Journal of Experimental Psychology: Animal Behavior Processes

Perceptual Learning and Inversion Effects: Recognition of Prototype-Defined Familiar Checkerboards

Ciro Civile, Di Zhao, Yixuan Ku, Heike Elchlepp, Aureliu Lavric, and I. P. L. McLaren

Online First Publication, December 23, 2013. doi: 10.1037/xan0000013

CITATION

Civile, C., Zhao, D., Ku, Y., Elchlepp, H., Lavric, A., & McLaren, I. P. L. (2013, December 23). Perceptual Learning and Inversion Effects: Recognition of Prototype-Defined Familiar Checkerboards. *Journal of Experimental Psychology: Animal Behavior Processes*. Advance online publication. doi: 10.1037/xan0000013

Perceptual learning and inversion effect:

Recognition of prototype-defined familiar checkerboards.

Ciro Civile[1], D. Zhao[2], Y. Ku[2], H. Elchlepp[1], A. Lavric[1] and I.P.L. McLaren[1]

[1]School of Psychology, College of Life and Environmental Sciences,

University of Exeter, UK

[2]School of Psychology and Cognitive Science,

East China Normal University, Shanghai, China

All correspondence concerning this article should be addressed to:

I.P.L. McLaren

Washington Singer Laboratories

University of Exeter

Perry Road

Exeter EX4 4QG

i.p.l.mclaren@exeter.ac.uk

Abstract

The face inversion effect is a defection in performance in recognizing inverted faces compared to faces presented in their usual upright orientation typically believed to be specific for facial stimuli. McLaren (1997) was able to demonstrate that 1) An inversion effect could be obtained with exemplars drawn from a familiar category, such that upright exemplars were better discriminated than inverted exemplars; and 2) That the inversion effect required that the familiar category be prototype-defined. He also provided some evidence that this effect would generalize to a recognition paradigm by showing that it could be demonstrated in a same / different judgment task, and there was an indication that the inversion effect was composed of two separable components, an advantage for upright, familiar exemplars, and a disadvantage for inverted familiar exemplars, although this latter effect was only significant in one of two experiments. In this paper we replicate and extend these findings. We show that the inversion effect can be obtained in a standard old/new recognition memory paradigm, demonstrate that it is contingent on familiarization with a prototype-defined category, and establish that the effect is made up of two components. We confirm the advantage for upright exemplars drawn from a familiar, prototype-defined category, and show that there is a disadvantage for inverted exemplars drawn from this category relative to suitable controls. We also provide evidence that there is an N170 ERP signature for this effect. These results allow us to integrate a theory of perceptual learning originally due to McLaren, Kaye and Mackintosh (1989) with explanations of the face inversion effect, first reported by Yin (1969).

Research key words: Perceptual learning; inversion effect; N170; face recognition; associative learning; expertise.

Face recognition is perhaps one of the most robust cognitive skills that individuals possess. We generally recognize faces with very little effort, despite large variations in lighting, viewpoint and expression. Discussion of the nature of face perception has been mainly divided between two interpretations: One of which makes the assertion that a large body of research supports the notion of specialized mechanisms used to process facial stimuli (Farah Tanaka and Drain, 1995; Valentine, 1988; Yin, 1969): Whilst the other points out that in the last thirty years there have been many studies showing that face recognition is actually based on general mechanisms that can also operate for other non-facial stimuli as well (Diamond & Carey, 1986; Tanaka and Farah, 1991). This paper attempts to adduce evidence for this latter position, by showing that inversion effects can be obtained under certain conditions with checkerboards. It then seeks to establish exactly what these conditions are, and to analyze the nature of the inversion effect obtained, so that the parallel with the face inversion effect can be better evaluated and to help characterize the mechanisms involved. We start with a brief review of some of the evidence for the expertise-based account of the FIE, and then report the results of four experiments that bear on these issues.

One the most reliable phenomena in the study of face perception, and central to this debate, is the face-inversion effect (FIE), which is the disproportionate drop in recognition performance for upside down (inverted) faces relative to upright faces (Yin, 1969). On its discovery, the FIE was described as a clear consequence of the specialized mechanisms used in face processing. This explained why the impairment in recognizing upside down faces was significantly larger than that for other objects (Yin, 1969). However, this interpretation has been challenged by the finding that an FIE as large as the one for faces can be also obtained with images of dogs (Diamond and Carey, 1986), which suggests that the face inversion effect may not be due to the fact that facial stimuli are subject to special processing because

they are facial in nature, but instead that there are other factors, such as expertise, which give rise to this effect.

Diamond and Carey suggested that large inversion effects are only obtained if three conditions are met. First, the members of the class of stimuli must share a basic configuration, the prototype. Second, it must be possible to individuate the members of the class by means of second-order information. And third, individuals must have the expertise (in other words, the experience with the stimuli) to exploit this configural information (spatial relationships between the main features within a face for example). Several additional studies have used Diamond and Carey's approach to explain robust cognitive effects such as the Thatcher illusion (Thompson, 1980), and other studies have investigated the role of feature and configural information under inversion, investigating the importance these types of information have in determining the FIE (Searcy and Bartlett 1996; Leder and Bruce, 1998; Leder, Huber and Bruce, 2001).

Tanaka and his colleagues (Tanaka & Farah, 1993; Tanaka & Sengco, 1997) asked participants to recognize parts of face images or control stimuli (scrambled faces, inverted faces, or houses) in a forced-choice procedure in which the parts were shown: **i)** in their original configuration (Jim's nose in Jim's face), **ii)** in a distorted configuration (Jim's nose in Jim's face with his eyes shifted apart), or **iii**) in isolation (Jim's nose alone). If participants used simple feature-based processing to perform the task, there should not have been any differences between the three stimulus' conditions. The results, however, showed that features of upright faces were better recognized in their original configuration than in the other two conditions.

Gauthier and Tarr (1997) showed that this advantage for a configuration (better feature identification in the original than in the distorted configuration) could also be obtained with Greebles (an artificial category) once the participants became experts with these stimuli. In

addition Gauthier and Tarr (1997) as well as Tanaka et al. (1997) showed that experts could display a whole/part advantage with stimuli such as Greebles, cars or cells. Tanaka and Gauthier (1997) suggested that, although novices may sometimes rely on first-order relational information (for a facial stimulus an example of this would be the nose in relation to the mouth) only experts seem to rely on second-order relational information (the variations in first-order relations relative to the prototype for that stimulus set).

In an evaluation of the expertise account proposed by Diamond and Carey (1986), Robbins and McKone (2007) attempted to replicate the results obtained by Diamond and Carey (1986, Experiment 3) using dog experts. Thus the experimental procedure used was very similar to that used before, however, the authors ensured that the experts could not name or access any other specific information for any of the dog images contained in the stimulus set. This was because in Diamond and Carey's (1986) study the dog images were taken from the archives of the American Kennel Club, and the experts were American Kennel Club judges. Thus, it could have been that some of the experts had seen some of those dogs before they entered the experiment. The results from Robbins and McKone's (2007) Experiment 1 showed that the inversion effect for dog experts with dog images was smaller than that for faces, in contradiction to what Diamond and Carey (1986) had previously found. Thus, the authors suggested that Diamond and Carey's (1986) results might have been influenced by the pre-experimental familiarity that some of the experts might had for some of the dogs in their upright orientation. This initial familiarity might have assisted experts with these stimuli in comparison to inverted dog images and to unfamiliar faces. Clearly Robbins and McKone's experiment raises the possibility that there is some special processing mechanism for faces, whilst leaving open the possibility of an expertise-based contribution to this effect.

Some of the strongest (and earliest) support for the expertise account of face inversion is reported in McLaren (1997). In that study the author demonstrated that inversion effects

similar to those found with faces could be obtained with mono-orientated artificial categories (that participants had never seen before entering the laboratory) once the participants in the experiment were familiarized with those categories. He generated checkerboard exemplars that defined a category, by starting with a randomly generated base pattern (that had 50% black squares and 50% white), and then randomly changing black squares to white and white squares to black. Participants were trained first of all to categorize these stimuli, by requiring them to learn, by trial and error, to distinguish between exemplars derived from one base pattern or prototype (A), and exemplars from another randomly generated base pattern (B) that was constrained to share exactly 50% of it's squares with A. This ensured that participants paid careful attention to the exemplars they encountered, whilst in no way encouraging them to differentiate between members of the same category, only between members of different categories. Nevertheless, on a subsequent discrimination task that involved new exemplars from these familiar categories, participants demonstrated an enhanced ability to learn to distinguish between upright exemplars from the familiar category (relative to performance on exemplars drawn from a novel category) as predicted on the basis of the results obtained by McLaren, Leevers and Mackintosh (1994), who observed a similar effect. This advantage for upright exemplars was also accompanied by a disadvantage for the inverted exemplars (again relative to control stimuli taken from a novel category) in Experiment 1 of this paper, but this result was not replicated in Experiment 2 (though the numerical effect was in the same direction). The net consequence was demonstration of a strong inversion effect in both experiments for the exemplars (which were themselves novel) drawn from the familiar category, in the absence of any such effect for exemplars drawn from a novel category derived from different base patterns that otherwise possessed the same type of stimulus structure.

McLaren was able to explain the advantage for upright exemplars from the familiar category in terms of perceptual learning. Exposure to exemplars from the familiar category led to greater within-category discriminability for exemplars from that category, as in McLaren, Leevers and Mackintosh (1994). The mechanism for this effect was taken to be that proposed in McLaren, Kaye and Mackintosh (1989), namely the differential latent inhibition of common elements. Exposure to exemplars from a prototype-defined category will, on this theory, lead to profound latent inhibition of the prototypical elements for that category. Once this has occurred, when an exemplar drawn from that category is encountered, the elements that it shares with the prototype (and there will be many of them) will be latently inhibited, making them relatively less salient. The elements (what McClelland and Rumelhart, 1985 call micro-features) that are unique to that exemplar will not have been encountered very often, will not suffer greatly from latent inhibition, and so will be relatively salient. As it is these features or elements that allow one exemplar to be discriminated from another (on this account it is the prototypical features that constitute the common elements that make exemplars confusable), discrimination between exemplars drawn from the familiar category will be enhanced.

This mechanism gives a good account of the advantage enjoyed by the upright exemplars, and it can explain the inversion effect by simply pointing out that this mechanism only applies to what it has been experienced, and participants have not experienced inverted exemplars during the earlier familiarization phase. It also predicts that if the category is not prototype-defined, but instead designed to a) allow categorization on the basis of similarity, but b) in such a way that the differential latent inhibition of common elements mechanism cannot gain any traction, then no inversion effect should be obtained, and this was found to be the case in Experiment 1b. Thus, these results provided an account of inversion effects for familiar prototype-defined categories that promised to generalize to other cases such as faces,

in line with the expertise-based explanation of the face inversion effect (FIE) pioneered by Diamond and Carey (1986).

As matters stand, however, there are some obstacles in integrating this theory of perceptual learning and these results with our understanding of the FIE. The first is that none of McLaren's (1997) experiments used the standard old/new recognition memory paradigm that is typically used to demonstrate the face inversion effect. To make the connection between inversion effects with this artificial category and faces, we need to demonstrate that, after exposure, an inversion effect can be obtained with checkerboards drawn from the familiar category in this standard paradigm. We also need to confirm that any such effect depends on the structure of the category in question, as this would be a result of considerable theoretical significance that would rule out a number of other accounts of the role of expertise in producing this phenomenon[1]. Another lacuna that needs to be addressed is whether or not the disadvantage observed for inverted exemplars taken from the familiar category is reliable. If it is, then this result would be of considerable significance. It would establish that the inversion effect had two components (rather than being solely due to an advantage for upright exemplars) something that could not easily be demonstrated by any other means given the difficulty in establishing the appropriate control baseline for comparison with faces for example. If both these issues could be addressed, then this would enhance our understanding of perceptual learning and inversion effects, and might also improve our understanding of the basis for the FIE. Experiment 1a set out to do just this.

[1] An example of such a theory would be one that claimed that simple familiarization with a category, subsequently enhanced attention to exemplars from that category, which improved later recognition performance on those exemplars.

Experiment 1a

Materials

The stimuli were 16 x 16 checkerboards containing roughly half black and half white squares. Four prototypes were randomly generated with the constraint that they shared 50% of their squares with each of the other prototypes, and were 50% black squares and 50% white. Exemplars were generated from these prototypes by randomly changing squares as described in Figure 1.

