Familiarity influences visual detection in a task that does not require explicit recognition

Pei-Ling Yang[1] · Diane M. Beck[1,2]

Abstract

The current study aims to explore one factor that likely contributes to these statistical regularities, familiarity. Are highly familiar stimuli perceived more readily? Previous work showing effects of familiarity on perception have used recognition tasks, which arguably tap into post-perceptual processes. Here we use a perceptual task that does not depend on explicit recognition; participants were asked to discriminate whether a rapidly presented image was intact or scrambled. The familiarity level of stimuli was manipulated. Results show that famous or upright orientated logos (Experiments 1 and 2) or faces (Experiment 3) were better discriminated than novel or inverted logos and faces. To further dissociate our task from recognition, we implemented a simple detection task (Experiment 4) and directly compared the intact/scrambled task to a recognition task (Experiment 5) on the same set of faces used in Experiment 3. The fame and orientation familiarity effect were still present in the simple detection task, and the duration needed on the intact/scrambled task was significantly less than the recognition task. We conclude that familiarity effect demonstrated here is not driven by explicit recognition and instead reflects a true perceptual effect.

Keywords Visual perception · Face perception

Introduction

We all agree that our experience not only shapes who we are, but also allows us to make inferences about the future. We learn to extract the relevant information from each event, scene, or situation that we encounter and then use that information to prepare for the next possible encounter. Although everyone agrees that prior knowledge informs our behavior and understanding, it is less widely believed that prior knowledge impacts perceptual processing.

One particularly influential model that predicts prior knowledge should impact perception is Rao and Ballard (1999) hierarchical predictive coding model. In their model, they posit that later regions of the visual pathway generate predictions based on prior knowledge to facilitate the processing of inputs that have been encountered before.

To achieve this facilitation, each level of the brain cascade computes prediction errors by comparing the current signals with the feedback predictions from later regions. Instead of processing and passing the raw input signals, each level of the brain thus only needs to iteratively reduce the prediction errors until the errors are minimized to settle on a representation (Friston, 2005; Rao & Ballard, 1999). Central to these models is a mechanism, usually unspecified, to extract statistical regularities from the world to serve as predictions. This model structure predicts that more experienced inputs, that is, inputs with which we are more familiar, should be better predicted and thus require less processing time than novel inputs. In other words, the visual system should more quickly settle on, and thus perceive, input that is more familiar.

In previous work, our lab developed an intact/scrambled paradigm to assess the effects of real-world statistical regularity on perceptual processing that precedes the need to label or explicitly identify a stimulus (Caddigan et al., 2017; Center et al., 2022; Greene et al., 2015). We use the term *real-world statistical regularity* to capture those regularities that are built up over a lifetime, rather than regularity introduced and learned within the experiment. In the intact/

✉ Pei-Ling Yang
plyang2@illinois.edu

[1] Department of Psychology, University of Illinois Urbana-Champaign, 603 E. Daniel St., Champaign, IL 61820, USA

[2] Beckman Institute, University of Illinois Urbana-Champaign, Champaign, IL, USA

scrambled task, participants simply discriminate whether a target is an intact or a scrambled image (Caddigan et al., 2017; Center et al., 2022; Greene et al., 2015; Smith & Loschky, 2019). The advantage of this task is that it can assess perceptual discrimination sensitivity without introducing the need for participants to explicitly identify what they see—just that they see something coherent rather than noise. In this paradigm the presentation duration is staircased to a threshold, previously between 70% and 82% for each participant that results in a target presentation duration that is often very brief, sometimes at the refresh rate of the monitor. This thresholding procedure means that participants very often are unsure whether an image or scrambled noise was presented, and instead experience a luminance flicker followed by the mask. Crucially though, on some trials participants clearly see an image (Caddigan et al., 2017). As such, the task is a proxy for whether a participant can "see" the image.

Using this intact/scrambled paradigm, researchers have manipulated real-world statistical regularity in a variety of ways and asked whether it impacts how readily participants see a rapidly presented image. Greene et al. (2015) manipulated the probability of natural scenes. Probability, in this case, was defined as "the probability of happening in daily life" and was assessed via four independent and naïve observers. Results showed that probable images were more easily discriminated from fully phased-scrambled images than less probable images. Caddigan et al. (2017) examined the category representativeness of natural scenes. Each scene was rated by separate participants as to how representative it is to its own category. They found that more representative scenes were better discriminated from fully phase-scrambled scenes than less representative scenes. Center et al. (2022) extended the paradigm to isolated objects. They assessed real-world statistical regularity by manipulating the typicality of the viewpoint of an object. Previous research has shown that it is easier to identify canonical views of an object (Palmer et al., 1981). Using the same intact/scrambled task, results showed that typically oriented objects were actually perceive, not just identified, more readily than atypical viewpoints.

Interestingly, in all of these experiments real-world statistical regularity impacted perception despite the fact that the exact stimuli used were unlikely to be previously experienced by the participants. However, given that predictive coding theory suggests that the predictions are based on statistical regularities extracted from the persons own personal experience, familiarity with a particular stimulus should be a strong modulator of perception. Familiarity, here, is defined as frequently encountered in one's own life. We ask whether the intact/scrambled discrimination advantage for statistical regular images holds for images with which we have more experience or are more personally familiar.

Numerous studies have reported familiarity effects on perception, although their status as true perceptual effects is questionable. For instance, words that are more familiar have lower recognition thresholds than non-words (Solomon & Postman, 1952) and letter sequences that are more similar to English grammar require less exposure time to be recognized than random strings of letters (Miller et al., 1954). In the object domain, Gollin (1960) showed that training and familiarization of object images can improve participants' recognition of line drawings in which the line segments delineating the object are fragmented. That is, familiar objects can be recognized with fewer line segments than unfamiliar objects. Critically, however, in these studies and many that followed, participants were asked to recognize objects (e.g., *visual search:* Hershler & Hochstein, 2009; Qin et al., 2014; Shen & Reingold, 2001; *semantic context:* Reingold & Jolicoeur, 1993; Snell & Grainger, 2017; see Baron, 2014, for review; *objects recognition:* Bülthoff & Newell, 2006; Honda et al., 2011; *general review:* Krueger, 1975). Pylyshyn (1999) has argued that recognition, the process interpreting our visual input, is better relegated to the realm of cognition, occurring after an "early-vision stage" in which properties such as color and shape are detected and individuating of object tokens occurs (see also Pylyshyn, 2001). Under this view then, the effects of familiarity assessed by recognition tasks should not be interpreted as evidence of familiarity's effect on perception, or at least not on early vision. This view of early vision as encapsulated from cognition predicts that we should be able to detect the presence of something before we can recognize it.

One study that does explore the effect of familiarity on perception and uses the same intact/scrambled paradigm described above is that of Smith and Loschky (2019). In particular, they investigated whether the expectation of familiar scene sequences can induce the same processing advantage as previous intact/scrambled studies. The sequences were constructed by a series of first-person-view pictures navigating from one location to another location. To manipulate expectation, the sequences could be coherent, in the order experienced as you move through the real world, or randomized. Participants could expect what the next scene should be in coherent sequences but not in the randomized sequences. In essence, they asked whether participants' intact/scrambled judgments were sensitive to expectation set up by the familiar sequence of scenes. Their results showed that target scenes are better detected when they are embedded in coherent familiar sequences than random sequences. Importantly, however, the scenes themselves were familiar in both the coherent and randomized sequence and so their study says less about the effect of familiarity and more about sequence prediction.

The current study aims to compare familiar and completely novel stimuli in the intact/scrambled paradigm to

directly test the effect of familiarity. We predict that participants will better at detecting the presence of familiar stimuli (as opposed to noise) than novel stimuli, indicating that familiarity impacts perceptual processes. In short, the current study uses this intact/scrambled paradigm to ask whether familiar objects are actually perceived, rather than recognized, more readily than unfamiliar ones.

Experiment 1: Intact/scrambled logos

To examine whether familiarity influences perceptual processes, we first compared famous and novel logos in an intact/scrambled discrimination task. Participants were asked to respond whether the target was an intact or a scrambled image under rapid presentation (Caddigan et al., 2017; Greene et al., 2015; Smith & Loschky, 2019). In addition to the fame factor, we also asked whether previous exposure impacted discrimination. In short, participants went through the full set of stimuli twice. We hypothesized that if a single repetition is enough to set up a memory trace, repetition might improve the discriminability in the repeated blocks. It is also possible that if the repetition makes the novel logos more familiar, repetition might reduce the advantage for famous over novel logos. Thus, we can look at two different effects of familiarity; familiarity established within the experiment with repetition, and familiarity established over the course of everyday life. Lastly, to verify that our stimuli were famous to our participants, they were asked to rate how familiar they were with all the famous logos. Novel logos were computer-generated for the experiment, and thus novel to all participants.

Participants

Twenty-six participants (18 females, mean age = 18.9 years) were recruited from the University of Illinois participant pool and were compensated with course credits. This sample size was determined based on the range of sample sizes typically used in prior studies (Caddigan et al., 2017). An a priori power analysis was not conducted, but this sample size is sufficient to detect an effect of dz = .69 (the effect observed by Caddigan et al. (2017) Experiment 1, when comparing d' values for representative and less-representative scenes using a paired t-test) with 92% power (*matched pairs t-test* in G power 3.1.9.4 (Faul et al., 2007)). All participants had self-reported normal or corrected-to-normal vision. Written informed consent was obtained in accordance with procedures and protocols approved by the University of Illinois Institutional Review Board.

Stimuli and procedure

Target images contained full-color famous and computer-generated novel logos (Fig. 1). Novel logos, 12 for the practice experiment, 94 for the staircase experiment, and 101 for the main experiment, were created using the following websites: https://emblemmatic.org/markmaker/#/, https://www.launchaco.com/logo, and https://www.freelogodesign.org/. Words that appeared in the famous logos (e.g., Adidas) were also included in the novel logos. The famous logos were selected based on the rating data of a separate pilot study in which six participants viewed 106 logos for unlimited time and rated them for familiarity on a 7-point scale. Only images with a mean rating greater than or equal to 5 were
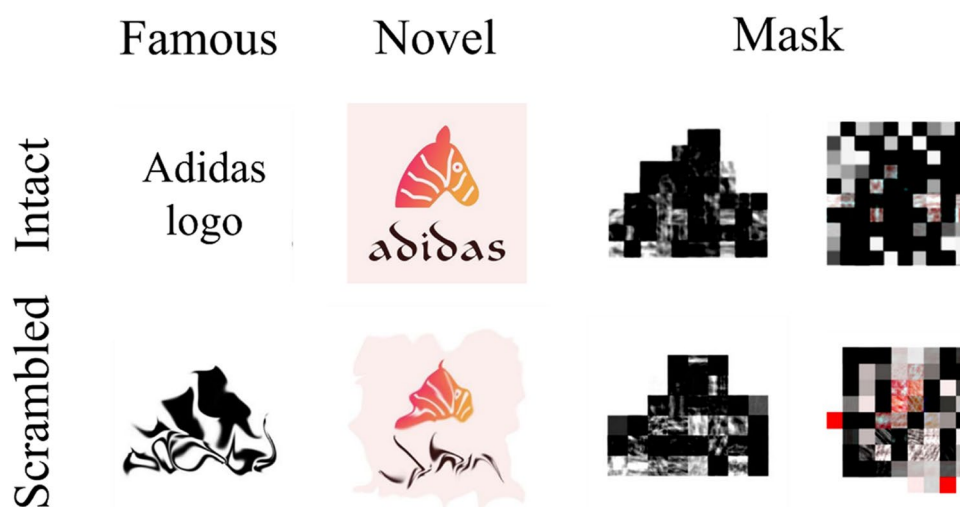

Fig. 1 Example stimuli used in Experiment 1. The first row shows the intact famous and computer-generated novel logos, and their corresponding masks. The famous logo is not depicted here because we do not have permission to use a trademarked image. The second row shows the scrambled form of the intact logos, and their corresponding masks

included in subsequent experiments. Scrambled versions of the famous and novel logos were created using a diffeomorphism with 25% distortion (see Stojanoski & Cusack, 2014) for the mathematical algorithm). The resulting images were no longer recognizable as famous logos but retained many of the same image qualities, including the centralized positioning of the artwork. We created masks by "grid scrambling" both intact and scrambled images. An invisible ten by ten grid was imposed on the images, and the images were phase scrambled within each grid. Hence, each target had its corresponding mask. The intact logos used in the staircase were computer-generated logos. The intact logos used in the practice session were all novel logos. Logos used in both the staircase and the practice sessions only appeared once in this experiment. Logos used in the main experiment were repeated twice, once in an initial experiment and again in a repeated experiment. All the targets and the masks were cropped to the same size of 320 px × 320 px square. Stimuli were presented on an 85-Hz CRT monitor of resolution 1,280 × 960 using the Psychopy package (Peirce et al., 2019) and Python (Python Software Foundation. Python Language Reference, version 3.7). Participants viewed the stimuli with their chin on a chinrest situated 59 cm from the monitor, and thus the images subtended approximately 9.69 degrees of visual angle.

To assess the visual salience of images, we used the saliency toolbox (Walther & Koch, 2006) for MATLAB (version R2021a). The mean and the maximum salience for each image were calculated. Famous logos had larger mean salience than novel logos (Famous (M = .053) > Novel (M = .045), $t(187.8) = 3.31$, $p = .001$, Cohen's d = .47), while novel logos had larger maximum salience than famous logos (Famous (M = 3.15) < Novel (M = 3.20), $t(199.8) = -2.27$, $p = .024$, Cohen's d = .32). Because salience was not perfectly equated between our variables of interest, salience values were included in hierarchical logistic linear models as predictors. In Experiments 3, 4, and 5, we controlled for the visual saliency across stimulus sets.

