.

Acknowledgements

We would like to thank Sebastian M. Frank, J. Swaroop Guntupalli, the members of the Gobbini lab and Haxby lab for helpful discussions on these results.

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

Burton, A. M., Schweinberger, S. R., Jenkins, R., & Kaufmann, J. M. (2015). Arguments Against a Configural Processing Account of Familiar Face Recognition. Perspectives on Psychological Science, 10 (4): 482–496.

Efron, B., & Tibshirani, R. J. (1994). An Introduction to the Bootstrap. Taylor & Francis.

Farah, M. J., Tanaka, J. W., & Drain, H. M. (1995). What causes the face inversion effect? Journal of Experimental Psychology. Human Perception and Performance, 21(3), 628–634.

Fox, J., & Weisberg, S. (2010). An R Companion to Applied Regression. SAGE Publications.

Freire, A., Lee, K., & Symons, L. A. (2000). The face-inversion effect as a deficit in the encoding of configural information: direct evidence. Perception, 29 (2), 159–170.

Freiwald, W. A., & Tsao, D. Y. (2010). Functional compartmentalization and viewpoint generalization within the macaque face-processing system. Science, 330(6005), 845–851.

Gobbini, M. I., Gors, J. D., Halchenko, Y. O., Rogers, C., Guntupalli, J. S., Hughes, H., & Cipolli, C. (2013). Prioritized Detection of Personally Familiar Faces. PloS One, 8(6), e66620.

Jaeger, T. F. (2008). Categorical Data Analysis: Away from ANOVAs (transformation or not) and towards Logit Mixed Models. Journal of Memory and Language, 59 (4), 434–446.

Jenkins, R., & Burton, A. M. (2011). Stable face representations. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 366(1571), 1671–1683.

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 & Review, 16(5), 778–797.

Rossion, B. (2009). Distinguishing the cause and the consequences of face inversion: the percpetual field hypothesis. Acta Psychol (Amst.), 132(3):300-312.

Tong, F., & Nakayama, K. (1999). Robust representations for faces: evidence from visual search. Journal of Experimental Psychology. Human Perception and Performance, 25 (4), 1016–1035.

Valentine, T. (1988). Upside-down faces: A review of the effect of inversion upon face recognition. British Journal of Psychology , 79(4), 471–491.

Visconti di Oleggio Castello, M., Guntupalli, J. S., Yang, H., & Gobbini, M. I. (2014). Facilitated detection of social cues conveyed by familiar faces. Frontiers in Human Neuroscience, 8(678), 1–11.

Visconti di Oleggio Castello, M., & Ida Gobbini, M. (2015). Familiar Face Detection in 180ms. PloS One, 10(8), e0136548.

Willenbockel, V., Sadr, J., Fiset, D., Horne, G. O., Gosselin, F., & Tanaka, J. W. (2010). Controlling low-level image properties: the SHINE toolbox. Behavior Research Methods, 42(3), 671–684.

Xu, B., & Tanaka, J. W. (2013). Does face inversion qualitatively change face processing: An eye movement study using a face change detection task. Journal of Vision, 13(2), 22-22.