Participants

32 students at the University of Exeter took part in the experiment. The study was counterbalanced across participants by splitting them into eight groups.

Figure 1 about here please

Procedure

The study consisted of a 'categorization phase', a 'study phase', and an old vs. new recognition 'test phase'. In the categorization phase, the subjects were instructed that once they pressed any key on the keyboard, a set of checkerboard stimuli would appear on the screen, one at a time in a random order. Their task was to sort these stimuli into two categories by pressing one of the two keys ("x" or "."), and they would get immediate feedback as to whether their response was correct or not. If they did not respond within 4 seconds, they would be timed out. The presentation of each stimulus was signaled by a warning cue (a fixation cross in the centre of the screen) presented for 1 sec. Each participant was shown 128 exemplars drawn from two different, prototype-defined categories, with 64

exemplars in each category. Subjects were encouraged to scan the whole of each checkerboard before categorizing it.

In order to counterbalance our stimuli, we used 8 participant groups. The first 4 of those were presented, during the categorization task, with 64 exemplars drawn from category A and 64 exemplars drawn from category B. The second 4 were presented with 64 exemplars drawn from each of the C and D categories. After the categorization phase concluded, participants proceeded to the study phase. For each participant, the task was to look at a number of new exemplars (i.e. exemplars not used during categorization) from one of the two familiar categories seen in the categorization task, plus novel exemplars from a category not previously encountered. Thus, for example, Participant Group 1 was presented with a set of stimuli that included 32 exemplars (16 upright and 16 inverted) drawn from category A (the familiar category for these participants) and 32 exemplars (16 in each of the two orientations) drawn from category C (which was novel for them). To counterbalance this, Participant Group 5 was presented with 32 exemplars (16 upright and 16 inverted) drawn from category C (familiar) and 32 exemplars (16 in both orientations) drawn from category A (novel for that group).

In the study phase each participant was shown with 4 types of exemplars each containing 16 stimuli giving a total of 64 exemplars. These were presented one at a time at random for 3 seconds. The study conditions were: Familiar Inverted exemplars (by which we mean exemplars that had been drawn from a familiar category and then inverted), Familiar Upright exemplars (exemplars drawn from the same familiar category without any inversion), Novel Inverted exemplars (by which we mean exemplars that had been drawn from a novel category and then inverted), and Novel Upright exemplars (exemplars that had been drawn from the same novel category and were not inverted). Following the study phase, participants were given an old/new recognition task. This involved the 64 exemplars seen in the study

phase (32 in an upright and 32 in an inverted orientation, as presented in the study phase), plus 64 new exemplars (32 in an upright and 32 in an inverted orientation) split across the same four conditions used in the study phase. Each participant saw the stimuli corresponding to their participant group in a random order. Participants in the test phase were asked to press "." on the computer keyboard if they had seen that checkerboard before in the study phase, or "x" if they had not seen it, and had 4 seconds in which to do so. Data was collected on accuracy and latency for recognition performance across the test recognition phase.

Results

In all the experiments reported in this paper, the analysis of the response latencies does not add anything to the analysis of the accuracy scores (which were our primary measure), and nor do they imply any speed/accuracy trade-off that might complicate our interpretation of the accuracy data. Following McLaren (1997), we expected an inversion effect (higher score for upright than for inverted) for exemplars drawn from the familiar category, no inversion effect for those from the novel category, and a significant difference between the effects of inversion for familiar and novel categories (i.e. an interaction). We also expected performance on upright exemplars from the familiar category to be better than on those drawn from the novel category, and inverted exemplars drawn from a familiar category to be worse than their controls based on the 1997 Experiment 1 results. In all the behavioral experiments reported here the statistical tests are one-tailed with an alpha of .05 unless otherwise noted, as in most cases we have a clear basis for predicting the direction of effect expected. We give the relevant F ratios and MSE, or the t value and the standard error for the effect tested. Simple effects analyses are uncorrected for multiple comparisons, and are also those used in McLaren (1997). We also offer an estimate of effect size based on Cohen's d for all comparisons that are $p<.1$ or better. This is computed using the pooled variance formula employing the between-subject variances (even when the comparison itself

uses a paired-samples test, as by doing this we avoid inflation of the effect size due to the correlation between samples), and we also give the appropriate confidence interval.

For completeness, the mean latencies for each condition in this experiment were (in msec): Familiar Upright = 2820; Familiar Inverted = 2817; Novel Upright = 2802; Novel Inverted = 2769. The data from all 32 participants were used in the signal detection *d'* analysis of the test phase where a *d'* of 0 indicates chance level performance. In the categorization phase, the mean percentage correct was 67% (but note that this is a figure across the entire 128 trials of trial and error learning, and that the purpose of this phase is to expose participants to the stimuli, we are not especially concerned about categorization accuracy). As predicted, ANOVA revealed a significant interaction between category type and orientation, $F(1, 31) = 3.64$, MSE = 0.170, p = .032, d = 0.53, 95% CI = 0.35, 0.71. Figure 2 gives the results from the test phase. As expected, a significant difference in *d'* was found for the upright versus inverted familiar category exemplars, $t(31) = 1.895$, SE = 0.098, p = .033, d = 0.40, 95% CI = 0.24, 0.56. No significant inversion effect was found for novel category exemplars, $t(31) = 1.08$, SE = 0.83, p = .29. To explore these results further the effect of category type on the recognition of upright exemplars was also analyzed by means of planned comparisons on *d'* scores. Familiar upright exemplars were not recognized significantly better than unfamiliar upright exemplars, $t(31) = 1.441$, d = 0.35 lower limit = 0.19 upper limit = 0.52, SE = 0.115, p =.08, though there was a clear trend in that direction. There was also a non-significant trend for familiar inverted exemplars to be worse than novel ones, $t(31) = 0.963$, SE = 0.116 , p = .171[2].

We also performed a complementary Bayesian analysis on our data to provide additional information on the extent to which they provide evidence that either increases or

[2] The reason why we compare the upright checkerboards from familiar and novel categories, and inverted exemplars in the same way, is that, because of the counterbalancing, these sets of stimuli correspond to one another across participants, thus controlling for random variations in difficulty due to fluctuations in stimulus similarity. In other words, the set of Upright Familiar exemplars for one participant are the Upright Novel exemplars for another etc.

decreases our confidence in the effects reported in the 1997 paper. Our assumptions in running these analyses were that 1) the direction of the effect could be specified and would be that observed in the 1997 paper, and 2) that the d' differences were normally distributed with a standard deviation corresponding to the differences observed in the most comparable study from the 1997 paper. We felt that this was the right approach given that the stimuli were generated according to the same principles as used in 1997, but the procedures used in this experiment are rather different to those employed by McLaren (1997). Using these assumptions, we employed the Bayes factor calculator provided by Dienes (2011) with a half-normal distribution, and this gave a Bayes factor of 3.18 for the contrast between upright familiar and inverted familiar category exemplars, confirming that we can be confident of this inversion effect. The Bayes factor for the comparison between upright familiar and upright novel exemplars was 1.75, again indicating that we have more evidence (though we could not describe it as compelling at this point) for the 1997 effect. The Bayes factor obtained by contrasting novel inverted and familiar inverted exemplars (1.35) suggests that we have some fairly weak additional evidence for this effect as well. We will continue to provide Bayes factors for these effects as we present the evidence from the studies reported in this paper, in order to allow a judgment to be made on whether the 1997 results are confirmed by the current set of experiments.

Figure 2 about here please

Discussion

We have replicated the original McLaren (1997) effect with checkerboards, but this time using exactly the same recognition paradigm as is normally used for face recognition studies. There is a significant inversion effect for the familiar checkerboards, but no inversion

effect for novel checkerboards. Numerically the results of this experiment are entirely in line with those of Experiment 1a in the 1997 paper, but the advantage for upright exemplars just fails to reach significance here, and the disadvantage is also unreliable. Nevertheless, these results do increase our confidence in the first of these effects, and do not undermine our confidence in the second (whilst not providing much additional support for it either). Before going on to consider these issues further, we first report a replication of Experiment 1b from the 1997 paper using the current recognition memory procedures, so that we are in a position to make as full a comparison as possible between our results now and the data reported then.

Experiment 1b

The same procedure was used again, but this time there was an alteration in the method used to generate the stimuli. We used a variant of the algorithm for generating 'shuffled' stimuli outlined in McLaren (1997), as this produced stimuli that were as easy to classify as prototype-defined stimuli, but they did not average to the base pattern used to generate them, and so did not, as a class, possess a prototype themselves. In McLaren (1997) all 16 rows of a base pattern were randomly re-ordered to create each exemplar, which guaranteed there was not prototype for that category. The result was no reliable inversion effect, and the inversion effect obtained with prototype-defined categories was significantly larger than the inversion effect with shuffled categories, confirming that this effect depended on familiarity with a prototype-defined category. This time we used a restricted version of this algorithm in which exemplars were constructed by performing a random permutation of 3 horizontal lines of a base pattern (we used the prototypes from Experiment 1a) to give an exemplar of that category. We only shuffled three rows, to keep the number of squares that (on average) changed the same as in Experiment 1a. The procedure was that two rows were identified at random and swapped, and then a new row was identified, and swapped with one

of the previous two. The result was that on average half the squares in each of the three rows would be altered, making 24 in all. Thus, this experiment is, in some sense, the control for Experiment 1a, though as will become apparent the ease of classification produced by using these materials more nearly matched that of Experiment 2. Because this algorithm was based on the McLaren (1997) shuffling algorithm, we predicted no inversion effect for either familiar or novel category exemplars, and predicted an interaction with Experiment 1a similar to that in McLaren (1997). We return to this point in the discussion.

Participants

32 students at the University of Exeter took part in the experiment. The study was counterbalanced, as in Experiment 1a, by splitting participants into 8 groups.

Results

The data from all 32 subjects was used for the analysis. In the categorization phase, the mean percentage correct was 77%. Thus, as we predicted, the stimuli were at least as easy to categorize as stimuli in Experiment 9a. Equally however, as Figure 3 suggests, there was no significant difference in d-prime means for familiar category exemplars, or for novel category exemplars, confirming our predictions. The crucial interaction, of course, is not that within Experiment 1b, but emerges when we compare the results of Experiment 1b with those of Experiment 1a. In McLaren (1997) a similar comparison showed that the inversion effect, defined as the familiarity by orientation interaction, obtained with exemplars drawn from a prototype-defined category, was significantly greater than that obtained with exemplars drawn from a category that was not defined by a prototype. If we compute a 2x2x2 ANOVA, to find the Experiment by Familiarity by Orientation interaction for Experiments 1a and 1b, we find a marginally significant result, $F(1, 62) = 2.721$, MSE = 0.215, p = .052, d = 0.50, 95% CI = 0.27, 0.91, suggesting that this might be the case here as well. Thus, we have some supporting evidence enabling us to claim that the inversion effect depends on both the

category being familiar and being based on a prototype. Calculating the Bayes factor for this analysis based on the 1997 priors gives a Bayes factor of 2.50, also suggesting that there is good, but not conclusive evidence for this effect.

Figure 3 about here please

Discussion

Experiment 1b did not produce any of the effects observed in Experiment 1a, despite our controlling for the number of squares changed to produce the exemplars and ensuring that the categorization task was at least as easy in 1b as it was in 1a. Our argument is that this is because the different algorithm for generating exemplars produces a category with a different structure, one in which the influence of any prototype is considerably weaker. We defer a detailed discussion of how the theory of perceptual learning in McLaren, Kaye and Mackintosh (1989), which was further developed in McLaren and Mackintosh (2002), and McLaren, Forrest and McLaren (2012) can explain these effects until the General Discussion later on, because our first priority is to establish this theoretically important result beyond reasonable doubt. Experiments 1a and 1b were a replication of Experiment 1 from McLaren (1997) using our recognition memory paradigm; we now varied our stimuli and experimental design so as to strengthen our effect size and eliminate some possible alternative explanations of this result.

The first issue we tackled was that the inversion effect in Experiment 1a was not as substantial as we would like, and we speculated that this may be because participants found it too hard to recognize the checkerboards in this experiment. Because of this, performance on all the different types of checkerboard became too close to floor to make it possible to decide if inverted checkerboards from the familiar category were actually harder to recognize than

inverted checkerboards from the novel category, i.e. to detect whether the disadvantage for inverted checkerboards drawn from a familiar category reported in McLaren (1997) was real. Experiment 2 aimed to address this issue.