The experiment had five sessions, all performed within the same hour: practice, staircase, main experiments (initial and repeated), and rating task. For all parts except the rating task, participants performed an intact/scrambled discrimination task. Each trial began with a fixation cross, then a target image (either intact or scrambled) appeared briefly in the middle of the screen followed by a mask (26 frames, 306 ms) (Fig. 2). The duration for the target image was determined for each participant by a staircasing procedure (see below). Participants were asked to respond "intact" or "scrambled" by pressing either the left or right "control" keys on a keyboard, the assignment of which was counter-balanced across participants. Participants were instructed to respond as fast as possible without sacrificing accuracy. The trial ended if participants did not make a response within 136 frames (1.6 s) after the onset of the mask.

Practice: Each participant first completed 24 trials with the target duration set at 118 ms to familiarize themselves with the task. They received feedback on these practice trials; the word "incorrect" appeared in red in the middle of the screen along with a beep sound (100-Hz tone) for incorrect responses, whereas a black "correct" (and no sound) appeared for correct responses. Feedback was used only in the practice session.

Staircasing: Because we have previously found that individuals vary greatly on this intact/scrambled task, duration was staircased for each individual. We used the Quest algorithm (Watson & Pelli, 1983) with a 71% accuracy threshold to estimate presentation duration, and the resulting duration was used in the main and repeated-main experiments. The staircase procedure contained 188 trials and the possible durations were set to range from one to 21 frames at 11.8 ms per frame.
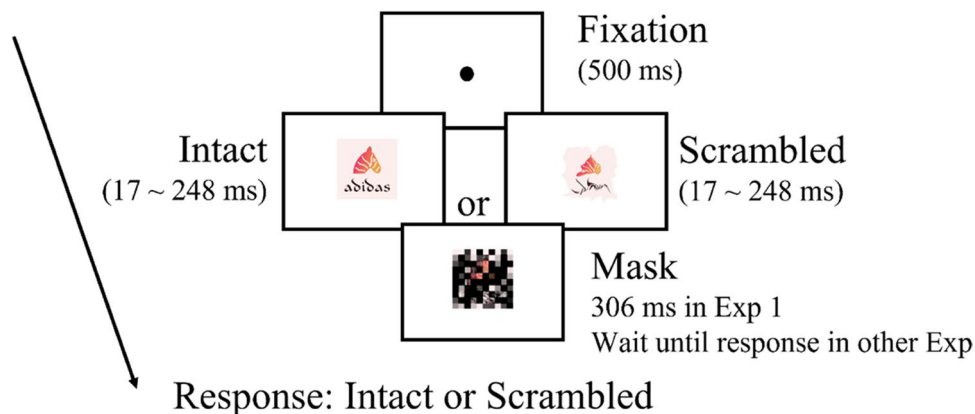

Fig. 2 The procedure for the intact/scrambled task. The target was presented in the middle of the screen and followed by a mask

Main experiments: Because we were interested in repetition, each participant initially completed 404 trials of 404 unique stimuli (no repeats). We refer to this phase of the experiment as the main experiment. Immediately after completing the first repetition, they completed a second repetition of the task with all the 404 stimuli, but in a different order. We refer to this condition as the repeated main experiment.

Rating: After completing the repeated main experiment, participants rated how familiar they were with each famous logo on a 7-point scale, 1 (never seen this logo) to 7 (very familiar).

Results

The durations that were obtained during staircasing ranged from one frame (~ 12 ms) to 20 frames (~ 235 ms) across participants, with a mean of 9.76 frames (~ 115 ms) and a standard error of 1.48 frames (~ 17 ms). For the main experiment, we first excluded trials in which participants did not respond. Most participants missed fewer than 2% of the trials in each condition. The highest missing rate was 23% of trials in the novel intact repeated condition and this only happened in one participant. In the context of signal detection theory, intact images were viewed as signal present while scrambled images were considered as signal absent. Hit rates, therefore, were defined as 'intact' responses when the targets were indeed intact. False alarm rates were defined as 'intact' responses when the targets were scrambled. A sensitivity measure, d', was calculated for both the famous and novel logo conditions in the main experiments. For both the initial and repeated phases of the main experiment, we observed higher d' for famous (Initial: M = 2.54, SE = 0.23; Repeated: M = 2.49, SE = 0.21) than novel logos (Initial: M = 2.31, SE = 0.26; Repeated: M = 2.23, SE = 0.25) (Fig. 3). A two-way repeated measures ANOVA (*afex* package (Singmann et al., 2021) in R version 3.6.3 (R Core Team, 2020)) of familiarity (famous vs. novel) and repetition (initial vs repeated) revealed a significant main effect of familiarity ($F(1, 25) = 18.38$, $p < .001$, partial $\eta^2 = .42$), but no significant main effect of repetition ($F(1, 25) = .07$, $p = .79$, partial $\eta^2 = .002$) nor an interaction with repetition ($F(1, 25) = .06$, $p = .81$, partial $\eta^2 = .002$).

To better understand the factors contributing to higher sensitivity, we compared both hit rates and false alarms as a function of fame and repetition (Table 1). We observed higher hit rates for famous logos than novel logos (Fame: $F(1, 25) = 52.00$, $p < .001$, partial $\eta^2 = .68$) and no effect of repetition ($F(1, 25) = .23$, $p = .64$, partial $\eta^2 = .01$) nor interaction with repetition ($F(1, 25) = .28$, $p = .60$, partial $\eta^2 = .01$). We also observed higher false alarm rates for the scrambled versions of the famous logos than novel logos (Fame: $F(1, 25) = 41.11$, $p < .001$, partial $\eta^2 = .62$). We suspect that the difference in false alarm rates between
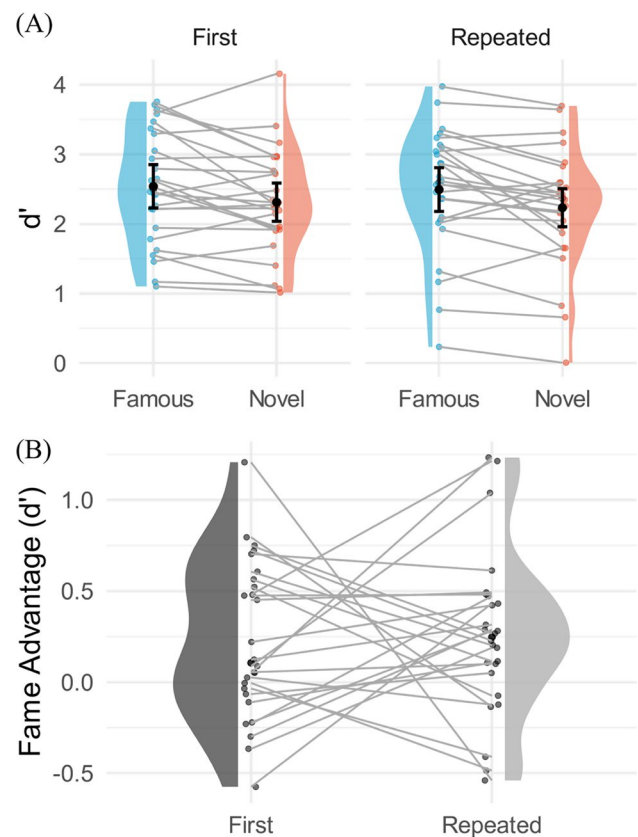

Fig. 3 **A** Within-subject comparison violin plots. Each dot represents one participant, and the performance of each participant is connected with a line. The distribution shows the density function of the d' distributions. The bar superimposed on the dots represents the 95% within-subject confidence interval based on the Cosineau-Morey-O'Brien method (Cousineau & O'Brien, 2014). The left-hand plots show the data from the initial presentation, and the right-hand plots show data from the repeated presentation. Blue represents famous logos, and red represents the novel logos. The main effect of fame is significant, but the main effect of repetition and the interaction effect are not. **B** The fame advantage of was calculated as the d' of the fame condition subtracts the d' of the novel condition. Each dot represents one participant, and the performance of each participant is connected with a line

famous and novel logos are due to low-level features that were retained in diffeomorphed famous logos (e.g., colors) causing participants to guess intact. There were no effects of repetition ($F(1, 25) = .01$, $p = .92$, partial $\eta^2 = .00$) nor an interaction with repetition for false alarms $F(1, 25) = .01$, $p = .92$, partial $\eta^2 = .00$). Higher hit rates and false alarm rates resulted in significant difference in bias between famous and novel logos ($F(1, 25) = 108.6$, $p < .001$, partial $\eta^2 = .81$).

There was no evidence of a speed/accuracy trade-off in these experiments (see Table 1); participants responded significantly faster to famous intact logos than to novel intact logos in initial and repeated presentations (Initial: 559 ms (SE = 21) vs. 593 ms (SE = 22), Repeated: 539 ms (SE = 17) vs. 557 ms (SE = 17); Fame: $F(1, 25) = 53.54$, $p < .001$,

Table 1 Summary of participant performance in Experiments 1, 2, and 3

	d' (sensitivity)	Hit rate	FA rate	Bias	RT (ms)
Exp. 1 (Logo in-lab)					
Initial - Famous	2.54 ± 0.16	0.90 ± 0.02	0.19 ± 0.04	-0.21 ± 0.09	559 ± 21
Initial - Novel	2.31 ± 0.15	0.82 ± 0.03	0.15 ± 0.03	0.12 ± 0.10	593 ± 22
Repeat - Famous	2.49 ± 0.17	0.89 ± 0.02	0.18 ± 0.04	-0.13 ± 0.09	539 ± 17
Repeat - Novel	2.23 ± 0.17	0.79 ± 0.03	0.14 ± 0.03	0.18 ± 0.10	557 ± 17
Exp. 2 (Logo online)					
Famous upright	1.70 ± 0.14	0.84 ± 0.02	0.34 ± 0.03	-0.32 ± 0.06	589 ± 41
Famous inverted	1.60 ± 0.12	0.82 ± 0.02	0.31 ± 0.03	-0.21 ± 0.06	602 ± 26
Novel upright	1.43 ± 0.11	0.72 ± 0.02	0.27 ± 0.03	0.04 ± 0.07	614 ± 22
Novel inverted	1.35 ± 0.11	0.69 ± 0.02	0.26 ± 0.03	0.11 ± 0.07	657 ± 53
Exp. 3 (Face online)					
Famous upright	1.88 ± 0.11	0.78 ± 0.02	0.21 ± 0.02	0.02 ± 0.06	584 ± 23
Famous inverted	1.60 ± 0.10	0.68 ± 0.02	0.20 ± 0.02	0.22 ± 0.06	618 ± 31
Novel upright	1.73 ± 0.10	0.75 ± 0.02	0.22 ± 0.02	0.07 ± 0.07	596 ± 28
Novel inverted	1.58 ± 0.10	0.66 ± 0.02	0.20 ± 0.02	0.26 ± 0.06	682 ± 66

Each cell represents mean ± standard error. The bias was calculated as -0.5*(Z(hit) + Z(FA)). Response time was calculated only for intact trials

partial η^2 = .68; Interaction: $F(1, 25)$ = 3.07, p = .092, partial η^2 = .11). That is, intact famous logos were detected both more accurately and more quickly than intact novel logos.

In follow-up analyses, we included salience in our statistical model to determine whether salience differences between intact famous and novel logos might be the cause of our familiarity effect. A random-intercept logistic hierarchical linear model was fitted to the accuracy (transformed into a continuous measure using a log of the odds ratio) of each intact trial (*lme4* package (Bates et al., 2015) in R version 4.1.2 (R Core Team, 2021)). Instead of using d' estimate, which aggregates across trials, accuracy on intact trials was chosen as the dependent variable so that we could model the salience of each image on each trial. This model contained five fixed-effect factors (fame, repetition, the interaction of fame and repetition, and the mean and maximum salience of each stimulus) and random intercept for each participant. Results showed that intercept, fame and repetition factors were the only significant predictors of accuracy in this model (Intercept: β = 2.09, SE = 0.90, Z = 2.31, p = .021; Fame: β = 0.75, SE = 0.08, Z = 9.16, p < .001; Repetition: β = -0.22, SE = 0.07, Z = -3.17, p = .002). The negative beta weight associated with repetitions indicates that participants were actually worse at discriminating the intact logos the second time they were encountered. This is more likely due to fatigue than actual logo repetition. Importantly, not only did salience not pick up significant variance, but the fame effect persisted when salience was included in the model, indicating that the familiarity effect cannot be attributed to low-level salience differences. Finally, the rating task confirmed that participants found our famous

logos familiar with average rating of 6.57 on a 7-point Likert scale (7 is very familiar and 1 is have never seen this logo).

Discussion

In Experiment 1, two aspects of familiarity were tested: familiarity built up through repeated exposure in our everyday lives (famous) and familiarity established within the experiment (repetition). An intact/scrambled task was used to access the sensitivity of participants to the presence of coherent (intact) stimuli under rapid presentation. Our data suggested that famous logos were better perceived than novel logos. A single repetition within the context of the experiment was not sufficient to impact sensitivity. This influence of familiarity then may require more frequent exposures than a single repetition within the experiment. We note, however, that in this case we were repeating images that participants may not have perceived clearly due to the brief and masked presentations, potentially preventing the formation of a familiarity signal. Future work is needed to determine whether exposure to a single clearly perceived exposure might impact the intact/scrambled discrimination.

In the next experiment, we use the same set of stimuli to replicate the effect of fame (the stronger of the two effects from Experiment 1) and also introduce another source of familiarity, orientation. We assume that subjects have more experience with upright logos than inverted ones, and thus inverted logos would be less familiar than upright ones.

Experiment 2: Intact/scrambled logos (online)

In Experiment 1, we demonstrated an effect of familiarity on perception (the detection of an intact as opposed to scrambled image) when comparing famous and novel logos. In Experiment 2, we asked whether this familiarity effect can be replicated, and further explored the effect of upright (familiar) versus inverted (unfamiliar) orientation of the logo. Specifically, we picture-plane rotated the logos 180°. The same intact/scrambled task and procedure as Experiment 1 was used. Experiment 2 and the following experiments were conducted online due to the COVID-19 pandemic. We note that we have less precise control of the duration of the stimuli in an online study than we do in the laboratory, due to differences in monitors and internet bandwidth and reliability. To tackle these limitations, we loaded the stimuli when the experiment was first initialized, we excluded participants based on dropped frames (see exclusion criteria below for description of how dropped frames were calculated) we used the within-subject design such that any idiosyncrasies of the monitor are shared across conditions (e.g., decay rate of monitors), and we increased the number of participants to offset the variability in the online studies. Furthermore, the performance variability induced by any participant's computer or internet should be viewed as variability that, if anything, reduces the effect size. Thus, if we find a significant effect of familiarity in the online studies, we can infer we would get similar or larger effects with better timing control. This experiment was pre-registered on the Open Science Framework (OSF) website (https://osf.io/mnpw6).

Participants

An online pilot study was conducted to estimate the required sample size needed give us 80% power to detect a familiarity effect. This pilot has the same design as the main experiment. Forty-one participants were recruited from the University of Illinois online participant pool and twenty-two of them passed the following exclusion criteria: (1) the accuracy of the last two blocks of the staircase was lower than 60% (chance is 50%), suggesting that staircase was not converging on an accurate duration; (2) the overall accuracy of the main task was lower than chance , 50%; (3) the dropped frame rate (due to the online nature of the task) in the staircase was higher than 5%, as such errors interfered with the accuracy of the staircase. A frame is marked as a dropped frame if the difference of

timestamps between two trials is smaller than 17 ms (the rounded-up refresh rate of 60 Hz monitor). The dropped frame rate then is calculated as the dropped frames counts over the total number of trials in staircase session.

A power analysis was conducted on the pilot data to estimate the needed sample size for the main experiment. The pre-registered sample size for detecting the main effect of fame was 24 and of orientation was 54 with 80% power. We recruited until we obtained data from 54 participants (38 females, mean age = 24.4 years) who met inclusion criteria (76 participants were recruited in total). Unfortunately, the sample size reported in the pre-registration form was incorrect due to an error in calculating the effect size that underestimated the actual effect size (for more information, please refer to the change of pre-registration form: https://osf.io/mnpw6). Using the correct effect size estimates for the power analysis, we needed 16 participants to detect, with 80% power, the main effect of fame, and 18 participants to detect the main effect of orientation for the repeated measure ANOVA (G power 3.1.9.4, Faul et al., 2007). With the 54 participants we actually recruited, we had 99% power to detect the effects of fame and orientation. Participants were recruited from the paid participant recruiting website, Prolific.org. All participants had self-reported normal or corrected-to-normal vision. Participants agreed to an online informed consent approved by the University of Illinois Institutional Review Board and presented on the screen before the start of the experiment.

Stimuli and procedure

Target images were the same set of images as those used in Experiment 1, with one exception—one vertically symmetric novel logo, which would be unchanged by our orientation manipulation, was substituted with another asymmetric novel logo. To introduce the orientation factor, the images were inverted by rotating the upright images by 180°. In this experiment, we had the following numbers of trials: 101 intact upright famous images for the main task, 60 intact upright images for the staircase session, and four intact upright images for the practice session. Overall, including scrambled (same diffeomorphed images) and inverted (180° picture-plane rotation) images, we had 808 trials for the main task, 480 trials for the staircase, and 32 trials for the practice. The same "grid-scrambled" masks were generated for all the intact logos. The intact logo and its corresponding scrambled logo (i.e., diffeomorphed intact logo) shared the same mask (generated from the intact logo). Stimuli were presented on participants' own monitor using javascript and the *jsPsych* package (De Leeuw, 2015, version 6.1.0). To control the size of the images that appeared on participants' screen, we asked participants to adjust a box on the screen to

be the same size as a credit card (width = 3.375 in.; height = 2.125 in.) held up to the screen.

As in Experiment 1, Experiment 2 contained four sessions, all performed within the same hour: practice, staircase, main task, and rating. The stimuli were the same as in Experiment 1 with four exceptions: (1) 180-degree picture-plane rotated images were included, (2) all the stimuli only appeared once before the rating session, (3) the same masks were used for each pair of intact and scrambled logos, and (4) the masks remained onscreen until response.

The same procedures as Experiment 1 were used apart from the following modifications:

> *Practice:* No audio feedback was given and instead of 24 trials, each participant completed 16 practice trials in total.
> *Staircasing:* This experiment had 240 trials in total, and we used a custom javascript version of Quest adapted from the MATLAB code in Psychtoolbox3 (Watson & Pelli, 1983; King-Smith et al., 1994). This version of Quest required a mean and standard deviation, rather than a range, to define the starting distribution. We used a starting distribution with a mean of 134 ms and standard deviation of 134 ms for this Quest algorithm, resulting in similar range of 1–21 frames as in Experiment 1.
> *Main experiments:* The orientation factor replaced the repetition factor and the experiment had 808 trials in total.
> *Rating:* We included two novel logos as catch trials. We expected lower ratings on these trials than for the famous logos.

Results

The duration obtained from the staircasing at 71% threshold ranged from 17.0 ms to 528 ms across participants, with a mean of 59 ms and a standard error of 12 ms. As in Experiment 1, d' was calculated for each condition (Famous upright: M = 1.69, SE = .14; Famous inverted: M = 1.60, SE = .12; Novel upright: M = 1.43, SE = .11; Novel inverted: M = 1.35, SE = .11) (Fig. 4). Further, a 2 (fame: famous / novel) x 2 (upright / inverted) factor repeated-measures ANOVA was conducted on the d' of each condition (*afex* package (Singmann et al., 2021) in R version 4.1.2 (R Core Team, 2021)). The main effects of fame and orientation were significant (Fame: $F(1, 53) = 39.74$, $p < .001$, partial $\eta^2 = .43$; Orientation: $F(1, 53) = 7.41$, $p = .009$, partial $\eta^2 = .12$) but no significant interaction effect was found ($F(1, 53) = 0.06$, $p = .80$, partial $\eta^2 = .00$).

Further analyses showed that both hit rate and false alarm rate contributed to the d' differences in fame and orientation. Specifically, hit rates were significantly higher when the target was famous or upright than when the target was novel or
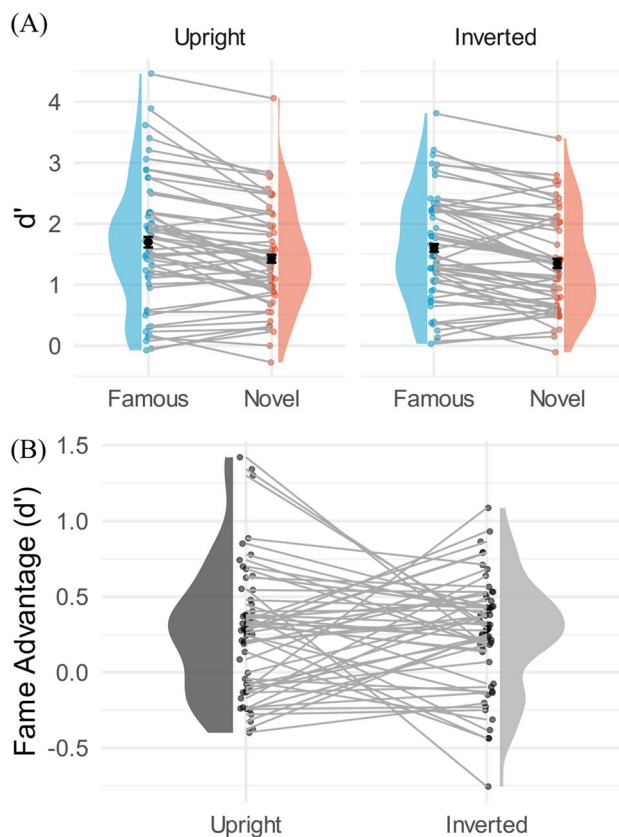

Fig. 4 A Within-subject comparison violin plots for Experiment 2. Y-axis is the d' measurements. The bar superimposed on the dots represents the 95% within-subject confidence interval based on the Cosineau-Morey-O'Brien method (Cousineau & O'Brien, 2014). The left-hand plots show the upright logos, and the right-hand plots show the inverted logos. Blue represents famous logos, and red represents the novel logos. The main effects of fame and orientation were significant, but no significant interaction was found. **B** The fame advantage in d', shown on the y-axis, was computed by subtracting d' of the novel condition from that of the famous condition. Each dot represents one participant, and each participant is connected with lines

inverted (Hit rate: Fame: $F(1, 53) = 76.1$, $p < .001$, partial $\eta^2 = .59$; Orientation: $F(1, 53) = 22.1$, $p < .001$, partial $\eta^2 = .29$; Interaction: $F(1, 53) = 1.25$, $p = .27$, partial $\eta^2 = .02$). Interestingly, the same was true for false alarm rates (Fame: $F(1, 53) = 25.7$, $p < .001$, partial $\eta^2 = .33$; Orientation: $F(1, 53) = 8.27$, $p = .006$, partial $\eta^2 = .13$; Interaction: $F(1, 53) = 2.09$, $p = .15$, partial $\eta^2 = .04$), indicating that participants were more likely to false alarm to scrambled stimuli made from the famous and upright logos than their unfamiliar counterparts. The diffeomorphic technique tends to preserve not only the color pattern of the intact logos but also some of the composition of the original logo. Hence, it may be that the familiar low-level features caused participants to guess that the logo was intact. We note, however, that this result does not undermine the d' result, as the increased sensitivity means that hits rates outweigh the increase in false alarms.

Higher hit rates and false alarm rates similar to Experiment 1 translated to a significant difference in bias estimates for the fame and orientation factors (Fame: $F(1, 53) = 82.7$, $p < .001$, partial $\eta^2 = .61$; Orientation: $F(1, 53) = 24.9$, $p < .001$, partial $\eta^2 = .32$).

Again, there was no evidence of a speed/accuracy trade-off. As in Experiment 1, participants had faster reaction times to detect famous than novel intact logos, although this difference did not reach significance ($F(1, 53) = 1.64$, $p = .21$, partial $\eta^2 = .03$). Similarly, participants were faster for upright than inverted intact logos, but again this difference did not reach significance (Orientation: $F(1, 53) = 1.08$, $p = .30$, partial $\eta^2 = .02$), nor did the Interaction: $F(1, 53) = .24$, $p = .63$, partial $\eta^2 = .004$). However, these experiments were not designed to find RT effects, and rather the above pattern of results indicates that no speed-accuracy tradeoff was observed.

To rule out the influence of low-level salience differences between famous and novel logos, we fitted a logistic hierarchical regression model to the accuracy of intact trials (*lme4* package, Bates et al., 2015; in R version 4.1.2, R Core Team, 2021). This model had five fixed effect factors including fame, orientation, the interaction between fame and orientation, mean salience and maximum salience of each stimulus and random intercept for each participant. Results showed that only intercept, fame, orientation, and the mean salience were significant predictors of accuracy for intact trials (Intercept: $\beta = 1.59$, SE = 0.58, $Z = 2.72$, $p = .007$; Fame: $\beta = 0.83$, SE = 0.05, $Z = 16.86$, $p < .001$; Orientation: $\beta = 0.19$, SE = 0.04, $Z = 4.29$, $p < .001$; Salience average: $\beta = -3.29$, SE = 1.54, $Z = -2.14$, $p = .03$). The significant fame and orientation predictors, with salience measures in the model, indicate that the familiarity effects cannot be solely explained by salience.

For the familiarity ratings, the average rating of famous logos was 6.41 out of the same 7-point Likert scale as Experiment 1, while the average rating of the two novel logos, the catch trials, was 2.54 on the same scale. This result confirmed that the famous logos used in the experiment were indeed more familiar than the novel logos.

Discussion

Experiment 2 manipulated familiarity via fame and orientation using the same intact/scramble task as Experiment 1. The results replicated the fame effect of Experiment 1, with participants perceiving famous logos better than novel logos. In addition, we also demonstrated a familiarity effect using orientation, such that upright logos were better perceived than the inverted logos. The significant affect orientation differed from that of Center and colleagues (Center et al., 2022). Although they found a significant effect of atypical viewpoints when the objects were rotated in depth, they also explore atypical viewpoints by rotating the objects in the picture plane. In contrast to the results here, they found no significant difference in detection between typical and picture-plane viewpoint objects (Center et al., 2022). This discrepancy cannot be due to famousness since orientation did not interact with fame; that is, the orientation effect did not vary with fame. Instead, a likely source of the discrepancy is the presence of words (i.e., company names) in many of our logos; Center and colleagues objects did not contain letters. Upright letters can also be viewed under the umbrella of a familiarity effect since we are more likely to view upright letters than inverted letters in daily life.

Before making strong conclusions that familiarity impacts perception, we should extend these results to another image domain. There are many possible ways in which famous and novel logos may differ, especially given that famous logos were designed by graphic artists and novel logos were computer generated. For example, graphic artists design logos to be eye catching whereas the automatic logo generating websites appear to be less good at capturing the eye-catching nature of artist designed logos. Thus, it is possible that the observed fame effect is not due to familiarity per se, but to factors employed by graphic artists that are not adequately captured by the automatic logo-generating websites.