Experiment 2

Experiment 2 was a replication of Experiment 1a, but this time we tried to make the checkerboards "clumpier", with the intention of making the stimuli easier to recognize (see Figure 4). We hoped to obtain a stronger inversion effect for familiar checkerboards than the one obtained in Experiment 1a.

Materials

In this experiment a randomly chosen 96 squares (up from the 48 used in Experiment 1a) were set at random to generate each exemplar from the base prototype, and the prototypes themselves had stronger differentiation into black and white areas (see Figure 4). This was accomplished by making the probability of a single square being a particular color depending on the color of its neighbors, so that if they were predominantly black, then that square had a greater chance of being black, and vv. for white. The result was a set of prototype patterns that were still 50% black and 50% white, and still overlapped 50% with one another, but with the squares of a particular color clumped together.

Figure 4 about here please

Participants

32 students at the University of Exeter took part in the experiment. The study was counterbalanced, as in Experiment 1a, by splitting participants into 8 groups.

Procedure

This was exactly the same as that used in Experiment 1a.

Results

For completeness, the mean latencies for each condition in this experiment were (in msec): Familiar Upright = 2787; Familiar Inverted = 2804; Novel Upright = 2796; Novel Inverted = 2860. The data from all 32 subjects was used in the analysis. In the categorization phase, the mean percentage correct was 77%, indicating that our manipulation of the stimuli had made them easier to classify compared to Experiment 1a (and identical, in terms of ease of classification, to Experiment 1b). Results from ANOVA once again showed a significant interaction between category type and orientation, $F(1, 31) = 4.13$, MSE = 0.182, p =.025, d = 0.50, 95% CI = 0.28, 0.71. Figure 5 gives the results for the mean d' score by stimulus type. Planned comparisons were used to examine whether or not there was a significant inversion effect for familiar category exemplars. A reliable difference in d' emerged for the upright versus the inverted familiar category exemplars, $t(31) = 2.845$, SE = 0.095, p = .004, d = 0.65, 95% CI = 0.50, 0.79. No significant inversion effect was found for novel category exemplars, $t(31) = 0.284$, SE = 0.119, p = .77. To explore this further the effect of category type on the recognition of upright exemplars was also analyzed. Familiar upright exemplars were not recognized significantly better than unfamiliar upright exemplars, $t(31) = 0.977$, SE = 0.091, p=.17, but novel inverted exemplars were recognized better than familiar inverted exemplars, $t(31) = 2.030$, SE = 0.106, p =.025, d = 0.49, 95% CI = 0.34, 0.64.

Figure 5 about here please

We also computed Bayes factors for these contrasts using priors based on Experiment 1a. The Bayes factor for the inversion effect in the familiar category was 20.88, indicating that we can now have a great deal of confidence in this finding. For the novel category, the Bayes factor for the inversion effect was 0.25, indicating that the evidence supports the null hypothesis of no effect. The interaction for these two effects has an associated Bayes factor of 4.47, suggesting that we can also be confident that the inversion effect in the familiar category is bigger than that in the novel category. The Bayes factor for the contrast between upright exemplars from the familiar and novel categories is 1.15, suggesting that we have no decisive evidence on this effect, but that for comparison of the inverted stimuli is 3.74, indicating that we are now in a position to be confident that performance in the recognition task on the inverted exemplars drawn from the novel category is superior to that on the inverted exemplars drawn from the familiar category.

Discussion

Experiment 2 replicated and strengthened the findings obtained in Experiment 1a. We were able to increase the size of the inversion effect for checkerboards drawn from a familiar category by making them easier to recognize. Thus, our results confirm that an inversion effect can be obtained after familiarizing participants with a prototype-defined category. We are also able to comment further on the basis for this effect. The trend for familiar upright exemplars to be better recognized than novel upright exemplars was not significant (though the effect was once again numerically in the same direction), but this time inverted familiar exemplars were significantly worse recognized than novel inverted exemplars. Our analysis is that the basis of the inversion effect obtained with prototype-definedcategories may well be, in part, due to some advantage for the upright exemplars from the familiar category, but also has a component due to the disadvantage suffered by inverted exemplars from a familiar category. The former effect was significant in McLaren's (1997) results, and though not

independently significant in the two relevant studies reported so far, the combined Bayes factor for this effect is now over 2. Thus, our evidence on this point, whilst still not compelling, is in line with that in McLaren (1997). If we perform a similar analysis for the disadvantage accruing to inverted exemplars drawn from a familiar category and combine the two experiments we get a Bayes factor of over 5, suggesting we can have confidence in this effect. This strongly suggests that the effect observed by McLaren (1997) in one of his experiments was real.

With this in mind, if we look back at Yin's (1969) study, and focus on his Experiment 1, we note that he found that normal faces in an inverted orientation were more difficult to recognize than other inverted sets of stimuli. A major criticism that can be (and was) leveled at this finding is that the sets of stimuli used in Yin's (1969) study i.e. for example pictures of houses or airplanes, did not match normal faces in terms of structure and the information contained in the stimuli. For example, the number of features and the complexity of their configurations varied widely across these stimuli. However, given that in our experiments we have found a substantial disadvantage for inverted checkerboards taken from a familiar category, this suggests that Yin's (1969) finding reflected some underlying reality, despite the lack of control in his experiment.

Experiment 3

Our conclusions from Experiments 1a and 1b depend, in part, on the cross-experiment comparison between them. The aim of Experiments 1a and 1b was to replicate McLaren's (1997) result, and demonstrate that the inversion effect requires that the exemplars belong to familiar categories defined by a prototype, and will not occur for rather similar sets of stimuli that lack this property. But, as matters stand, it could be that our result (and McLaren's original result) was due to the two experiments drawing on different populations by virtue of

their being run at different times. We would also concede that the effect size in these experiments was not large, and that this is probably due to using the less "clumpy" checkerboards employed by McLaren (1997). Thus, in the next experiment (Experiment 3) the logical step was to run a replication of Experiments 1a and 1b, but this time using the type of checkerboards used for Experiment 2, and to ensure that the two conditions were run together within the same experiment.

Materials

For Experiment 3 the categories of checkerboards used in were the same as used in Experiment 2. Thus, we already had the four sets of exemplars generated from Category prototypes A, B, C and D. Our next task was to construct matched sets of shuffled exemplars. Recall that in Experiment 1b, to change 24 squares (on average) we shuffled three rows. First two were selected and swapped (let's call them 1 and 2), then a new row was selected (call it 3) and 2 and 3 were swapped. This meant that a maximum of 48 squares could change, and on average 24 would. In what we will call the New Shuffled exemplars used in Experiment 3b, on average 48 squares are changed, with a maximum of 96 possible, in order to match the similarity structure of the prototype-defined sets of exemplars. This was done by iterating the procedure employed in Experiment 1b. Given that each time we selected another row, we changed (on average) another 8 squares, 6 rows had to be altered in total to change an average of 48 squares.

Participants

48 students at the East China Normal University took part in the experiment.

Procedure

The procedure within the different experimental conditions was exactly the same as in the previous experiments reported in this paper. Thus, the participants were presented with a

categorization task, followed by a study phase and finally there was an old/new recognition task. However this time 24 of the participants performed the task with prototype-defined checkerboards, and 24 participants performed the task with shuffled checkerboards. The participants in each group were matched in terms of the stimuli used (i.e. in the allocation of prototypes to conditions for that participant). The only difference is in how the exemplars were generated from those prototypes.

Results

The data from all 48 subjects was used in the analysis. The results from Experiment 3a show that in the categorization phase, the mean percentage correct was 82%, confirming what was already obvious in Experiment 2, thus that our construction of the stimuli had made them easier to classify compared to Experiment 1a. The mean latencies for each condition in this experiment were (in msec): Familiar Upright = 2785; Familiar Inverted = 2851; Novel Upright = 2820; Novel Inverted = 2720. ANOVA on the test d' revealed a significant interaction between category type and orientation, $F(1, 23) = 5.389$, MSE = 0.430, p =.014, d = 0.74, 95% CI = 0.41, 1.08. Figure 6 gives the results for the mean d' score by stimulus type for this experiment. Planned comparisons were used to examine whether or not there was a significant inversion effect for familiar category exemplars. A reliable difference in d' emerged for the upright versus the inverted familiar category exemplars, $t(23) = 3.381$, SE = 0.146, p = .001, d = 0.94, 95% CI = 0.73, 1.15. No significant inversion effect was found for novel category exemplars, $t(23) = 0.659$, SE = 0.192, p = .51. To explore the reliable inversion effect further the effect of category type on the recognition of upright exemplars was also analyzed. Familiar upright exemplars in this experiment were not recognized significantly better than unfamiliar upright exemplars, $t(23) = 0.737$, SE = 0.226, p = .23, but novel inverted exemplars were recognized better than familiar inverted exemplars, $t(23) = 2.295$, SE = 0.198, p =.015, d = 0.73, 95% CI = 0.47, 0.98.

Turning to Experiment 3b, in the categorization phase, the mean percentage correct was 88%. Thus, these stimuli were at least as easy to categorize as the stimuli in Experiment 3a. However, as Figure 6 suggests, there was no significant difference in d-prime means for familiar category exemplars, or for novel category exemplars, confirming our predictions. The 2x2x2 (Experiment by Familiarity by Orientation) ANOVA gave an $F(1, 46) = 3.253$, MSE = 316, p = .039, d = 0.52, 95% CI = 0.07, 0.97. However, in this case because 3a and 3b are actually in the same experiment, and, as mentioned previously, each participant in one group can be matched in terms of the stimuli used (i.e. the allocation of prototypes to condition) to a participant in the other, we can also analyze this result by using a matched samples test. The results of the Experiment (a vs. b) by Familiarity by Orientation interaction using this approach gives a $t(46) = 2.060$, SE = 0.276, p = .022, strongly supporting the claim that the inversion effect depends on both the category being familiar and its being based on a prototype.

Figure 6 about here please

Calculating the Bayes factor for this analysis based on the Experiment 1a and 1b priors now gives a Bayes factor = 6.11, suggesting that there is strong evidence for this effect. The Bayes factor for the inversion effect in the familiar category in Experiment 3a was 55.38, indicating that we can have a great deal of confidence in this finding. For the novel category, the Bayes factor for the inversion effect was 1.12, indicating that the evidence is indifferent with respect to the null hypothesis of no effect. The interaction for these two effects has an associated Bayes factor of 6.92, suggesting that we can also be confident that the inversion effect in the familiar category is bigger than that in the novel category. The Bayes factor for

the contrast between upright exemplars from the familiar and novel categories is 1.17, suggesting that we have no evidence either way on this effect from this result (but the cumulative effect of these Bayes factors being over 1 is starting to tell as we shall see). The Bayes factor for the comparison of the inverted stimuli in Experiment 3a is 3.90, confirming that performance in the recognition task on the inverted exemplars drawn from the novel category is superior to that on the inverted exemplars drawn from the familiar category.

Discussion

Experiment 3 has essentially confirmed our predictions. Thus, the results show that the inversion effect requires that the exemplars are drawn from a familiar, prototype-defined category. Experiment 3 supports the main findings of Experiments 1a and 1b, and of Experiment 1 in McLaren (1997). It also confirms that the sets of checkerboards used in Experiment 2 are easier to recognize and show stronger effects compared to the effect size in Experiment 1a. We defer any further discussion of these effects until we have considered our final experiment, which seeks to investigate the neural correlates of the inversion effect with checkerboards drawn from a familiar, prototype-defined category, now that we are quite sure that this combination of factors is necessary to produce this effect. Our hypothesis is that we will see an effect of inversion in the N170 to the familiar checkerboards, and we begin with a short review of the literature germane to this hypothesis.

Experiment 4

Several studies on face recognition using ERPs have demonstrated a difference in the ERPs to faces and objects at between 150-200 ms in bilateral occipito-temporal regions (Eimer, 2000; Rossion, Gauthier, Tarr, Despland, Bruyer, Linotte, & Crommelinck, 2000). Bentin, Allison, Puce, Perez, McCarthy (1996) investigated the response characteristics of the

N170 using a target detection task in which various pictures of faces and other objects (e.g. flowers, cars) were presented and participants were monitored for the appearance of butterflies (target) in the sequence. In that early study, the N170 response was small for non-face stimuli, as the negative deflection did not cross the zero baseline level of the EEG. Since then, ERP studies have obtained large and clear negative N170 components following the presentation of faces as well as non-face objects (De Haan, Pascalis, & Johnson, 2002; Tanaka & Curran, 2001; Rossion, Gauthier, Goffaux, Tarr, & Crommelinck, 2002). N170 responses have been shown for objects (e.g. houses, chairs, cars) at latencies comparable to those for images of faces, but always at smaller amplitudes.