Experiment 3: Intact/scrambled faces (online)

In Experiment 3, we extended the familiarity factor to faces (famous or novel). Faces are not only more natural (i.e., not artificially designed) than logos, but they are much more similar to each other than logos are. Famous and novel faces differ only in their familiarity. Moreover, one could argue that we are all experts at faces. Thus, better detection of famous faces would provide a strong test of our hypothesis that familiarity affects perceptual processes that precede explicit recognition.

Participants

A pilot study of 36 participants was run to estimate the required sample size. Seventy-one participants were initially recruited from the University of Illinois online participant pool, but only 36 of them passed our exclusion criteria, which were the same as in Experiment 2. A repeated-measures ANOVA (fame × orientation) conducted on sensitivity revealed main effects of fame ($F(1, 35) = 5.23$, $p = .028$, partial $\eta^2 = .13$) and orientation ($F(1, 35) = 23.56$, $p < .001$, partial $\eta^2 = .40$) but no interaction of fame and orientation ($F(1, 35) = .032$, $p = .86$, partial $\eta^2 = .00$). Using the effect sizes of this pilot, we conducted a power analysis using the repeated ANOVA function in G power 3.1.9.4, which

indicated we needed 18 participants to detect the main effect of orientation and 75 participants to detect the main effect of fame at 80% power. We recruited until 75 participants (45 females, mean age = 23.1 years) met inclusion criteria (107 participants were recruited in total). Participants were recruited from either the University of Illinois online participant pool (18 participants, compensated with course credits) or from the paid participant recruiting website, Prolific.org (89 participants, compensated with money). All participants had self-reported normal or corrected-to-normal vision. Participants agreed to a written informed consent approved by the University of Illinois Institutional Review Board and presented on the screen before the start of the experiment.

Stimuli and procedure

Stimuli in the main task consisted of color pictures of 90 famous people and 90 computer-generated novel faces (Fig. 5). Novel faces were generated by a Generative adversarial network (GAN) model implemented on the website, thispersondoesnotexist.com (Karras et al., 2020) and were selected by the first author to be visually matched for age range, sex, and race of the famous faces.

The same orientation factor as Experiment 2 was included in this experiment. The scrambled images were created through the same diffeomorphic technique as in previous experiments. The masks were generated through total phase-scrambling of the intact images. For the staircase session, 52 pictures of less well-known stars were used, saving our most famous faces for the main experiment. For the practice session, 4 novel faces were used. No face with the same fame and orientation factor was repeated, prior to the rating experiment. In this experiment, intact and scrambled images shared the same mask to eliminate the possibility of making the judgment based on the masks. All the images were 320 px × 320 px. The same credit card method was used to control the size of the stimuli presented on participants' screens, resulting in a visual angle of approximately 5.2° × 5.2° assuming an approximate distance of 60 cm from the computer screen. The program was written in javascript and using the *jsPsych* package (de Leeuw, 2015, version 6.1.0), and hosted on Pavlovia.org.

The task and procedure were the same as Experiment 2 except that in this case the stimuli were faces and the experiment had slightly fewer trials.
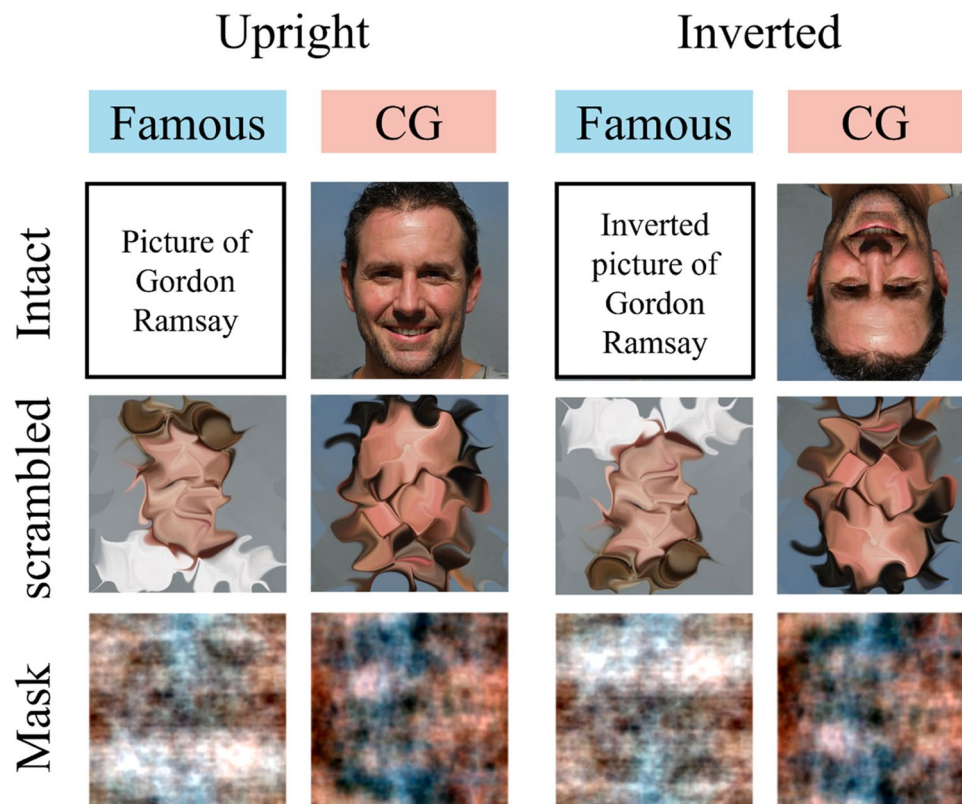

Fig. 5 Example stimuli used in experiment 3. The first row shows the upright or inverted intact famous and computer-generated novel faces. The famous person (Gordon Ramsey) is not depicted here because we do not have permission to use his image. The second row shows the scrambled form of the intact faces as scrambled images. The last row shows the phase-scrambled form of the intact faces as masks

Practice: Practice had 16 trials in total including four upright novel faces and their inverted and diffeomorphed versions.

Staircasing: The staircase had 208 trials in total including 52 upright less famous faces and the inverted and diffeomorphed versions.

Main experiment: The main experiment had 720 trials including 90 famous upright faces and 90 novel upright faces, and their inverted and diffeomorphed version.

Rating: Rating had the same ninety famous upright faces from the main experiment and two additional novel faces were included as catch trials. Like in Experiment 2, we expect the novel faces to have lower familiarity ratings.

The mean and maximum salience, computed using the saliency toolbox (Walther & Koch, 2006) for MATLAB (version R2021a), were not significantly different between famous and novel faces (Mean: Famous (M = 0.053) vs. Novel (M = 0.053), $t(177.7) = -0.23$, $p = 0.82$; Maximum: Famous (M = 3.17) vs. Novel (M = 3.15), $t(177.8) = 0.55$, $p = 0.58$), suggesting that differences observed in participants intact/scrambled performance cannot be explained by the visual salience.

Results

The duration obtained from staircasing at 71% accuracy ranged from 17 ms to 277 ms across participants, with a mean of 35 ms and a standard error of 5 ms. A two-way repeated-measures ANOVA (*afex* package (Singmann et al., 2021) in R version 4.1.2 (R Core Team, 2021)) was conducted on d' with fame (famous vs. novel) and orientation (upright vs. inverted) as factors (Famous upright: M = 1.88, SE = 0.11; Famous inverted: M = 1.60, SE = 0.10; Novel upright: M = 1.73, SE = 0.10; Novel inverted: M = 1.58, SE = 0.10) (Fig. 6). Results showed a significant main effect of fame ($F(1, 74) = 10.45$, $p = .002$, partial $\eta^2 = .12$), a significant main effect of orientation ($F(1, 74) = 10.44$, $p = .002$, partial $\eta^2 = .12$), and a significant interaction of fame and orientation ($F(1, 74) = 4.77$, $p = .032$, partial $\eta^2 = .06$). To follow up the interaction, a Tukey HSD test (*emmeans* package (Lenth, 2022) in R version 4.1.2 (R Core Team, 2021)) was conducted comparing upright famous and novel faces and comparing inverted famous and novel faces. The fame effect was significant for upright faces ($t(74) = 3.91$, $p = .001$, SE = .039) but not for inverted faces ($t(74) = .71$, $p = .89$, SE = .040). This interaction is likely explained by the fact that face recognition is impaired by inversion (see Valentine, 1988, for review), rendering famous faces less familiar. Indeed, the familiarity of inverted faces is delayed relative to upright faces (Marzi & Viggiano, 2007).

To determine whether the d' results reflected differences in hit rate, false alarms, or both, follow up ANOVAs of hit
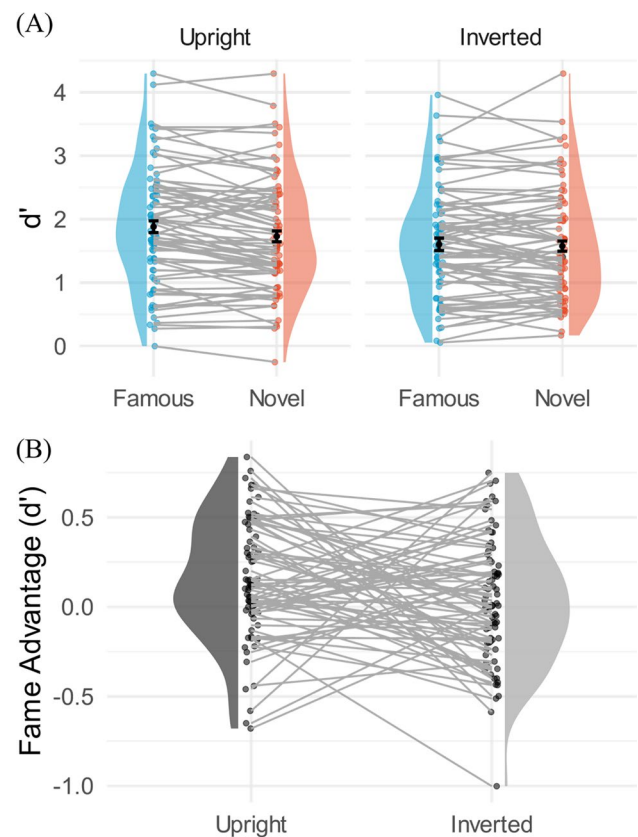

Fig. 6 A Within-subject comparison violin plots for the d' estimates in Experiment 3. The superimposed error bar represents the 95% within-subject confidence interval based on the Cousineau-Morey-O'Brien method (Cousineau & O'Brien, 2014). The left-hand plots show the upright faces, and the right-hand plots show the inverted faces. Blue represents famous faces. Red represents novel faces. The main effects of fame and orientation, as well as their interaction were significant. **B** The d' fame advantage is calculated by subtracting the d' for the novel condition from the d' for the famous condition, just as it was in Experiment 2. Each dot represents one participant, and each participant is connected with a line

rate and false alarms were conducted. Hit rates significantly differed in both fame and orientation factors (Fame: $F(1, 74) = 19.54$, $p < .001$, partial $\eta^2 = .21$; Orientation: $F(1, 74) = 24.66$, $p < .001$, partial $\eta^2 = .25$; Interaction: $F(1, 74) = 2.55$, $p = .11$, partial $\eta^2 = .03$), while false alarm rates only significantly differed in the orientation factor but not in the fame factor (Fame: $F(1, 74) = .37$, $p = .54$, partial $\eta^2 = .01$; Orientation: $F(1, 74) = 9.11$, $p = .003$, partial $\eta^2 = .11$; Interaction: $F(1, 74) = .66$, $p = .42$, partial $\eta^2 = .01$). Since the diffeomorphed scrambled faces retain some aspect of facial structure, participants might have been more likely to guess *intact* when the scrambled faces were presented upright than inverted, resulting in the significant effect of orientation on false alarms. On the other hand, the visual information preserved in the diffeomorphed images cannot indicate whether a face is famous

or novel. Hence, unlike logo experiments, fame did not significantly modulate false alarms. Moreover, the main effects of fame and orientation on bias estimates were significant (Fame: $F(1, 74) = 9.98$, $p = .002$, partial $\eta^2 = .11$; Orientation: $F(1, 74) = 16.14$, $p < .001$, partial $\eta^2 = .18$; Interaction: $F(1, 74) = .32$, $p = .58$, partial $\eta^2 = .00$). That is, the bias was more conservative for inverted than upright faces and for novel than famous faces, indicating that participants were more likely to respond scrambled when the stimulus was unfamiliar.

RTs to intact trials again support that no speed/accuracy tradeoff occurred. Participants were faster in the same conditions they were more accurate in although that RT difference only reached significance for orientation ($F(1, 74) = 5.02$, $p = .028$, partial $\eta^2 = .06$) and not for Fame ($F(1, 74) = 1.79$, $p = .19$, partial $\eta^2 = .02$) nor the Interaction ($F(1, 74) = .96$, $p = .33$, partial $\eta^2 = .01$).

To further ensure that our fame effects were not attributable to differences in salience, logistic hierarchical linear models that included both mean and maximum image salience, as well as fame and orientation, were fitted to predict accuracy of intact trials using the *lme4* package (Bates et al., 2015) in R version 4.1.2 (R Core Team, 2021). The participant factor was modeled under a random intercept design for all the models tested. Results confirmed the effect of familiarity, with higher accuracy for famous and upright faces than novel and inverted faces when controlling for the salience factors (Fame: $\beta = 0.10$, SE = 0.04, $Z = 2.55$, $p = .01$; Orientation: $\beta = 0.46$, SE = 0.04, $Z = 11.11$, $p < .001$). Interestingly, in this model the interaction effect was no longer significant ($\beta = 0.11$, SE = 0.06, $Z = 1.94$, $p = .053$), in contrast with the significant interaction effect from the repeated-measure ANOVA on d'.