Rossion et al. (2000) found that the N170 is both increased and delayed when faces were presented after inversion, but that this difference was not obtained for inverted classes of objects for which participants were not experts e.g., shoes. This effect of inversion on the N170 for faces is robust, and it has been obtained in several ERP studies (Rossion, Delvenne, Debatisse, Goffaux, Bruyer, Crommelinck, Guerit, 1999; Rossion et al., 2000; Taylor, McCarthy, Saliba, & Degiovanni, 1999).

In 2002 Rossion et al., investigated whether this inversion effect reflects mechanisms specific to faces, or whether it could be extended to other stimuli as a function of visual expertise. ERPs were recorded in response to upright and inverted images of faces and novel objects (Greebles). The study used 10 subjects before and after 2 weeks of expertise training with Greebles. The N170 component was elicited for both faces and Greebles. The results confirmed the previous findings in the literature, in that the N170 was delayed and larger for inverted faces compared to upright ones at recording sites (T5 and T6) in both hemispheres. The new finding was that Greebles elicited the same effect on inversion, but only for experts and primarily in the left hemisphere. The authors proposed that the mechanisms underlying the FIE on the N170 can be extended to visually homogeneous (share a configuration) non-

face object categories, at least in the left hemisphere, but only when such mechanisms are recruited by expertise. However the main issue that arises from these results is that it could be objected that stimuli such as Greebles are still quite similar to faces; in the way that they share a basic configuration of features which varies, and in that both can be considered mono-oriented stimuli. More recently Busey and Vanderkolk (2005) investigated the effect of visual expertise on the N170 using mono-oriented fingerprint stimuli. In Experiment 2 of their study, upright and inverted images of faces and fingerprints were shown to experts and novices. The N170 component was reliably elicited but somewhat delayed over the right parietal-temporal regions when faces were presented in an inverted orientation, confirming other the findings in the literature. In addition the inverted fingerprints elicited a similarly delayed N170 over the right parietal-temporal region in experts, but not in novices.

Thus, in Experiment 4, we investigated the electrophysiological correlates of our behavioral results, by looking to see if we could find an N170 for our familiar checkerboards that was increased and delayed by inversion. We also noted any left-right localization for any such effect.

Materials

The categories of checkerboards used in Experiment 4 were the same as those used in Experiments 2 and 3a.

Participants

32 students at the University of Exeter took part in the experiment.

Procedure

The experimental procedure was identical to the one used in Experiment 2, except that we doubled up the number of trials to allow for better signal averaging and to obtain a more reliable ERP. To make this possible, the experiment was split into two parts: each

including a categorization task followed by a study phase and an old / new recognition task. Straight after the first part participants were presented with the second part that used a different set of stimuli. The categories of stimuli were counterbalanced across the two experimental parts in such a way that if a category was processed in the first part then that category was not presented again in the second part.

EEG apparatus.

The EEG was sampled continuously at 500 Hz with a bandpass of 0.016-100 Hz, the reference was at Cz and the ground at AFz using 64 Ag/AgCl active electrodes (ActiCap, Brain Products, Munich, Germany) and BrainAmp amplifiers (Brain Products, Munich, Germany). There were 61 electrodes on the scalp in an extended 10-20 configuration and one on each earlobe. Their impedances were kept below 10kΩ. The EEG was filtered offline with a 20 Hz low-pass filter (24 dB/oct) and re-referenced to the linked ears.

EEG analysis.

Peak amplitudes of the N170 in study and recognition phase were examined for differences between the experimental conditions. To improve the estimates of N170 amplitude and latency given the relatively small number of ERP segments in each condition (leading to a low signal-to-noise ratio), N170 extraction was aided by linear decomposition of the EEG by means of Independent Component Analysis (ICA, Bell & Sejnowski, 1995; for the application of ICA for the identification of ERP components, see Debener, Ullsperger, Siegel, Fieler, Von Cramon, & Engel, 2005, and Lavric, Bregadze, & Benattayallah, 2010). ICA is predicated on the assumption that the EEG at each electrode represents a mixture of temporally independent signals (components). It thus attempts to determine the ‘unmixing’ square matrix whose multiplication with the data results in the ‘original’ independent components. The number of entries for each dimension of the unmixing matrix is equal to the number of EEG electrodes, meaning that each row is a spatial filter that ‘unmixes’ one

independent component from the EEG electrode data and the number of recovered components is equal to that of the electrodes. Because the unmixing matrix values relate electrodes to components, they are also referred to as 'ICA weights'. An important aspect of the procedure is what constitutes 'independence'of the extracted components. We used the Infomax version of ICA (Bell &Sejnowski, 1995; implemented in the Brain Analyzer software), which minimizes the mutual information (maximizes entropy) between components. Infomax comprises a neural network algorithm, with the EEG data at each electrode as input, a sigmoidal function of each independent component as output, and the unmixing matrix as the input-output connection weights. The algorithm iteratively adjusts the weights using gradient descent to maximize the entropy (independence) of the output (the components) (see Brown, Yamada, & Sejnowski, 2001).

ICA was run separately for each participant using all scalp channels and the entire dataset (not only the target ERP segments). The resulting ICA components were segmented into 600-ms epochs time-locked to stimulus onset and baseline-corrected relative to the mean amplitude in the 100 ms preceding the stimulus. For analyses of the recognition phase, segments associated with incorrect responses were discarded (there were no responses in the study phase). The remaining EEG segments were averaged for every participant and experimental condition. In each participant, we identified ICA components that: (1) showed a deflection (peak) in the N170 time-range (at 150-200 ms following stimulus onset), and (2) had a scalp distribution containing an occipital-temporal negativity characteristic of the N170 (the scalp distributions of components are the columns of the inverted unmixing matrix). This resulted in 1-4 ICA components corresponding to the N170 identified in most participants (mean 2.6; SD 1) - these were back-transformed into the EEG electrode space (by multiplying the components with the inverted unmixing matrix that had the columns

corresponding to other components set to zero) and submitted to statistical analysis of N170 peak amplitude and latency.

Results

Behavioral results.

The data from all 32 subjects was averaged across the two parts of the experiment and used in the analysis. In the categorization phase, the mean percentage correct was 88%. The mean latencies for each condition in this experiment were (in msec): Familiar Upright = 2920; Familiar Inverted = 2847; Novel Upright = 2877; Novel Inverted = 2893. ANOVA on the d' scores from the test phase showed a significant interaction between category type and orientation, $F(1, 31) = 4.91$, MSE = 0.101, p =.017, d = 0.51, 95% CI = 0.34, 0.68. Planned comparisons were used to examine whether or not there was a significant inversion effect for familiar category exemplars. A reliable difference in d' emerged for the upright versus the inverted familiar category exemplars, $t(31) = 2.988$, SE = 0.083, p=.002, d = 0.76, 95% CI = 0.64, 0.87. No significant inversion effect was found for novel category exemplars, $t(31) = 0.003$, SE = 0.087, p = .498. To explore this further, the effect of category type on the recognition of upright exemplars was also analyzed. Familiar upright exemplars were not recognized significantly better than unfamiliar upright exemplars, $t(31) = 1.179$, SE = 0.076, p = .12, but there was a significant tendency for novel inverted exemplars to be better recognized than familiar inverted exemplars, $t(31) = 1.803$, SE = 0.088, p = .04, d = 0.47, 95% CI = 0.35, 0.48. The upshot of these results is that the behavioral data for this experiment strongly resembles that of our previous experiments when no ERPs were recorded.

A complementary Bayesian analysis using Experiment 2 to generate priors revealed that the Bayes factor for the effect of inversion on familiar checkerboards was 35.74, the effect of inversion on novel checkerboards had an associated Bayes factor of 0.50 and that

the Bayes factor for the interaction was 5.97. The Bayes factor for the comparison between upright stimuli was 1.59, and that for inverted stimuli was 2.86. Thus, the overall Bayes factor (obtained by multiplying the individual Bayes factors from each experiment) for the comparison between familiar upright stimuli and novel upright stimuli is now 3.74 and hence we can be confident that our procedures produce an inversion effect that has a component attributable to an advantage for the familiar upright stimuli; and as the Bayes factor for the comparison between the inverted stimuli now comfortably exceeds 10 (it's actually 56.31), we can be very confident that there is a component attributable to a disadvantage for familiar inverted stimuli. Figure 7 shows the results for the mean d' score by stimulus type, and the pattern is very similar to that obtained in Experiments 1a, 2 and 3a.

Figure 7 about here please

N170 analysis.

N170 latency and amplitude analyses were run on electrode PO7 (Left occipito-temporal site) and on electrode PO8 (Right occipito-temporal site). We report the analysis from the study phase of Experiment 4. This is because significant differences on the N170 were not found in the recognition task; not an entirely unexpected result given that, if the modulation of the N170 reflects an effect of perceptual expertise, then this should occur when simply perceiving the stimulus, and should be easiest to detect during the study phase. This is because the effect should not be tied to having to do anything in particular, except perhaps attend to the stimulus, and by the recognition phase our familiarity manipulation will have been somewhat diluted by experience of all the stimuli in the study phase. Figure 7 shows the N170 recorded during the study phase of this experiment. Table 1 gives latency and

amplitude values for both the study and recognition phases of Experiment 4. All analyses reported are two-tail (because we do not have prior data for ERPs to these stimuli).

Latency analysis on PO7.

There was a trend for the Orientation by Familiarity interaction, F (1, 31) = 2.884, MSE = 65.931, p = .09, d = -0.20, 95% CI = -4.95, 4.51. There was also trend for the N170 to familiar inverted stimuli to peak later than the one for familiar upright stimuli, t(31) = 1.656, SE = 2.302, p = .10, d = 0.31 95% CI = -3.90, 4.52. No significant difference in latency was found for novel stimuli t(31) = 0.419, SE = 2.538, p = .339.

Peak amplitude analysis for PO7.

ANOVA revealed a trend for the main effect of Familiarity, F(1, 31) = 3.730, MSE = 0.563, p = .06, and a main effect of Orientation, F(1, 31) = 13.094, MSE =0.188, p = .001. The difference in peak amplitudes between upright and inverted checkerboards was significantly larger when the stimuli were drawn from a familiar category than from a novel one, Orientation by Familiarity interaction, F(1, 31)= 4.469, MSE = 0.282, p = 0.033, d = 0.61, 95% CI = 0.34, 0.87. The effect of inversion was reliable for familiar categories, t(31) = 3.934, SE = 0.123, p < .001, d = 0.73, 95% CI = 0.50, 0.96; with more negative amplitudes for inverted (-0.558μV) compared to upright (-0.072μV) checkerboards. For novel categories the inversion effect did not approach significance t(31) = 0.574, SE = 0.118, p = .285. Finally, there was a significant difference between novel inverted stimuli and familiar inverted stimuli, t(31) = 2.605, SE = 0.178, p = .014, d = 0.62, 95% CI = 0.36, 0.88; with more negative amplitudes for familiar inverted (-0.558μV) compared to novel ones (-0.093μV).

Latency analysis on PO8.

ANOVA revealed a trend for the Orientation by Familiarity interaction F(1, 31) = 3.619, MSE = 63.123, p = .06, d = -0.35, 95% CI = -.5.54, 4.82. A significant delay in the

N170 was found for familiar inverted checkerboards, with them peaking 6ms later than familiar upright stimuli, $t(31) = 2.539$, SE = 2.215, p = .016, d = -0.43, 95% CI = -.4.01, 4.88. No significant difference in latency was found for novel stimuli $t(31) = 0.093$, SE = 3.014, p = .46.

Peak amplitude analysis for PO8.