To examine whether the salience measures can account for the insignificant interaction effect, we removed the salience measures from the full model. The reduced model only contained fame, orientation, and their interaction with a random participant design. After fitting this model to predict the accuracy of intact trials, we continued to see a significant main effects of fame and orientation (Fame: $\beta = 0.10$, SE = 0.04, $Z = 2.55$, $p = .01$; Orientation: $\beta = 0.46$, SE = 0.04, $Z = 11.11$, $p < .001$) and insignificant interaction effect ($\beta = 0.11$, SE = 0.06, $Z = 1.94$, $p = .053$). In fact, the results were unchanged whether salience was included or not, indicating that salience not only did not account for the insignificant interaction effect but also did not share variance with familiarity factors. Instead, the insignificant interaction of fame and orientation in the hierarchical model likely stems from the accuracy measure, which does not include false alarms that are part of the d' measure.

Ratings confirmed that the famous faces we chose (on the basis of pilot data) were more familiar to our current participants than the novel faces (Famous faces: 5.35 on the same 7-point Likert scale as previous experiments; Novel faces (two catch trials): 2.43 out of 7 points).

Discussion

In this experiment, we extended the familiarity effect on perception to famous faces. Results showed that famous faces were better perceived than novel faces (fame effect) and that upright faces were better perceived than inverted faces (orientation effect). These results extend the findings in Experiment 2 to a domain in which we are all experts. Importantly, both the famous and novel faces are clearly faces and our task did not require participants to recognize the face, but simply discriminate an intact face from one that is clearly not an intact face. In other words, these results indicate that our ability to detect the presence of a face is better if that face is known to us. The significant orientation effect also reflects familiarity, because here the very same features are present for both upright and inverted faces. The only difference is that we more frequently encounter upright faces than inverted faces. The interaction effect between fame and orientation reflected a significant fame effect for upright faces only, which is also aligned with the hypothesis that we have greater expertise with upright faces. Inversion can affect recognition of objects in general, but face recognition is especially impaired (see Valentine, 1988, for review). It is also consistent with data showing that inverted faces are less readily perceived as familiar (Marzi & Viggiano, 2007). Finally, this experiment extends the effects of familiarity, induced by either fame or orientation, to more natural stimuli.

Experiment 4: Face detection (online)

The fact that participants false alarmed more to upright diffeomorphed faces than inverted ones suggests that the diffeomorphs may have retained some face structure. We chose to scramble the faces using a diffeomorphed procedure because they retained the central concentration of features (as opposed to 100% phase scrambling) but yet, at least with extended viewing, they clearly look different than intact faces. However, if diffeomorphed faces retained some face structure, it is possible that participants adopted a conservative strategy to respond scrambled unless they explicitly recognized the person when the stimuli were presented only briefly. In other words, although the task did not explicitly require recognition, it is possible participants nonetheless chose to make it a recognition test. Hence, Experiments 4 and 5 were designed to rule out the possibility that explicit recognition was responsible for the superior performance on familiar faces. In Experiment 4, we removed the diffeomorphed faces and simply included a second mask, making the

task more similar to a pure detection task: was there a face present or not. In Experiment 5, we directly compared the time needed to make an intact/scrambled judgment to the time needed to explicitly recognize a face.

Participants

This experiment was pre-registered on the Open Science Framework (https://osf.io/r9c2g). A pilot study was conducted to estimate the required sample size at 80% power. Twenty-nine participants were recruited from the University of Illinois online participant pool. We planned to drop participants who met the following pre-registered exclusion criteria: (1) the overall accuracy of all the staircases was 50% (the chance level) or lower. (2) the dropped frame rate (calculated as in Experiment 2) in the staircase experiment was larger than 5%. However, no pilot participant's data was excluded given these criteria. A 2 (fame) × 2 (Orientation) repeated-measures ANOVA was conducted on the duration estimated from the staircase procedure (for details, please see Results section below). Results showed that only the main effect of orientation was significant (Fame: $F(1, 28) = 3.67$, $p = .066$, partial $\eta^2 = .12$; Orientation: $F(1, 28) = 8.48$, $p = .007$, partial $\eta^2 = .23$; Interaction: $F(1, 28) = .49$, $p = .49$, partial $\eta^2 = .017$). The power analysis focused on the number of participants needed to detect the main effect of fame. To reach power at 0.8 based on the pilot results, we needed 62 participants for the main effect of fame. Thus, 62 participants (43 females, mean age = 23.5 years) were recruited from Prolific.org. All participants had self-reported normal or corrected-to-normal vision and agreed to the consent approved by the University of Illinois Institutional Review Board before the start of the experiment. The same exclusion criteria as in pilot study were applied. All the recruited participants, however, passed these criteria.

Stimuli and procedure

The stimuli used in this experiment were the same intact faces used in the main task of Experiment 3. Instead of an intact/scrambled judgment, however, participants performed a detection task in which they were instructed to simply respond whether or not they saw a face (Fig. 7). Four staircases were conducted to determine the duration needed for each face type (famous-upright, famous-inverted, novel-upright, novel-inverted). The order of the staircases was random. Each staircase had 90 faces and 180 trials in total. When a face was not present, an additional phase-scrambled mask was presented in place of the face. This mask, however, differed from the final mask so ask to create a visual flicker similar to that experience in the face present trials. The program was written in javascript and using the *jsPsych* package (de Leeuw, 2015, version 6.3.0), and hosted on Pavlovia.org. The same credit card size control method was used.

Practice: The same procedure as in Experiment 3 was used except instead of making intact/scrambled judgment, participants made face/no face judgment. The same four upright novel faces and their corresponding inverted version were used as in Experiment 3.

Rating: The procedure was the same as in Experiment 3 except we increased the number of catch trials to 30 novel faces.
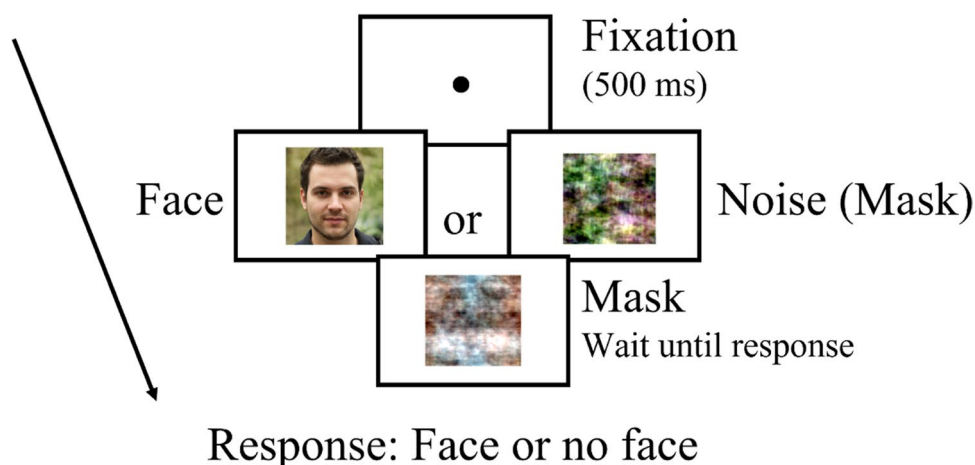

Fig. 7 The procedure of the detection staircase. After a fixation, a face or a noise appeared. The task is to respond whether a face appear or not. The duration to achieve 71% accuracy was estimated by the staircases

Results

In this experiment, the duration estimated by the Quest algorithm for a 71% accuracy threshold was used as the dependent variable. In particular, we compared the estimated duration needed to detect a familiar face from that needed to detect an unfamiliar face. We note that if a participant has very high performance and is sufficiently above 71% accuracy at the refresh rate of the monitor (the shortest duration at which we can present) then it is possible that the duration estimated by Quest will be negative. If Quest estimated a duration lower than the refresh rate of the monitor, then the actual presentation time was set to the refresh rate. It is this structural limit that will cause the estimate to end up negative; that is, Quest will continue to request a shorter duration than is possible and thus the participant will continue to respond correctly (because the refresh rate is adequate) and thus the estimated duration will continue to trend downward until it is in negative territory. For instance, with a threshold set at 71%, the

duration suggested by the Quest algorithm could range from -857 (100% accuracy) to 1,124 (0% accuracy). In short, a more negative duration, although not itself possible, does reliably indicate that participants find those trial types easier and that they would need shorter duration to perform equivalently to a trial type whose estimated duration is higher. The famous upright condition had the shortest estimated duration while the novel inverted condition had the longest estimated duration (Famous upright: M = -263, SE = 41.0; Famous inverted: M = -180, SE = 41.4; Novel upright: M = -216, SE = 42.7; Novel inverted: M = -133, SE = 34.4) (Fig. 8).

A 2 × 2 repeated-measures ANOVA (*afex* package (Singmann et al., 2021) in R version 4.1.2 (R Core Team, 2021)) was conducted on the estimated duration and revealed significant main effects of fame and orientation but no interaction of fame and orientation (Fame: $F(1, 61)$ = 5.83, p = .019, partial η^2 = .09; Orientation: $F(1, 61)$ = 21.8, p < .001, partial η^2 = .26; Interaction: $F(1, 61)$ = .001, p = .98, partial η^2 = .00) (Fig. 9).
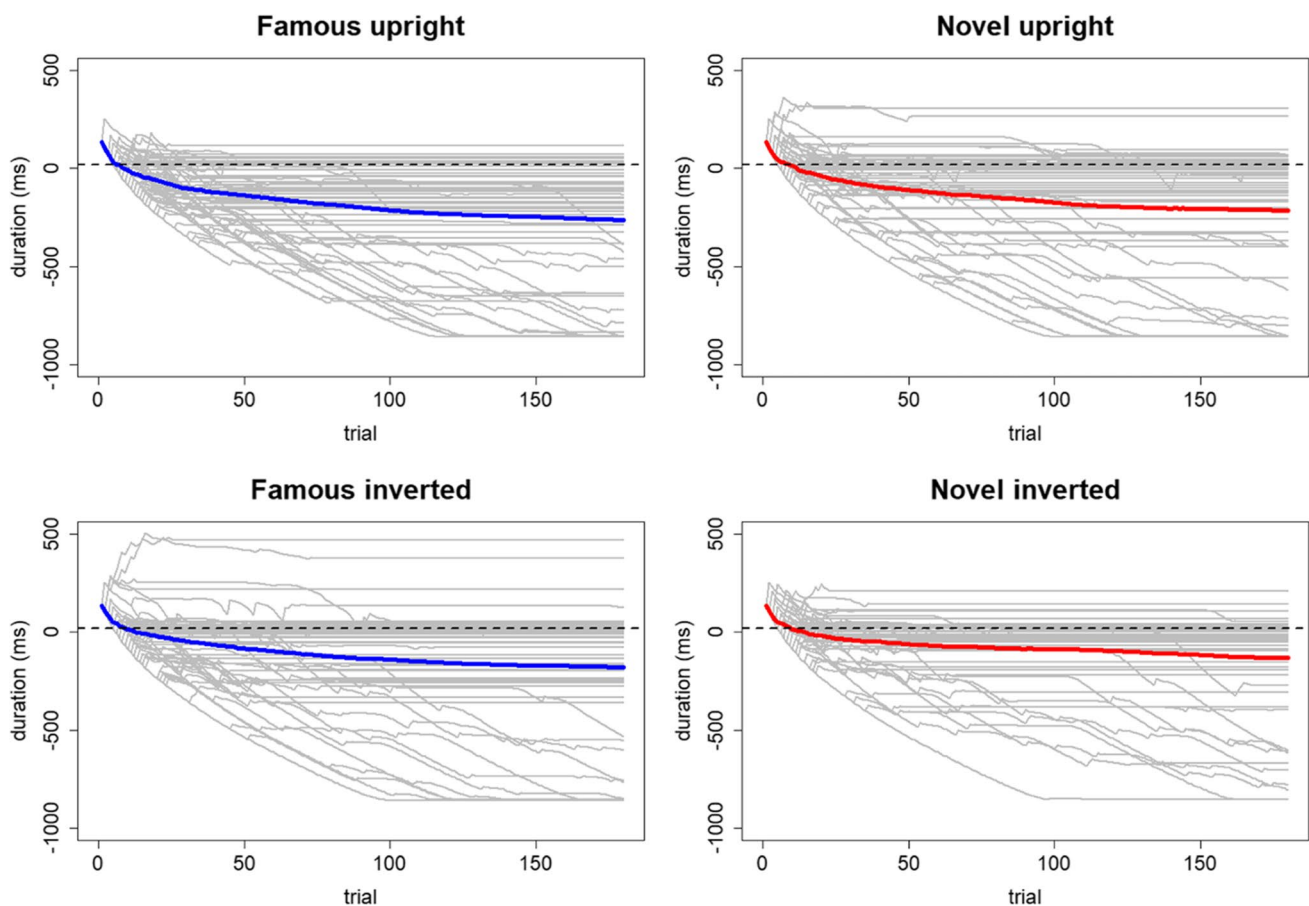