ANOVA revealed a main effect of Familiarity, $F(1, 31) = 6.077$, MSE = 0.384, p = .019, and a main effect of Orientation, $F(1, 31) = 7.229$, MSE = 0.370, p = .011. Here as well the difference in peak amplitudes between upright and inverted checkerboards was significantly larger when the stimuli were drawn from a familiar category rather than from a novel one, Orientation by Familiarity interaction $F(1, 31) = 6.66$, MSE = 0.360, p = .015, d = 0.64, 95% CI = 0.34, 0.93. The effect of inversion was reliable for familiar categories, $t(31) = 4.178$, SE = 0.134, p < .001, d = 0.71, 95% CI = 0.44, 0.98; with more negative amplitudes for inverted (-0.557μV) compared to upright (0.005μV) checkerboards. For novel categories the inversion effect did not approach significance $t(31) = 0.094$, SE = 0.165, p = .46. Finally there was a highly significant difference between novel inverted stimuli compared to familiar inverted stimuli, $t(31) = 3.800$, SE = 0.143, p < .001, d = 0.67, 95% CI = 0.40, 0.96; with more negative amplitudes for familiar inverted (-0.557μV) compared to novel inverted stimuli (-0.013μV).

Figure 8 about here please

Table 1 about here please

Discussion

The behavioral results of Experiment 4 confirm that we can obtain a significant inversion effect using stimuli drawn from familiar prototype-defined categories of checkerboards, and that this inversion effect is significantly greater than that for novel categories of checkerboards. Correspondingly, the ERP results show that checkerboards from familiar categories elicit a significant inversion effect in the N170, which is larger than that elicited by checkerboards from novel categories. It would seem, then, that the N170 can serve as a neural signature of the inversion effect obtained with checkerboards drawn from familiar, prototype-defined categories. Hence, we conclude that we have clear evidence of an electrophysiological inversion effect on the N170 for a set of stimuli entirely different from faces, other "natural" categories or Greebles. Additionally, the effect on the N170 found for Greebles was typically limited to the left hemisphere, whereas the analogous effect for faces is usually bilateral; our results show a strong effect of inversion on the N170 for both left and right occipito-temporal sites, providing a good match to the face data. We also note that the inverted checkerboards drawn from the familiar category produce a larger and delayed (in PO8) N170, also in line with the face inversion literature. The final point to make is that upright familiar categories and novel categories in both orientations elicited a similar N170. The real difference in the N170 is between the ERP to inverted checkerboards drawn from familiar category and the other stimuli in this experiment, which suggests that it may be driven by the disadvantage consequent on seeing familiar checkerboards presented upside down that we also see in our behavioral data.

General Discussion

Experiments 1a & 1b and 3a & 3b support the hypothesis that familiarity with a category defined by a prototype leads to an inversion effect in standard recognition

paradigms with novel stimuli drawn from that category, and this does not happen after experience with a category that cannot be defined in terms of a prototype. Before accepting this assertion, however, we must establish that the pattern of performance seen in Experiment 1 was not simply a floor effect. In fact, the data argue against this interpretation. Performance overall in Experiment 1a was only marginally better than chance $F(1, 31) = 2.64$, $p = .057$, confirming the impression that the participants found the task very difficult. Overall performance in Experiment 1b was, however, significantly above chance, $F(1, 31) = 8.00$, $p < .005$, so, if anything, the task with the shuffled checkerboards was easier, and this is consistent with the categorization data as well. It is unlikely, therefore, that the lack of an inversion effect with the shuffled checkerboards is due to a floor effect. Experiments 3a and 3b offer additional reassurance that this result was not due to any artifact induced by near floor performance.

The conclusion that the inversion effect that can be obtained with checkerboards drawn from a familiar category (even though those checkerboards are themselves novel) depends on the category being defined in terms of a prototype has strong theoretical consequences. It immediately invalidates any theory that proposes that the effect is simply consequent on experience with a category to the point that new exemplars can be identified as being from that category. For example, it might be supposed that being able to label an exemplar as a member of category A resulted in more attention being paid to that exemplar, and that this aided subsequent recognition. But this account would also predict a similar effect in the shuffled checkerboards (that were just as easy to classify as the prototype-defined ones) and this was not found to be the case. As far as we are aware, this is the only result of this type in the literature, where what are in some sense the same type of stimuli (they are all checkerboards), which are experienced in the same way, gives rise to such different consequences for later learning and memory. If the previous sentence seems to

overstate our result, consider this: Any participant given just one exemplar from either a shuffled or a prototype-defined category could not tell which type of category it came from in the absence of any further information. In this sense, the two sets of stimuli, shuffled and prototype-defined, act as perfect controls for one another. It is only experience with a set of exemplars that can bring the category structure into play, and here we have the clearest evidence possible that the effects of that experience depend crucially on category structure.

There is, of course, one aspect of the stimulus construction used this time in the case of the shuffled checkerboards that does require further discussion. In the McLaren (1997) experiments, the rows were shuffled completely, and as such the likelihood of any given row remaining in its base position was rather low. This meant that the average of all the shuffled patterns was a set of vertical bands of varying degrees of grey (depending on the proportion of black squares in any given column), and this average was not actually a checkerboard, and so could not be considered as a prototype of the category. In the current experiments we only shuffled three rows (in 1b) and six rows (in 3b), to equate the number of squares changed (on average) with Experiments 1a and 3a, and this means that the chance of a row not being changed from its base position is rather high. Given this, the average of all the shuffled exemplars will now approximate a (somewhat blurry) checkerboard, and the claim that this is no longer a prototype-defined category is harder to sustain. Nevertheless, there is no doubt that the procedures with these stimuli lead to a quite different set of results to those obtained with the standard prototype + noise stimuli used in Experiments 1a and 3a.

A more detailed application of the MKM (McLaren, Kaye and Mackintosh, 1989; further developed in McLaren and Mackintosh, 2000; and McLaren, Forrest and McLaren, 2012) model to these stimuli helps make it clear why this should be so (see also Wills, Suret and McLaren, 2004 for a discussion of these issues in the context of categorisation rather than recognition). Take the stimuli of Experiment 1a first. Starting with a base pattern (the

prototype), 48 squares are randomly chosen and then set to black or white at random to create each exemplar that will, on average, differ by 24 squares from the prototype. Consider a typical changed square in the middle of the stimulus. It will be surrounded by 8 squares that will mostly be those of the base pattern (on average 0.75 of a square of these 8 will have been changed). As a consequence of category pre-exposure, the MKM model tells us that the elements of a stimulus associate to one another, and that this allows them to predict one another, reducing their error scores, and that as a consequence their salience decreases. But, for a changed square, the predictions from the surrounding elements (which as near neighbours we assume will be important predictors for this square) will be incorrect, and so the square will have a relatively high salience because of its relatively high error score. This facilitates discrimination and recognition based on these changed features (which uniquely define the exemplars). In the case of the shuffled stimuli in Experiment 1b and 3b, because a row is moved as a whole, the squares either side of a changed square will be the same as usual for that square, and are good predictors of that square, even though its location in the stimulus has altered. The other surrounding squares are not such good predictors, and hence their influence will be less. The essential difference captured by this analysis is that shuffling rows leaves the changed squares in an exemplar relatively well predicted by other squares nearby, and this acts to mitigate any salience increase that would be gained from location specific prediction effects. Thus, category pre-exposure will not be expected to be that beneficial in the shuffled case, especially if we add in the fact that the location-specific predictions are themselves somewhat degraded by the shuffling process. Our conclusion from this analysis is that, despite the somewhat restricted shuffling algorithm used in the current experiments, the prediction that there should be no perceptual learning, and hence no inversion effect for the shuffled checkerboards in Experiment 1b and 3b still holds. This brings us to the basis for the inversion effect in Experiment1a.

Experiments 2 and 3a and 4 confirm the existence of the inversion effect found in Experiment 1a, and strongly suggest that it is made up of two components. These seem to be an advantage for the upright exemplars from the familiar category, and a disadvantage for the inverted exemplars taken from the familiar category. The explanation of the advantage for the upright exemplars drawn from the familiar category has already been given, but bears some repetition. During categorization, the prototypical elements common to the exemplars of a given category will be routinely exposed, and so will lose salience according to the MKM model. By way of contrast, the elements unique to each exemplar (which the subjects will have less exposure to and will be less well predicted by other elements of the stimulus), will still have relatively high salience. Hence, the structure of this prototype-defined category will ensure that differential latent inhibition of common and unique elements can occur, and this leads to perceptual learning, which in turn leads to an improved ability to recognize upright exemplars of the familiar category, because this depends on using the unique elements of exemplars rather than the ones they share in common. This simply represents an instance of the type of effect reported by McLaren, Leevers and Mackintosh (1994), and also seen in Graham and McLaren (1998). This advantage would be lost on inversion, because participants are not familiar with those exemplars in an inverted orientation, and hence the unique elements of an exemplar would no longer enjoy any salience advantage over the elements common to most exemplars and the prototype. On the other hand, when subjects are presented with exemplars of a novel category that they have not been pre-exposed to, these mechanisms do not apply (at least not straight away), so there will not be any benefit in recognizing exemplars of that novel category in their upright orientation. Thus, no significant inversion effect would be expected, because an inverted novel checkerboard is just another novel checkerboard.

The explanation just given of the inversion effect found with exemplars drawn from a familiar category in terms of perceptual learning works well when considering the advantage enjoyed by upright exemplars drawn from that category. And we now have further evidence that this advantage is real. But this leaves us with evidence across five studies that familiarity with a prototype-defined category will lead to inverted members of that category being less easily discriminated (McLaren, 1997) or recognized (Experiments 1a, 2, 3a and 4 of this paper) than novel controls. The implications of this finding are far-reaching, because they suggest that the standard face inversion effect could also be due to a combination of two components, an advantage for upright faces (relative to other classes of stimuli), and a disadvantage for inverted faces. In fact, we note that it is only experiments of the type reported here which can establish this possibility, as the baseline for standard face inversion experiments is otherwise hard to determine. But how are we to explain this disadvantage? The perceptual learning analysis offered so far simply suggests a return to baseline performance, not something worse. What is it about familiarity with a prototype-defined category that leads to poorer discrimination or recognition of inverted exemplars drawn from that category?

McLaren (1997) speculated that this effect might be connected with the finding that participants were able to categorize exemplars even when they were inverted. Tests administered at the end of the experiments in that paper revealed that for both prototype-defined categories and shuffled categories, participants were able to classify inverted exemplars as members of the correct category with above chance accuracy (59% in both cases, compared to 66% and 70% for upright exemplars). We can see two possible mechanisms that might follow from this and explain the disadvantage for inverted exemplars drawn from a familiar category. The first is that, if participants are able to classify inverted exemplars as an "A" or "B", then this in its own can have consequences. If discrimination

between an exemplar from one category and an exemplar from the other is required then a "learned distinctiveness" effect (Honey and Hall, 1989) can be expected, whereby the different labels attached to each exemplar aid in their discrimination. But when the discrimination is within category, a "learned equivalence" effect can be expected instead (*ibid*), which enhances generalization between the stimuli making discrimination more difficult. Admittedly this effect can be expected for both upright and inverted exemplars drawn from a familiar category, but the upright exemplars benefit from perceptual learning as already outlined, which more than compensates for this effect. When this compensatory effect disappears on inversion, the cost of "equivalence" manifests and this is why the familiar inverted exemplars are poorly recognized compared to novel exemplars. This account is plausible, and there is evidence for the mechanisms involved, but it does suffer from the observation that no such effect was noticeable in Experiments 1b and 3b (or in Experiment 1b in 1997), and the effect would not be expected to be dependent on the category being prototype-defined.

The second possible mechanism does depend on category structure. In the case of prototype-defined categories, the ability to categorize inverted exemplars implies that features present in the these exemplars are capable of calling to mind some representation of the structure of that category, which will correspond to the upright, prototypical structure experienced during training. According to the MKM theory, it is exactly this ability that allows the differential salience of the unique elements of an exemplar to manifest and support better learning and memory. But, in the case of the inverted exemplars, the predictions made by retrieval of the prototypical structure will often be incorrect, and will not correspond to the layout of the black and white squares. Thus the elements that become differentially salient will be randomly determined, and will more often be those that are common to most exemplars (simply because there are more of them), rather than unique to any one of them.

This will have the effect of adding unwanted noise to the discrimination, making it more difficult – hence a disadvantage for inverted exemplars drawn from a familiar category will emerge. Note that this effect will be beyond that expected for novel stimuli, as in that case the elements will be uniformly unpredicted rather than randomly (and often incorrectly) predicted.