Fig. 8 The estimated duration traceplots of each participant under each condition. The y-axis is the estimated duration by the Quest algorithm. The x-axis is the trial numbers. Each grey line represents one participant. The average of each condition was the thick colored line. The estimated duration needed for famous faces (blue trace) was shorter for novel faces (red trace), and the estimated duration for upright faces (top row) was shorter than inverted faces (bottom row). See Results section for explanation of the negative values
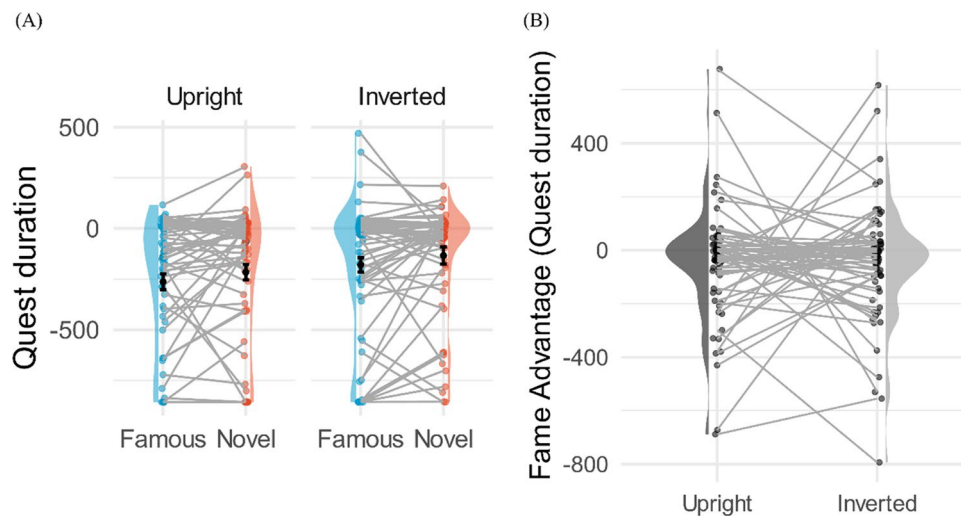
(A)

(B)

Fig. 9 A Within-subject comparison violin plots for Experiment 4. Figure legends are the same as previous within-subject comparison violin plots. The y-axis shows the duration estimated by Quest algorithm. The bar within each raincloud represents the 95% within-subject confidence interval. The left-hand plots show the upright faces, and the right-hand plots show the inverted faces. Blue represents famous faces. Red represents novel faces. Main effects of fame and orientation were significant, and no interaction was observed. **B** The duration fame advantage is calculated by subtracting the duration estimated by Quest algorithm for the novel condition from the duration for the famous condition. Each dot represents one participant, and each participant is connected with a line

For the rating results, again, famous faces were rated as more familiar (M = 5.66, SE = 0.15) than novel faces (M = 1.86, SE = 0.17). In this experiment, we increased the catch trials in the rating task from two faces to 30 novel faces. Thus, a paired t-test can be conducted on the rating to show that the rating difference was indeed significant ($t(61) = 17.5$, $p < .001$, Cohen's d = 3.08).

Discussion

Experiment 4 results indicated that participants could detect famous and upright faces more quickly than novel and inverted faces; that is, the duration estimated by Quest to achieve 71% accuracy was shorter when the faces were famous or upright than when they were novel and inverted. In other words, we again found the famous faces, whether upright or inverted, were more readily perceived, but this time using a more traditional detection task and duration, rather than accuracy, as our dependent variable. Importantly, in this task the unfamiliar faces look nothing like a phase scrambled mask and so the participants should not be biased to say no face when an unfamiliar face is present. Interestingly, the significant interaction of fame and orientation found in Experiment 3 was not replicated in the current experiment. We note, however, that we did not power our study to detect the interaction, as our main interest was in the effect of familiarity.

Experiment 5: Recognition versus intact/scrambled

To further demonstrate that the intact/scrambled judgment did not depend on explicit recognition, in Experiment 5, we compared the duration needed to reach 71% accuracy on the intact/scrambled task to that needed to explicitly recognize 71% of the famous faces. If participants were using explicit recognition to complete the intact/scrambled task, then the two tasks should require similar durations. This experiment was pre-registered on the Open Science Framework (osf.io/y3ta8).

Participants

We conducted a pilot study to estimate the duration difference between the intact/scrambled task and explicit recognition task for the purposes of a power calculation. Twenty-seven participants were recruited from the University of Illinois online participant pool to perform the explicit recognition staircases and familiarity ratings on target faces (117 famous faces and 20 sampled novel faces). If the number of dropped frames exceeded 5% or the familiarity rating difference between famous and novel faces was less than 1, a participant was excluded, resulting in exclusion of one participant. The rating exclusion

criterion was added in current experiment because the ground truth of recognition staircasing is based on subjective familiarity ratings from the other set of participants. If current participants disagree with the ground truth that we assumed, their accuracy of the recognition judgment will be low and thus erroneously result in longer estimated duration. This is not an issue in previous experiments because the familiarity level of the targets we chose could not have produced the familiarity effect observed in d'; underestimating familiarity in those experiments would only serve to weaken any familiarity effect we observed.

To estimate the effect size, an independent t test was conducted to compare estimated duration under the intact/scrambled task (Experiment 3 staircasing data) and under the explicit recognition task. Results showed that estimated duration under the intact/scrambled task was significantly faster than the explicit recognition task ($t(27.68) = 2.74$, $p = .01$, Cohen's d = 0.88). Although our effect size was computed using independent samples, our current design was within-subject (i.e., the same participant went through both intact/scrambled and explicit recognition tasks). We would need a sample size of 13 to achieve 80% power to detect an effect of this magnitude using a paired t-test (conducted with the *pwr.t.test* function in *pwr* package in R version 3.6.3). However, because our effect size was initially computed from independent samples and our actual design is repeated-measured we deemed it safer to increase the targeted sample size to 20. Thus, 22 participants were recruited either from the University of Illinois online participant pool (eight participants, compensated with course credits) or from Prolific.org (14 participants). All participants had self-reported normal or corrected-to-normal vision and agreed to the consent approved by the University of Illinois Institutional Review Board before the start of the experiment. We dropped the data from two participants who met the following pre-registered exclusion criteria: (1) the dropped frame rate in the staircase experiment was larger than 5%; (2) the rating difference between the familiar faces and novel faces was smaller than 1; (3) The accuracy of the intact/scrambled quest or the recognition quest was 0.5 (chance level) or lower. Twenty participants (11 females, mean age = 22.2 years) passed the criteria.

Stimuli and procedure

Two staircases and one rating task were included in this experiment. The first staircase required participants to make the same intact/scrambled judgment as in Experiments 1, 2, and 3. The second staircase, a recognition staircase, asked participants to respond whether they have seen the face before or not (Fig. 10). In the recognition staircase, participants were explicitly instructed to respond based on their daily experiences and not on the basis of the experiment (i.e., having only seen the face in the intact/scrambled experiment). The order of the two staircases was fixed. The intact/scrambled staircase was always first followed by the recognition staircase to make sure that intact/scrambled results were not influenced by the recognition task. The upright faces in both the intact/scrambled staircase and the recognition staircase were the same as in Experiment 3 (90 famous faces, 90 novel faces). In addition, the intact/scrambled staircase included the same scrambled faces as in Experiment 3. The tasks and procedures for the practice and the intact/scrambled staircase were the same as in Experiment 3. The recognition staircase required participants to answer whether they have seen the face before or not and included famous and novel upright faces. The familiarity rating task was the same as in Experiment 4. The same *jsPsych* package (de Leeuw, 2015, version 6.3.0) and javascript language were used for programming the experiment, and it was hosted on
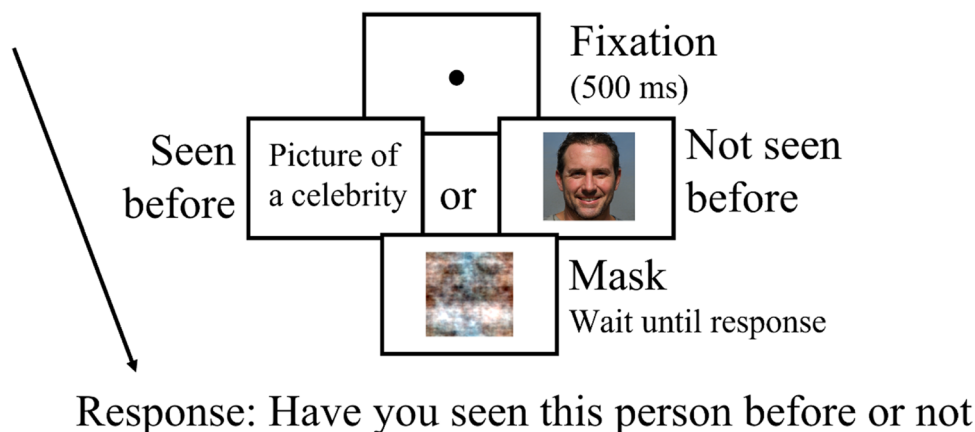

Fig. 10 The procedure of the recognition staircase. After a fixation, an intact face appeared. Participants were instructed to respond whether they have seen this person before or not. The duration to achieve 71% accuracy threshold is estimated

Pavlovia.org. The same credit card size method was used to control visual angle.

Results

A paired t-test was conducted on the duration estimated by Quest for the two staircases, one for recognition and one for intact scrambled. Importantly, the same intact faces were used in both experiments. The duration estimated by Quest needed to reach 71% accuracy on the recognition judgment (M = 51 ms, SE = 51) was significantly longer than the duration needed to reach the same level of accuracy on the intact scrambled judgment (M = -197 ms, SE = 77) ($t(19)$ = -3.39, p = .003, Cohen's d = .83) (Fig. 11). Even if we based the analysis on the actual presentation time (i.e., the minimum duration is the refresh rate), the recognition judgment (M = 97 ms, SE = 21) still required a longer duration than the intact/scrambled judgment (M = 19 ms, SE = 1). Rating results again confirmed that the famous faces were significantly more familiar to participants than the novel faces ($t(19)$ = 12.55, p < .001, Cohen's d = 3.68).

Discussion

The current experiment estimated the required duration to perform either intact/scrambled judgment or recognition judgment. Results demonstrated that the recognition judgment required a significantly longer duration to reach the same accuracy threshold as the intact/scrambled judgment. In fact, Quest estimated that participants needed almost 250 ms longer to recognize the face than to say that an intact face was present. This result also accords with previous

work comparing identification and detection performance of the same faces (Grill-Spector & Kanwisher, 2005; Mack & Palmeri, 2010). The researchers briefly presented faces or nonobject texture followed by a mask and asked participants to perform two tasks on the same trial: a detection task in which they had to indicate which of two intervals contained the object (in this case a face) and an identification task in which they had to determine whether that face (regardless of in which intervals it was presented) was Harrison Ford or a non-famous male. Both studies found that performance on the detection task did not depend on performance on the identification task; that is, the probability of a hit on the detection task was no higher when the participant correctly identified the face than when they were not able to correctly identify the face. Thus, like our data, these results suggest that participants do not need to explicitly identify a face in order to detect it. We note, however, that in these papers (Grill-Spector & Kanwisher, 2005; Mack & Palmeri, 2010) the researchers did not analyze detection as a function of whether the face was Harrison Ford or a non-famous male. If they had, our data would suggest that participants should be better at detecting a face when the face was Harrison Ford than when it was a non-famous face.

Together with Experiment 4, the results of Experiment 5 aligned with our hypothesis that the intact/scrambled judgment does not depend on explicit recognition processes. Instead, any explanation must appeal to an effect of familiarity that precedes explicit recognition. We suggest, in line with predictive coding models, that our results instead reflect a match to a prediction (or template matching) process, in which our visual system's predictions are informed by familiarity and thus famous faces generate less prediction error
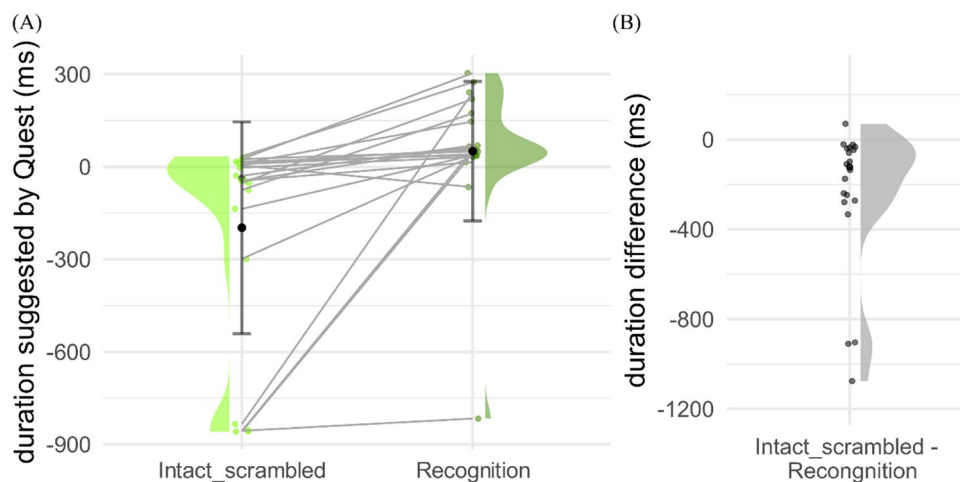

Fig. 11 **A** Within-subject comparison violin plots for Experiment 5. The results of the duration estimated by Quest algorithm for the intact/scrambled staircase (left) and the recognition staircase (right). The interval on each bar represents one standard deviation. Estimated duration was significantly lower for the intact/scrambled staircase

than the recognition staircase. **B** The subtraction of the duration estimated by Quest algorithm of the intact/scrambled staircase to of the recognition staircase. All the participants except one had shorter estimated duration in the intact/scrambled staircase than the recognition staircase

than novel ones and thus resolve more quickly. Identification then follows the detection process.

General discussion

In this study, we investigated the influence of familiarity on participants ability to discriminate intact images from scrambled images. To achieve this goal, we compared famous logos (Experiments 1 and 2) and faces (Experiments 3, 4, and 5) with computer-generated, and thus novel, logos and faces, respectively. In additional to the fame factor, we included an orientation factor in Experiments 2 and 3 to explore the familiarity of viewpoint on detection. Results showed that highly familiar intact stimuli, defined by fame or orientation were better discriminated from scrambled stimuli than unfamiliar intact stimuli were. This result indicates that familiarity can modulate perceptual discrimination, even when the task is as simple as determining an image is intact.