The ERP results from Experiment 4 also allow us to say a little more about the effect of familiarity with a prototype-defined category on inverted exemplars drawn from that category. It would seem to strongly affect the N170 for those stimuli, delaying it and increasing its amplitude. We speculate that this may be a direct neural correlate of the recognition / discrimination disadvantage suffered by these stimuli, but this is an issue that will have to await further research for confirmation. What we can say is that the effect on the N170 clearly correlates with the inversion effect found with familiar checkerboards, and that this replicates and extends the earlier demonstrations of such a correlation with Greebles (Rossion et al., 2002). It also fits in rather well with the effect on the N170 of disrupting the configural information in a face recently reported by ourselves (Civile, Elchlepp, McLaren, Lavric and McLaren, 2012), which is to attenuate the effect of inversion on the N170 by bringing the ERPs for the disrupted faces nearer to the upright normal face. We showed that by presenting different categories of novel faces in both upright and inverted orientations the FIE was reduced, and the elicited N170 was significantly smaller compared to normal inverted faces. Roxane Latimus, Taylor (2006), found a larger N170 for inverted faces compared to other objects (e.g. chairs, houses, cars) and animals (apes) in both upright and inverted orientations, all of which elicited a significantly smaller N170 (in some cases even smaller than that for upright human faces). More evidence in support of this finding comes from studies of inversion as modulated by ethnicity on the N170. Vizioli, Foreman, Rousselet, Caladara (2010) showed that the N170 amplitude for same race inverted faces was

significantly larger compared with upright same race and upright and inverted other race stimuli. Thus, the presentation of unfamiliar faces, in this case faces taken from an unfamiliar ethnic group, attenuates the effect of inversion on the N170. Additionally, the behavioral test showed that accuracy for inverted familiar race faces was significantly lower than for the other stimuli. These results, taken together with ours (Experiment 4) seem to suggest that the largest amplitude for the N170 is correlated with the lowest behavioral performance, in this case for familiar inverted exemplars.

We will conclude by considering some of the implications of our research for theories of perceptual learning in general. The advantage of the approach we have taken in extending the account given by McLaren (1997) to our current data, is that it does seem to have the potential to explain our results and to explain the role of perceptual learning in the face inversion effect. Can other theories of perceptual learning provide different explanations of the phenomena reported in this paper? We particularly have in mind here recent research by Mundy, Dwyer and Honey (2006), Mundy, Honey and Dwyer (2007), Dwyer and Vladeanu (2009) and Mundy, Honey and Dwyer (2009) that makes a case for a comparison process in perceptual learning in humans. This research with human participants (often using faces as stimuli) shows that simultaneous or alternated presentation of similar stimuli leads to better discrimination in a subsequent test phase. The inference is that the ability to compare the stimuli that have to be discriminated later leads to stronger perceptual learning compared to controls that are exposed to these stimuli equally often, but without the opportunity for comparison (often referred to as a "blocked" schedule of exposure).

Our response to these studies is to first note that McLaren, Forrest and McLaren (2012) have shown that the most recent version of MKM, in the form of the MKM-APECS hybrid model, can simulate the blocked vs. alternated effect. Thus, the evidence for a comparison process based on this type of result is not compelling. But if we take the

comparison process account as given, what are the implications for our present results? During categorization, our participants have the opportunity to compare exemplars both across and within categories. This would lead to them being better able to discriminate both within and between categories as a result of this opportunity for comparison (assuming it somehow generalizes to new exemplars) and could then predict an advantage for upright exemplars drawn from a familiar category in later recognition. In this respect the comparison account's predictions do not differ greatly from those already in plays, and they would doubtless go on to predict that inversion would lead to a loss of perceptual learning. We cannot, however, find any particular reason for them to predict that inversion of exemplars drawn from a familiar category would lead to worse performance than to novel exemplars (beyond some general account in terms of learned equivalence), and this seems to us something of a challenge for this class of theory. Given that this result is now established, it would be greatly to the credit of any theory of perceptual learning to offer at least some explanation for the effect. At present, to our knowledge, only the MKM based theories seem capable of doing this. A further difficulty for the comparison account would then be to explain why experience with the shuffled stimuli did not lead to perceptual learning? Surely, if people can learn to categorize these stimuli, then they should benefit from the opportunity to compare them in the same way that those exposed to prototype-defined exemplars would? We will leave it for the researchers actively involved in developing these theories to respond to these challenges, but at present do not feel able to give an account of our results in these terms.

Our final issue concerns a number of recent studies that have shown that perceptual learning can, under some conditions, simply involve participants learning where to look, rather than in any way implying some enhancement of stimulus discriminability (Jones and Dwyer, 2013; Wang, Lavis, Hall and Mitchell, 2012). Could this explanation apply to the

experiments reported in this paper, so that the inversion effect is due to participants learning where to look during categorization training, and applying this strategy during the recognition experiment, with some success in the case of the upright familiar exemplars, but suffering because of it when dealing with inverted familiar exemplars? Whilst we would readily acknowledge that strategies of this kind are possible in many perceptual learning experiments, we do not believe that they are in play here. The stimuli are all randomly generated in our experiments, and there is no particular region to focus on to detect individuating features for any stimulus set. The squares changed vary from stimulus to stimulus, making any such strategy unlikely to succeed. It might be that categorization training encourages participants to look at certain regions of a stimulus to distinguish members of category A from those of category B, but this is unlikely to have any relevance to discrimination within one of these categories, which is what is tested in the recognition phase. We ensured that only one of the familiar categories was ever used in a given study / recognition phase cycle of our experiments, so that any enhancement of the ability to distinguish between categories A and B as a result of experience with them, would not in itself improve recognition performance. Finally, given that the shuffled stimuli were more easily categorized, and if this is to be taken as an index of success in learning the necessary strategy, surely the inversion effect should have been larger in Experiment 1b rather than non-existent? We conclude that we have no evidence in our data that the participants' enhanced performance on exemplars drawn from a familiar category is due to learning where to look, or that their impaired performance on inverted exemplars from the same category is due to this type of learning. Rather, it would seem that an explanation in terms of enhanced stimulus discriminability is to be preferred.

In conclusion: In four experiments we have demonstrated that we can obtain a strong inversion effect in a recognition task contingent on use of exemplars drawn from a familiar

prototype-defined category. This effect can be decomposed into an advantage for upright exemplars drawn from a familiar category, and a disadvantage for inverted exemplars drawn from a familiar category, and has a neural correlate in the N170, which seems to predominantly reflect the disadvantage for inverted exemplars drawn from a familiar category. An explanation based on enhanced stimulus discriminability as a consequence of the differential latent inhibition of common elements is our current best attempt at explaining these phenomena, and our next step will be to implement further tests of this theoretical account.

Acknowledgments

The research reported in this paper was supported by: a Postgraduate studentship and an Exeter Graduate Fellowship awarded to Ciro Civile; a Laboratory study visit scholarship awarded to Ciro Civile by the International Office, University of Exeter; and this work was also supported by the National Key Fundamental Research (973) Program (2013CB329501) of China to Yixuan Ku.

References

Bartlett, J. C., & Searcy, J. (1993). Inversion and configuration of faces. *Cognitive Psychology*, *25*, 281–316.

Bartlett, J. C., Searcy, J., & Abdi, H. (2003). What are the routes to face recognition? In M. A. Peterson & G. Rhodes (Eds.), *Perception of faces, objects, and scenes: Analytic and holistic processing* (pp. 21-52). Oxford: Oxford University Press.

Bell, A. J, & Sejnowski, T. J. (1995). An information–maximization approach to blind separation and blind deconvolution. *Neural Computation*, *7*, 1129-59.

Bentin, S., Allison, T., Puce, A., Perez, E., McCarthy, G. (1996). Electrophysiological studies of face perception in humans, *Journal of Cognitive Neurosc*ience. 551-565.

Brown, G. D., Yamada, S., & Sejnowski, T. J. (2001). Independent component analysis at the neural cocktail party. *Trends in Neurosciences*, *24*, 54-63.

Busey, T.A. & Vanderkolk, J.R. (2005). Behavioral and electrophysiological evidence for configural processing in fingerprint experts. *Vision Research 45*, 431-448.

Civile, C., Elchlepp, H., McLaren, R., Lavric, A & McLaren, I.P.L. (2012). The face inversion effect and evoked brain potentials: Complete loss of configural information affects the N170. In N. Miyake, D. Peebles, & R. P. Cooper (Eds.), *Proceedings of the 34th Annual Conference of the Cognitive Science Society* (pp. 1416-1421). Austin, TX: Cognitive Science Society.

Debener, S., Ullsperger, M., Siegel, M., Fieler, K.,. Von Cramon, D. Y., & Engel, A. (2005). Trial-by-trial coupling of concurrent electroencephalogram and functional magnetic resonance imaging identifies the dynamics of performance monitoring. *Journal of Neuroscience*, *25*, 730-737.

Diamond, R. & Carey, S.(1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology*: *General*, *115*, 107-117.

Dienes, Z.(2011). Bayesian Versus Orthodox Statistics: Which Side Are You On? *Perspectives on Psychological Science , 6*, 274–290.

Dwyer, D.M. and Vladeanu, M.C. (2009). Perceptual learning in face processing: comparison facilitates face recognition. *The Quarterly Journal of Experimental Psychology*, *62*, 2055-2067.

Eimer, M. (2000). The face-specific N170 component reflects late stages in the structural encoding of faces. *NeuroReport, 11*, 2319-2324.

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

Gauthier I., Skudlarski P., Gore J.C. & Anderson A.W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience, 3,* 191–197.

Gauthier I. & Tarr M.G. (1997). Becoming a "Greeble" expert: exploring mechanisms for face recognition. *Vision Research, 12,* 1673-1682.

Graham, S. and McLaren, I.P.L. (1998). Retardation in human discrimination learning as a consequence of preexposure: Latent inhibition or negative priming? *The Quarterly Journal of Experimental Psychology*, *51B*, 155-172.

Honey, R.C. and Hall, G. (1989). Acquired equivalence and distinctiveness of cues. *Journal of Experimental Psychology: Animal Behavior Processes*, *15*, 338-346.

Jones, S.P. and Dwyer, D.M. (2013). Perceptual learning with complex visual stimuli is based on location, rather than content, of discriminating features. *Journal of Experimental Psychology: Animal Behavior Processes*, Online First Publication.

Lavric, A., Bregadze, N., &Benattayallah, A. (2010). Detection of experimental ERP effects in combined EEG–fMRI: Evaluating the benefits of interleaved acquisition and

Independent Component Analysis. *Clinical Neurophysiology*, doi:10.1016/j.clinph.2010.06.033.

Leder, H., & Bruce, V. (1998). Local and relational aspects of face distinctiveness. *The Quarterly Journal of Experimental Psychology* , *51A*, 449 – 473.

Leder, H., & Bruce, V. (2000). When inverted faces are recognised: the role of configural information in face recognition. *The Quarterly Journal of Experimental Psychology*, *53*, 513-536.

Leder, H., Huber, O., and Bruce, V. (2001) Configural features in the context of upright and inverted faces. *Perception, 30*, 73–83.

McClelland, J.L. & Rumelhart, D.E. (1985). Distributed memory and the representation of general and specific information. *Journal of Experimental Psychology: General*, *114*, 159-197.

McLaren, I.P.L. (1997). Categorization and perceptual learning: An analogue of the face inversion effect. *The Quarterly Journal of Experimental Psychology 50A*, 257-273.

McLaren, I.P.L., Forrest, C.L., McLaren, R.P. (2012). Elemental representation and configural mappings: combining elemental and configural theories of associative learning. *Learning and Behavior*, *40*, 320-333.

McLaren, I.P.L., Kaye, H. & Mackintosh, N.J. (1989). An associative theory of the representation of stimuli: Applications to perceptual learning and latent inhibition. In R.G.M. Morris (Ed.) *Parallel Distributed Processing - Implications for Psychology and Neurobiology*. Oxford, Oxford University Press.

McLaren, I.P.L., Leevers, H.L. & Mackintosh, N.J. (1994). Recognition, categorisation and perceptual learning. In C. Umilta & M. Moscovitch (Eds.) *Attention & Performance XV*. Cambridge, MA: MIT Press.

McLaren, I.P.L. and Mackintosh, N.J. (2000). An elemental model of associative learning: Latent inhibition and perceptual learning. *Animal Learning and Behavior*, *38*, 211-246.

McLaren I.P.L. and Mackintosh NJ. (2002). Associative learning and elemental representation: II. Generalization and discrimination. *Animal Learning and Behavior, 30*, 177–200.

Mundy, M. E., Dwyer, D. M. and Honey, R. C. (2006). Inhibitory associations contribute to perceptual learning in humans. *Journal of Experimental Psychology: Animal Behaviour Processes*, *32*, 178-184.

Mundy, M. E., Honey, R. C. and Dwyer, D. M. (2007). Simultaneous Presentation of Similar Stimuli Produces Perceptual Learning in Human Picture Processing. Journal of Experimental Psychology: Animal Behaviour Processes, *33*, 124-138.

Mundy, M. E., Honey, R.C., and Dwyer, D. M. (2009). Superior discrimination between similar stimuli after simultaneous exposure. *The Quarterly Journal of Experimental Psychology*, *62*, 18-25.

Robbins, R. & McKone, E. (2007). No face-like processing for objects-of-expertise in three behavioural tasks. *Cognition*, 103, 34-79.

Rossion, B., Delvenne, J., Debatisse, D., Goffaux, V., Bruyer, R., Crommelinck, M., Guerit, M. (1999). Spatio-Temporal localization of the face inversion effect: an event-related potentials study. *Biological Psychology, 50*, 173-189.

Rossion, B., Gauthier, I., Goffaux, V., Tarr, M.-J., Crommelinck, M. (2002). Expertise training with novel objects leads to face-like electrophysiological responses. *Psychological Science*, *13*, 250-257.

Rossion, B., Gauthier, I., Tarr, M.J., Despland, P., Bruyer, R., Linotte, S., & Crommelinck, M. (2000). The N170 occipito-temporal component is delayed and enhanced to

inverted faces but not to inverted objects: An electrophysiological account of face-specific processes the human brain. *NeuroReport, 11*, 69–74.

Roxane J., Latimus, M., Taylor, M.(2006). Face, eye and object early processing: What is the face specificity? *NeuroImage*, 29, 667 – 676.

Tanaka, J. W., & Farah, M. J.(1991). Second-order relational properties and the inversion effect: Testing a theory of face perception. *Perception & Psychophysics*, *50*, 367–372.

Tanaka, J.W., & Gauthier, I.(1997). Expertise in object and face recognition. In R.L. Goldstone, P.G. Schyns, & D.L. Medin (Eds.), *Psychology of learning and motivation: Vol. 36. Perceptual mechanisms of learning* (pp. 83–125). San Diego, CA: Academic Press.

Tanaka, J.W., & Sengo, J.A.(1997). Features and their configuration in face recognition. *Memory & Cognition*, 25, 583–592.

Tanaka, J.W. and Curran, T. (2001) A neural basis for expert object recognition. *Psychological Science*, 12, 43–47.

Tarr, M. J., & Gauthier, I. (2000). FFA: A flexible fusiform area for subordinate-level visual processing automatized by expertise. *Nature Neuroscience*, *3*, 764-769.

Taylor, M.J., McCarthy, G., Saliba, E., & Degiovanni, E. (1999). ERP evidence of developmental changes in processing of faces. *Clinical Neurophysiology*, *100*, 910–915.

Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology*, *81*, 141-145.

Valentine, T., & Bruce, V. (1988). Mental rotation of faces. *Memory & Cognition, 16*, 556-566.

Vizioli, L., Foreman, L., Rousselet, A., Caladara, R.(2010). Inverting faces elicits sensitivity to race on the N170 component: A cross-cultural study. *Journal of Vision 10*, 1–23.

Wang, T., Lavis, Y., Hall, G. and Mitchell, C.J. (2012). Location and salience of unique features in human perceptual learning. *Journal of Experimental Psychology: Animal Behavior Processes*, *38*, 407-418.

Wills, A.J., Suret, M.B. and McLaren, I.P.L. (2004). The role of category structure in determining the effects of stimulus pre-exposure on categorisation accuracy. *The Quarterly Journal of Experimental Psychology*, *57B*, 79-88.

	Study Phase				*Old/New Recognition task*			
Stimulus type	**Familiar**		**Novel**		**Familiar**		**Novel**	
	Upright	Inverted	Upright	Inverted	Upright	Inverted	Upright	Inverted
Latencies (ms)								
PO7 (Left)	168.562	172.375	171.937	170.875	159.187	161.062	162.312	161
PO8 (Right)	166.437	172.062	168.562	168.843	160.125	160.562	160.062	161.875
Amplitudes[μV]								
PO7 (Left)	-0.072	-0.558	-0.025	-0.093	0.007	-0.002	-0.165	0.122
PO8 (Right)	0.005	-0.557	0.002	-0.013	0.064	0.086	0.197	-0.154

Table 1: Latency and Amplitude data for the N170 in Experiment 4.

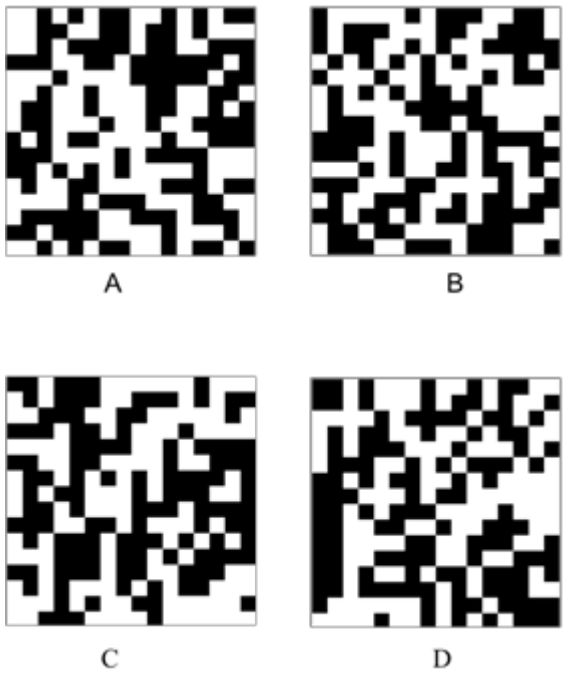

Figure 1: The prototypes for categories A, B, C and D. 48 of the squares were randomly changed to generate each exemplar, thus, on average, 24 squares would be expected to alter from black to white or white to black.

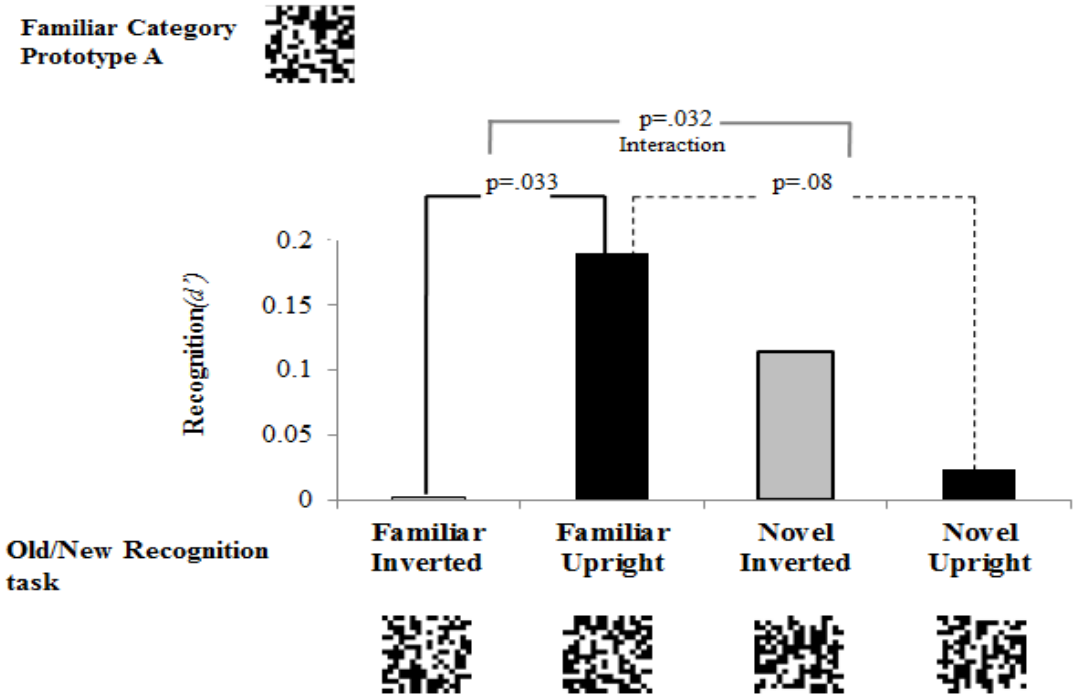

Figure 2: The x-axis represents the four different stimulus conditions, and the y-axis gives the mean d' for the old/new recognition task of Experiment 1a.

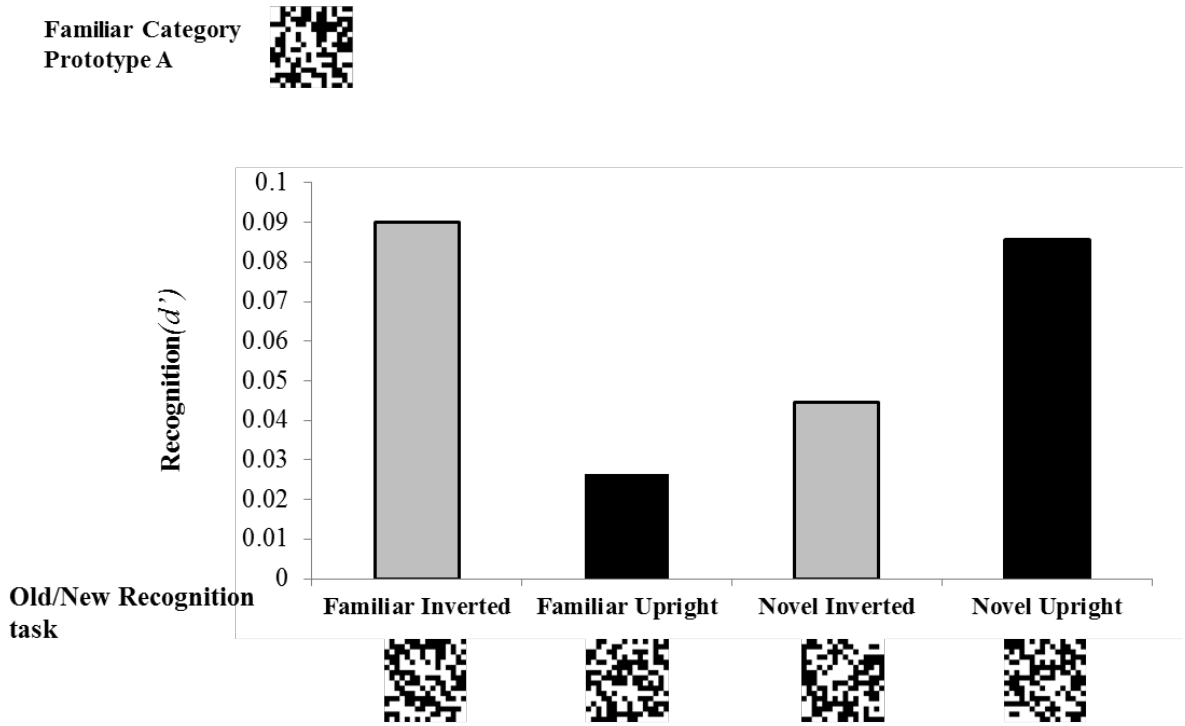

Figure 3: The x-axis shows the four different stimulus conditions, and the y-axis shows the d' scores for the test phase of Experiment 1b.

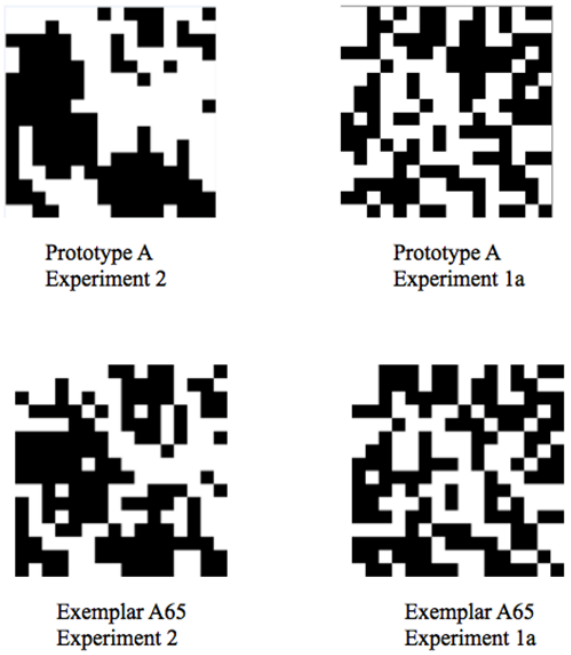

Figure 4: The prototypes and some example exemplars for category A in Experiment 1a and for category A in Experiment 2.