Both familiarity effects of fame and orientation reflect life-long experiences. Throughout our lives we experience many more upright faces than inverted ones. Famous faces and logos are generally ones we have encountered many times. In Experiment 1 we also explored more short-term familiarity, that is, repetition within the experiment. This factor had no effect on sensitivity, but it was admittedly a very weak manipulation of familiarity. Further work will be needed to determine how much exposure to logo, or face, is needed to produce a familiarity effect on intact/scrambled sensitivity.

Unlike other tasks that have investigated the effect of familiarity on perception, our task does not require an explicit recognition judgment. However, in Experiments 4 and 5 we explored the possibility that participants nonetheless approached the task as a recognition task, withholding response until they recognized the logo or person. This strategy maybe particularly likely given that our scrambled images shared lots of the same features as our intact ones. Thus, in Experiment 4 we removed the diffeomorphed faces and instead conducted a simple detection task in which participants had to respond whether a face present or not. When the face was not present an additional phase-scrambled mask was presented in place of the face, to control for the brief flash that was experienced when the face was presented. Again, we replicated the finding that familiar faces (famous and upright) required shorter durations to detect than unfamiliarity (novel or inverted) faces. Thus, our results in Experiment 3 cannot be explained by confusion between intact faces and diffeomorphs causing participants to adopt an explicit recognition strategy. To rule out an explicit face recognition strategy more generally, we directly compared the duration needed to accurately perform an intact/ scrambled judgment with the duration needed to accurately

perform an explicit recognition judgment (Experiment 5). Results of Experiment 5 indicated that the recognition judgment needed 248 ms longer duration to achieve the same level of accuracy as the intact/scrambled task, suggesting that participants were not waiting to explicitly recognize the faces in the intact/scrambled task before responding intact. Thus, we concluded that familiarity effect is not driven by explicit recognition per se, but instead reflects an ease of processing induced by familiarity that precedes recognition.

Past research on familiarity

Importantly, the familiarity effect demonstrated here goes beyond that established in the literature via recognition tasks (Bruner, 1957). Others have raised comprehensive theoretical challenges with respect to the claim that familiarity can influence what we see (i.e., perception; Pylyshyn, 1999; Firestone & Scholl, 2016). Viewing human perception as constructed by multiple stages, Pylyshyn (1999) argued that the impact of knowledge, including familiarity, on vision stems from the more cognitive stages of recognition, identification, or response selection, instead of what he refers to as "early vision," although he includes "some local vision-specific memory" (p. 344) in his conception of early vision. Because the term "early vision" has since come to imply V1-V4, and we do not mean to imply that our effect is occurring in V1-V4, we refer to Pylyshyn's "early vision" stage as simply perceptual (as opposed to cognitive). In fact, if we were to conjecture as to where in the brain our perceptual effect resides, we would suggest mid-level vision, beyond V1-V4 (for more on this, see section entitle When is familiarity having its effect below). Returning to Pylyshyn's argument, explicit recognition is a cognitive process that involves access to multi-modal memory and knowledge. Hence, merely showing the knowledge can influence the recognition stage is not equivalent to demonstrating of the effect of familiarity on visual perception, or as he would call it "early vision." Our task does not require explicit recognition and thus cannot be relegated to a cognitive stage of processing. Instead, we show that familiarity affects how readily you can even determine an intact image was presented and thus can be construed as a perceptual, as opposed to cognitive, effect.

Firestone and Scholl (2016) further explicated the possible pitfalls of the experimental designs used to support claims that knowledge can influence perception. With the current experimental design, we believe we have overcome the pitfalls raised in Firestone and Scholl (2016) to demonstrate that the familiarity can indeed facilitate the perception of known stimuli. First, instead of making subjective reports such as asking participants to estimate the distance of a ball traveled (Witt et al., 2004), this study used objective performance-based measurements (e.g., d' and estimated duration) under rapid presentations. Second, the familiarity

manipulation was never mentioned to the participants. Even if the participants noticed the manipulation, it is still irrelevant to the intact/scrambled judgment. Thus, the current experiment is not susceptible to demand characteristics, in which participants change their answer to suit the experimenters' expectations. Lastly, our effects are unlikely to be due to low-level feature differences. For instance, we used hierarchical linear models to examine the effects of familiarity while modeling out any salience differences in the stimuli. Furthermore, the familiarity effect of orientation was still present even though upright and inverted stimuli share the same sets of low-level features.

As mentioned above, although we believe recognizability plays a critical role in our effects, we do not believe that our effects depend on explicit recognition. First, our task does not explicitly require that the participant recognize the stimuli. This fact is particularly salient for the face experiments; although the novel faces are unfamiliar, they are still very obviously faces. In fact, most of us encounter unknown faces every day and we do not struggle to see them as faces. Thus, it is not logically necessary that participants recognize the individual to know a face was present. Indeed, in our control experiment we showed that participant needed more time to explicitly recognize an individual than they did to indicate a face was present. Yet, participants need much less time and are more accurate at detecting the presence of a known individual. Coupled with the avoidance of the pitfalls described by Firestone and Scholl (2016), we believe our familiarity effect reflects a true perceptual effect; that is, familiar stimuli, along with other statistically regular stimuli, are perceived more rapidly than unfamiliar stimuli.

Familiarity has also been shown to impact visual search. No matter whether participants are searching for simple symbols (Rauschenberger & Chu, 2006), single letters (Malinowski & Hübner, 2001; Shen & Reingold, 2001; Wang et al., 1994), words or non-words (Flowers & Lohr, 1985), cars or birds in photos (Hershler & Hochstein, 2009), everyday objects (Mruczek & Sheinberg, 2005), logos (Qin et al., 2014), or faces (Tong & Nakayama, 1999), the search speed is decreased by distractor or target familiarity. When search efficiency was enhanced by distractor familiarity, authors argued that the enhancement demonstrated a more efficient background grouping process, making it easier to reject familiar distractors which were outnumbered targets in most conditions. (Malinowski & Hübner, 2001; Shen & Reingold, 2001). On the other hand, when the target was familiar, the authors argued that this improvement was due to a more readily available target templates that can make the attention allocation more efficient (Hershler & Hochstein, 2009). With a combination of the two situations, Mruczek and Sheinberg (2005) showed that the most efficient case is "… when targets previously searched for are located among familiar distractors." (p. 1031). This suggests that

the expectation of target along with the experiences of the target and distractors can enhance the search speed of the target. Importantly, however, all of these explanations would fall into the cognition category of Pylyshyn's (1999) model, as they reflect top-down effects of familiarity on attention mechanisms (not perception itself). Given our findings, however, it is also possible that such effects reflect true perceptual effects, with more efficient processing of familiar stimuli than novel stimuli.

Real-world statistical regularity

Our data add to a growing body of work showing that statistically regular images are more readily detected than statistically irregular images. In particular, familiarity can be added to a list of types of prior knowledge that influence perception, along with the probability of occurrence (Greene et al., 2015), category representativeness (Caddigan et al., 2017), expectation (Smith & Loschky, 2019), and typicality of viewpoint (Center et al., 2022). We refer to all such images that more readily match our prior knowledge or expectations as statistical regular. These statistical regularities allow the visual system to predict incoming stimuli and ultimately recognize the input. Such predictions are key to predictive coding models. With higher real-world statistical regularity or better templates, the prediction error in the predictive coding models should, on average, be minimized. That is, the more familiar inputs result in a smaller prediction error relative to a hypothesis about *what* is out there because they are better predicted by templates that capture the statistical regularities of the world. Further, minimization of the prediction error will be faster when the input is a closer match to the prediction (statistically regular) than when the prediction error starts off large (statistically irregular). Therefore, high real-world statistical regularity allows the brief and masked targets used here to more quickly minimize the prediction error and reach the stage of being discriminated or detected.

When is familiarity having its effect?

Although the underlying processes of when familiarity has its effect in the brain is unknown in the current study, we can make an informed guess. As alluded to before, we posit that the observed familiarity effect may be occurring in mid-level visual areas responsible for the N300 ERP component. In other work, we link the N300 to other statistical regularities (Kumar et al., 2021), and indeed familiarity also impacts the N300 amplitude (Scott et al., 2005), such that unfamiliar stimuli evoke a greater N300 indicating more processing for the unfamiliar than the familiar stimulus. The N300 component has been source localized to lateral occipital cortex (LOC; Sehatpour et al., 2006) and the posterior fusiform gyrus (Schendan & Lucia, 2010), depending on the stimulus.

Both of these regions are later visual areas than V1-V4. In addition, the onset of the observed familiarity effect arises after the sensory processes (indexed by visual mismatch negativity (vMMN), for an example) and before meaning assessment (indexed by N400 component; Kutas, & Federmeier, 2011). The relation of the N300 to other ERP components thus also support the claim that the effect emerges at mid-level visual areas.

One might also ask when this effect occurs relative to feedback or recurrent signals. Our results are revealed using backwards masking and some forms of backward masking are thought to interfere with recurrent signals. We have suggested that our masks interfere with the brains ability to minimize prediction error in a predictive coding framework. This iterative process of prediction error minimization is a type of recurrence and depends in part on "feedback" signals (i.e., the prediction is thought to "feedback"). Such a claim would accord well with our masks interfering with recurrence or feedback. However, we note that it is not clear whether our form of backward masking interferes with feedback. Interference with feedback is typically inferred when masking produces a u-shaped curve with respect to the stimulus-onset asynchrony (SOA) between stimulus and mask. We did not vary SOA here and so we cannot be sure whether the mask is having its effect on feedforward or feedback processes. Thus, it is equally likely that the mask interferes with the ability of the stimulus to make contact with a prediction, or indeed any template or memory representation, on the feedforward sweep or during recurrent processing.

Familiarity or frequency?

Because, according to predictive coding models, statistical regularities are meant to minimize prediction error in general, there are a number of factors that may go into their construction. Our study, along with the previous studies to use a similar paradigm (Caddigan et al., 2017; Center et al., 2022; Greene et al., 2015; Smith & Loschky, 2019), allow us to postulate the factors that go into their construction. One obvious factor is frequency of occurrence; that is our visual system should be more likely to predict something that occurs frequently in our environment. Familiarity in the form of things frequently encountered in daily life, clearly falls into this category. This does not necessarily imply that familiarity should be considered as one and the same as frequency of occurrence, however. Familiarity may also reflect recognizability that is not fully dependent on frequency of occurrence. For example, even someone or something that we do not encounter very often (e.g., Albert Einstein), might still be famous enough that they are highly recognizable to us. Thus, it is possible that the system chooses familiar people or objects as possible predictions because our goal is recognition and so it starts with things it knows it has

recognized before, whether they are frequently encountered or not.

Smith and Loschky (2019) also used familiar scenes, although they manipulate sequential expectation, which adds another dimension to the idea of frequency; it is not just the frequency of being encountered in general but the frequency of being encountered in a given context. The probable images of Greene and colleagues (2015) also bear some relation to our familiarity effects in that they also likely are informed by frequency of occurrence. Although in this case it would not be the frequency of specific inputs; their stimuli were not specifically familiar. Instead, their probable stimuli could be described as being pulled from frequently encountered classes of stimuli or from stimuli that frequency of co-occur with each other. For example, cars frequently co-occur with paved roads. While participants may not have seen the exact "probable" image before, they will have many experiences of cars co-occurring with paved roads. In short, the probable images from Greene et al. (2015) may not specifically be familiar but instead represent the frequency of co-occurrence or familiar classes of stimuli, capturing the frequency of more conceptual representations of our past experience rather than specific instances.

Detection versus identification

At first blush, our data may seem to conflict with the literature investigating the time course of object detection and object categorization. Mack and Palmeri (2010) showed that object detection precedes basic-level categorization. That is, when a target briefly flashes (e.g., 17 ms), the accuracy of detection task is above chance while the performance of basic-level categorization task is at chance, suggesting that the visual information gleaned during this rapid presentation is sufficient to support object detection but not basic-level categorization. Similar results were observed in reaction time (Bowers & Jones, 2008); basic-level categorization requires longer reaction time than object detection. The current study, on the other hand, shows that object detection can be modulated by the level of familiarity of targets. How might we reconcile these two results?

First, we note that familiarity may precede basic-level categorization. That is, incoming stimuli may make contact with a prior knowledge (a familiarity signal or statistical regularity) before basic-level categorization can occur. Relatedly, these basic-level categorization studies are investigating the time course of an explicit recognition judgment, which we have shown require more time than our intact/scrambled task (Experiment 5). Similarly, Mack and Palmeri (2010) have shown that participants do not need to identify a face before they can detect it. Since our intact/scrambled task does not require participants to explicitly report the identity of targets, it is not surprising that our detection

task might precede basic level categorization. To reconcile these two results, then, one must only presume that incoming stimuli make contact with a familiarity signal or statistical regularity before either detection or basic-level categorization. In other words, although the limited processing that occurs under brief presentation may not suffice to make an explicit basic-level categorization judgment it may be sufficient to produce a familiarity signal or be compared to a statistical regularity and thus influence detection sensitivity.

Related work and models

Although predictive coding models provide an adequate framework for explaining the observed familiarity effect to perceptual detection in the current study, it is not the only model that can explain our data. Our data are consistent with any model that appeals to prediction or some form of perceptual hypothesis testing (PHT) in which perception is conceived of as a process of generating a hypothesis about the sensory input and that those hypotheses are informed by past experience (Clark, 2013; Gregory, 1980; Helmholtz, 1925; Hochberg, 1981; Huang & Rao, 2011; Kersten et al., 2004; Rock, 1983). PHT has had a long history starting with Helmholtz, who argued that perception was a process of interference. These ideas have resurfaced periodically (Gregory, 1980; Hochberg, 1981; Rock, 1983) with the latest versions appealing to priors in a Bayesian framework (Kersten et al., 2004) or predictions in a predictive coding framework (Clark, 2013; Friston, 2005; Huang & Rao, 2011; Rao & Ballard, 1999). The advantages of the predictive coding theory over the earlier theories is that it both makes clear that these hypotheses reside (although not necessarily exclusively) in visual cortex and it provides an explanation for why statistically irregular stimuli should both take longer to perceive and evoke more activity (Kumar et al., 2021; Schendan & Kutas, 2002, 2003, 2007; Scott et al., 2005; Torralbo et al., 2013; Vo & Wolfe, 2013).

Our findings are also related to the perceptual learning (e.g., Ahissar & Hochstein, 2004) and statistical learning (e.g., Aslin, 2017) literatures in that these rely on repeated practice or exposure. Familiar stimuli become familiar through repeated exposure, and as in those literatures, that repeated exposure results in changes within the visual system that enable perceptual fluency. One important difference between those studies and the current studies are that perceptual and statistical learning are generally produced and assessed within the same experiment; that is, unlike our real-world regularities which are learned over years, the learning is specific to the experiment.

Our paradigm and results may remind readers of the recognition memory literature. Jacoby and colleagues proposed that recognition memory judgments stem from two factors: a sense of familiarity and direct memory of the items (Johnston et al., 1985). They hypothesize that the sense of familiarity reflected the perceptual fluency that is gained from repetition of as stimulus (Jacoby & Dallas, 1981; Johnston et al., 1985). They showed, for instance, that more fluently perceived items were more likely to be judged as more familiar (i.e., more likely to be judged "old"). We, on the other hand, showed that stimuli that participants agree are highly familiar (i.e., they are sure the know) are processed more fluently (i.e., require less time to perceive). The results presented here can be seen as complementary to theories claiming that familiarity is driven by perceptual fluency. Another way in which to connect these literatures, is to say that prior experience in the form of repetitions leads both to experiences of familiarity and perceptual fluency.

Conclusion

This study demonstrates that familiarity, like other real-world statistical regularities, impacts mere detection. We interpret these findings within the context of predictive coding, such that familiarity contributes to the predictions made by the visual system that speed the processing of incoming stimuli. Rather than a serial model in which the input first coheres into a potential object and then makes contact with prior knowledge, our data are better explained by a model that uses predictions, in the form of statistical regularities extracted from past experience, to perceptually organize the input. Incoming stimuli that more closely match the predictions undergo less iterative updating to minimize prediction error and are resolved sooner than less statistically regular stimuli. In short, our visual system is more likely to predict stimuli that we are familiar with, and thus we more readily perceive things that we are familiar with.

Data availability The deidentify raw data collected and analyzed in the current study are available on the Open Science Framework repository at the following site: OSF_familiarity_project_site. Experiments 2, 4, and 5 were preregistered on the Open Science Framework.

Declarations

Ethics approval All the experiments were approved by University of Illinois Institutional Review Board.

Consent to participate Informed consent was obtained from all individual participants included in the study.

Conflicts of interest/Competing interests The authors have no relevant conflicts of interest to be declared.

References

Ahissar, M., & Hochstein, S. (2004). The reverse hierarchy theory of visual perceptual learning. *Trends in cognitive sciences, 8*(10), 457–464.

Aslin, R. N. (2017). Statistical learning: A powerful mechanism that operates by mere exposure. *Wiley Interdisciplinary Reviews: Cognitive Science, 8*(1–2), e1373.

Baron, J. (2014). The word-superiority effect: Perceptual learning from reading. *Handbook of learning and cognitive processes, 6*, 131–166.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). *Lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1–7. 2014*.

Bowers, J. S., & Jones, K. W. (2008). Short article: Detecting objects is easier than categorizing them. *Quarterly Journal of Experimental Psychology, 61*(4), 552–557.

Bruner, J. S. (1957). On perceptual readiness. *Psychological Review, 64*(2), 123.

Bülthoff, I., & Newell, F. N. (2006). The role of familiarity in the recognition of static and dynamic objects. *Progress in Brain Research, 154*, 315–325.

Caddigan, E., Choo, H., Fei-Fei, L., & Beck, D. M. (2017). Categorization influences detection: A perceptual advantage for representative exemplars of natural scene categories. *Journal of Vision, 17*(1), 21–21.

Center, E. G., Gephart, A. M., Yang, P. L., & Beck, D. M. (2022). Typical viewpoints of objects are better detected than atypical ones. *Journal of Vision, 22*(12), 1–1.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and brain sciences, 36*(3), 181–204.

Cousineau, D., & O'Brien, F. (2014). Error bars in within-subject designs: A comment on Baguley (2012). *Behavior Research Methods, 46*(4), 1149–1151.

De Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a Web browser. *Behavior Research Methods, 47*(1), 1–12.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191.

Firestone, C., & Scholl, B. J. (2016). Cognition does not affect perception: Evaluating the evidence for "top-down" effects. *Behavioral and Brain Sciences*, *39,* e229.

Flowers, J. H., & Lohr, D. J. (1985). How does familiarity affect visual search for letter strings? *Perception & Psychophysics, 37*(6), 557–567.

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B: Biological Sciences, 360*(1456), 815–836.

Gollin, E. S. (1960). Developmental studies of visual recognition of incomplete objects. *Perceptual and Motor Skills, 11*(3), 289–298.

Greene, M. R., Botros, A. P., Beck, D. M., & Fei-Fei, L. (2015). What you see is what you expect: Rapid scene understanding benefits from prior experience. *Attention, Perception, & Psychophysics, 77*(4), 1239–1251.

Gregory, R. L. (1980). Perceptions as hypotheses. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences, 290*(1038), 181–197.

Grill-Spector, K., & Kanwisher, N. (2005). Visual recognition: As soon as you know it is there, you know what it is. *Psychological Science, 16*(2), 152–160.

Hershler, O., & Hochstein, S. (2009). The importance of being expert: Top-down attentional control in visual search with

photographs. *Attention, Perception, & Psychophysics, 71*(7), 1478–1486.

Helmholtz, H. V. (1925). Treatise on psychological optics. *Optical Society of America, 3*, 482.

Hochberg, J. (1981). On cognition in perception: Perceptual coupling and unconscious sinference. *Cognition, 10*(1–3), 127–134.

Honda, H., Abe, K., Matsuka, T., & Yamagishi, K. (2011). The role of familiarity in binary choice inferences. *Memory & cognition, 39*(5), 851–863.

Huang, Y., & Rao, R. P. (2011). Predictive coding. *Wiley Interdisciplinary Reviews: Cognitive Science, 2*(5), 580–593.

Jacoby, L. L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General, 110*(3), 306.

Johnston, W. A., Dark, V. J., & Jacoby, L. L. (1985). Perceptual fluency and recognition judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11*(1), 3.

Karras, T., Laine, S., Aittala, M., Hellsten, J., Lehtinen, J., & Aila, T. (2020). Analyzing and improving the image quality of stylegan. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 8110–8119).

Kersten, D., Mamassian, P., & Yuille, A. (2004). Object perception as Bayesian inference. *Annual Review of Psychology, 55*, 271–304.

King-Smith, P. E., Grigsby, S. S., Vingrys, A. J., Benes, S. C., & Supowit, A. (1994). Efficient and unbiased modifications of the QUEST threshold method: Theory, simulations, experimental evaluation and practical implementation. *Vision Research, 34*(7), 885–912.

Krueger, L. E. (1975). Familiarity effects in visual information processing. *Psychological Bulletin, 82*(6), 949.

Kumar, M., Federmeier, K. D., & Beck, D. M. (2021). The N300: An index for predictive coding of complex visual objects and scenes. *Cerebral Cortex Communications, 2*(2), tgab030.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event related brain potential (ERP). *Annual review of psychology, 62*, 621.

Lenth, R. V. (2022). *emmeans: Estimated marginal means, aka least-squares means*. R package version 1.8.1-1. https://CRAN.R-project.org/package=emmeans

Mack, M. L., & Palmeri, T. J. (2010). Decoupling object detection and categorization. *Journal of Experimental Psychology: Human Perception and Performance, 36*(5), 1067.

Malinowski, P., & Hübner, R. (2001). The effect of familiarity on visual-search performance: Evidence for learned basic features. *Perception & Psychophysics, 63*(3), 458–463.

Marzi, T., & Viggiano, M. P. (2007). Interplay between familiarity and orientation in face processing: An ERP study. *International Journal of Psychophysiology, 65*(3), 182–192.

Miller, G. A., Bruner, J. S., & Postman, L. (1954). Familiarity of letter sequences and tachistoscopic identification. *The Journal of General Psychology, 50*(1), 129–139.

Mruczek, R. E., & Sheinberg, D. L. (2005). Distractor familiarity leads to more efficient visual search for complex stimuli. *Perception & Psychophysics, 67*(6), 1016–1031.

Palmer, S., Rosch, E., & Chase, P. (1981). Canonical perspective and the perception of objects. In J. Long & A. Baddeley (Eds.), *International Symposium on attention and performance (attention and performance IX)* (pp. 135–151). Lawrence Erlbaum Associates.

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods, 51*(1), 195–203.

Pylyshyn, Z. W. (1999). Is vision continuous with cognition?: The case for cognitive impenetrability of visual perception. *Behavioral and Brain Sciences, 22*(3), 341–365.

Pylyshyn, Z. W. (2001). Visual indexes, preconceptual objects, and situated vision. *Cognition, 80*(1–2), 127–158.

Qin, X. A., Koutstaal, W., & Engel, S. A. (2014). The hard-won benefits of familiarity in visual search: Naturally familiar brand logos are found faster. *Attention, Perception, & Psychophysics, 76*(4), 914–930.

R Core Team. (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing. https://www.R-project.org/

R Core Team. (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing. https://www.R-project.org/

Rao, R. P., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience, 2*(1), 79–87.

Rauschenberger, R., & Chu, H. (2006). The effects of stimulus rotation and familiarity in visual search. *Perception & Psychophysics, 68*(5), 770–775.

Reingold, E. M., & Jolicoeur, P. (1993). Perceptual versus postperceptual mediation of visual context effects: Evidence from the letter-superiority effect. *Perception & Psychophysics, 53*(2), 166–178.

Rock, I. (1983). *The logic of perception*. MIT Press.

Schendan, H. E., & Kutas, M. (2002). Neurophysiological evidence for two processing times for visual object identification. *Neuropsychologia, 40*(7), 931–945.

Schendan, H. E., & Kutas, M. (2003). Time course of processes and representations supporting visual object identification and memory. *Journal of Cognitive Neuroscience, 15*(1), 111–135.

Schendan, H. E., & Kutas, M. (2007). Neurophysiological evidence for the time course of activation of global shape, part, and local contour representations during visual object categorization and memory. *Journal of cognitive neuroscience, 19*(5), 734–749.

Schendan, H. E., & Lucia, L. C. (2010). Object-sensitive activity reflects earlier perceptual and later cognitive processing of visual objects between 95 and 500 ms. *Brain Research, 1329*, 124–141.

Scott, L. S., Shannon, R. W., & Nelson, C. A. (2005). Behavioral and electrophysiological evidence of species-specific face processing. *Cognitive, Affective, & Behavioral Neuroscience, 5*(4), 405–416.

Sehatpour, P., Molholm, S., Javitt, D. C., & Foxe, J. J. (2006). Spatiotemporal dynamics of human object recognition processing: An integrated high-density electrical mapping and functional imaging study of "closure" processes. *Neuroimage, 29*(2), 605–618.

Shen, J., & Reingold, E. M. (2001). Visual search asymmetry: The influence of stimulus familiarity and low-level features. *Perception & Psychophysics, 63*(3), 464–475.

Singmann, H., Bolker, B., Westfall, J., Aust, F., Ben-Shachar, M., Højsgaard, S., Fox, J., Lawrence, M., Mertens, U., Love, J., & others. (2021). *afex: Analysis of factorial experiments (0.28-1) [computer software]*.

Smith, M. E., & Loschky, L. C. (2019). The influence of sequential predictions on scene-gist recognition. *Journal of Vision, 19*(12), 14–14.

Snell, J., & Grainger, J. (2017). The sentence superiority effect revisited. *Cognition, 168*, 217–221.

Solomon, R. L., & Postman, L. (1952). Frequency of usage as a determinant of recognition thresholds for words. *Journal of Experimental Psychology, 43*(3), 195.

Stojanoski, B., & Cusack, R. (2014). Time to wave good-bye to phase scrambling: Creating controlled scrambled images using diffeomorphic transformations. *Journal of Vision, 14*(12), 6–6.

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

Torralbo, A., Walther, D. B., Chai, B., Caddigan, E., Fei-Fei, L., & Beck, D. M. (2013). Good exemplars of natural scene categories elicit clearer patterns than bad exemplars but not greater BOLD activity. *PloS One*, 8(3), e58594.

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

Vo, M. L. H., & Wolfe, J. M. (2013). Differential electrophysiological signatures of semantic and syntactic scene processing. *Psychological Science, 24*(9), 1816–1823.

Walther, D., & Koch, C. (2006). Modeling attention to salient protoobjects. *Neural Networks, 19*(9), 1395–1407.

Wang, Q., Cavanagh, P., & Green, M. (1994). Familiarity and pop-out in visual search. *Perception & Psychophysics, 56*(5), 495–500.

Watson, A. B., & Pelli, D. G. (1983). QUEST: A Bayesian adaptive psychometric method. *Perception & Psychophysics, 33*(2), 113–120.

Witt, J. K., Proffitt, D. R., & Epstein, W. (2004). Perceiving distance: A role of effort and intent. *Perception, 33*(5), 577–590.