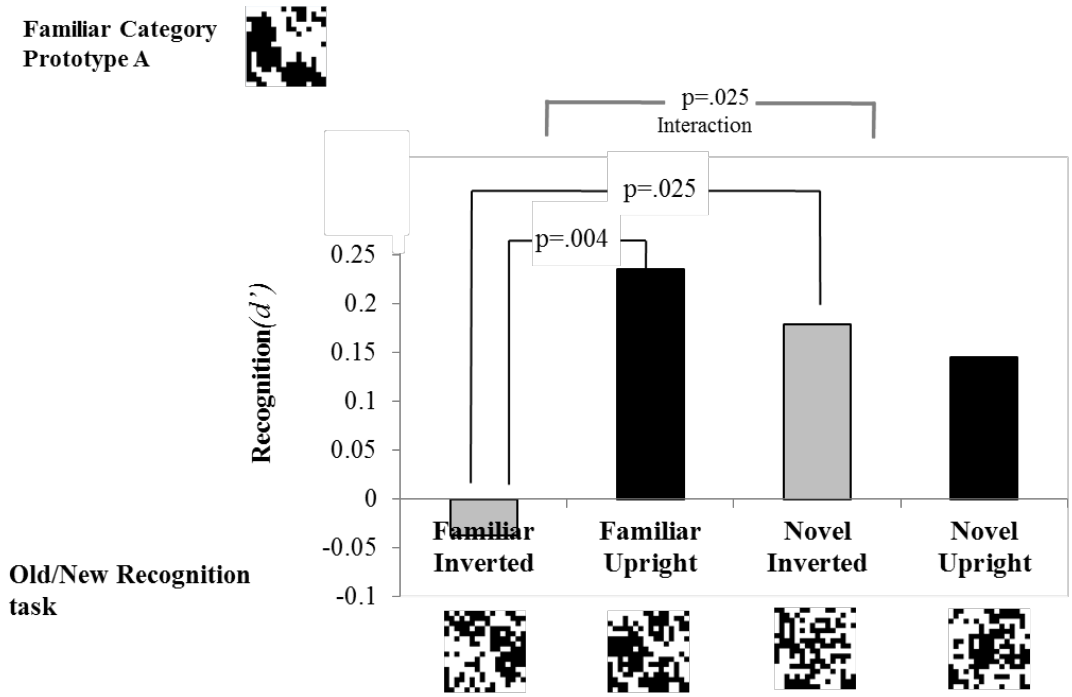

Figure 5: The x-axis gives the four different stimulus' conditions, and the y-axis shows the mean d-prime scores for the old/new recognition phase in Experiment 2.

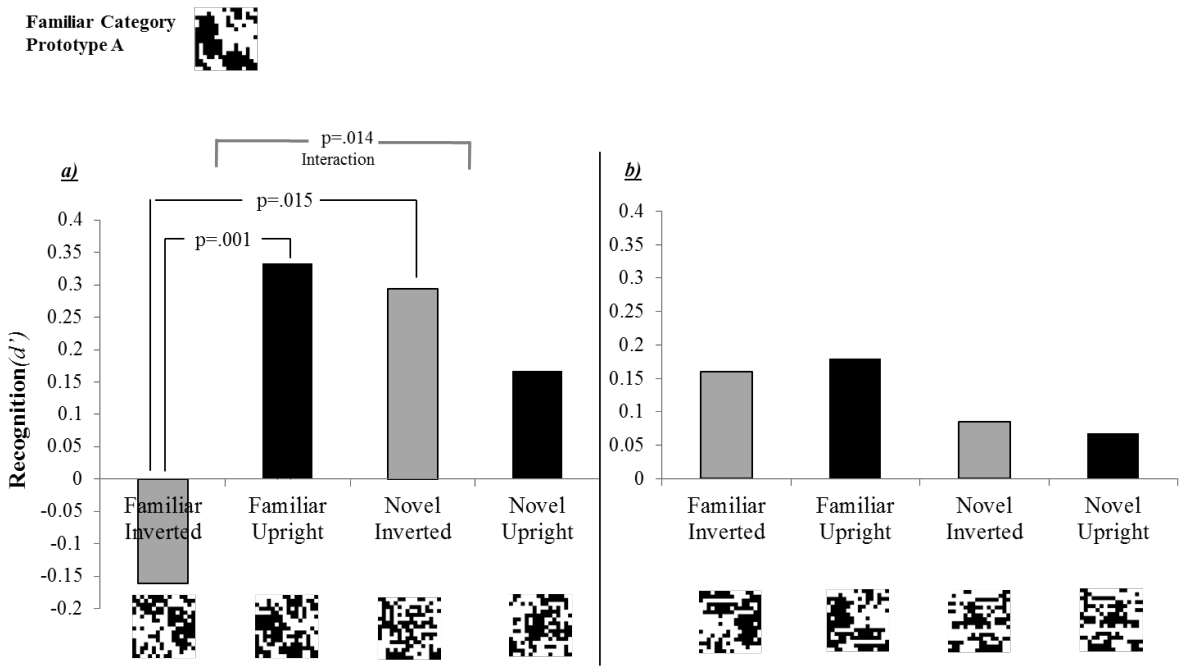

Figure 6: Panel *a)* represents the results obtained in Experiment 3*a*. Panel *b)* represents the results obtained in Experiment 3*b*. Thus in both panels the x-axis gives the four different stimulus' conditions, and the y-axis shows the mean d-prime scores for the old/new recognition phase in the experiment.

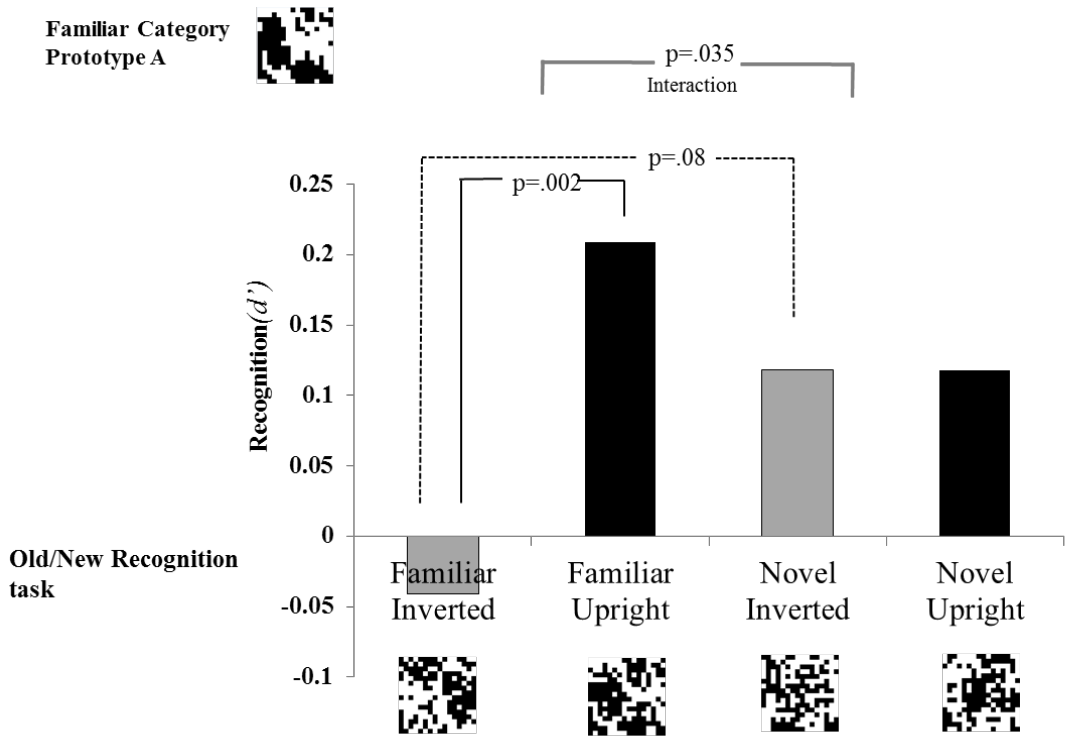

Figure 7: The x-axis gives the four different stimulus' conditions, and the y-axis shows the mean d-prime scores for the old/new recognition phase in Experiment 4.

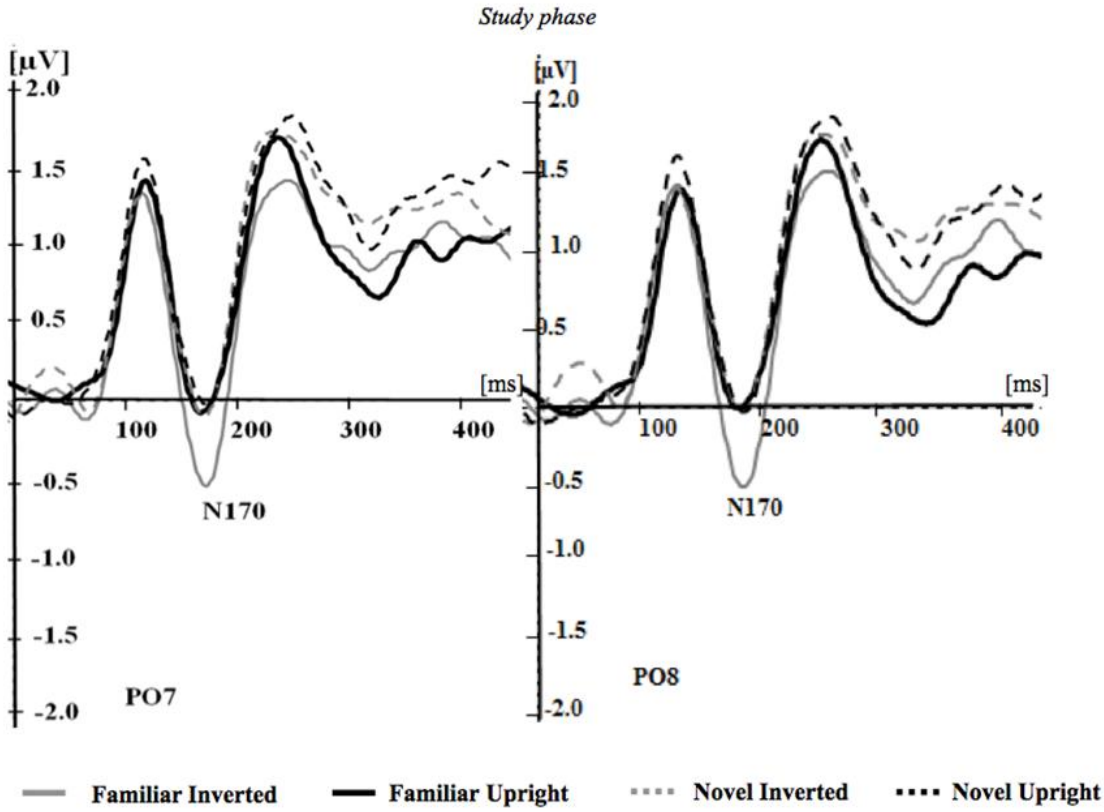

Figure 8: Waveforms obtained at electrode PO7 and PO8 during the study phase of Experiment 4.

Table

	Study Phase				*Old/New Recognition task*			
Stimulus type	**Familiar**		**Novel**		**Familiar**		**Novel**	
	Upright	Inverted	Upright	Inverted	Upright	Inverted	Upright	Inverted
Latencies (ms)								
PO7 (Left)	168.562	172.375	171.937	170.875	159.187	161.062	162.312	161
PO8 (Right)	166.437	172.062	168.562	168.843	160.125	160.562	160.062	161.875
Amplitudes[μV]								
PO7 (Left)	-0.072	-0.558	-0.025	-0.093	0.007	-0.002	-0.165	0.122
PO8 (Right)	0.005	-0.557	0.002	-0.013	0.064	0.086	0.197	-0.154

Table 1: Latency and Amplitude data for the N170 in Experiment 4.

Figure

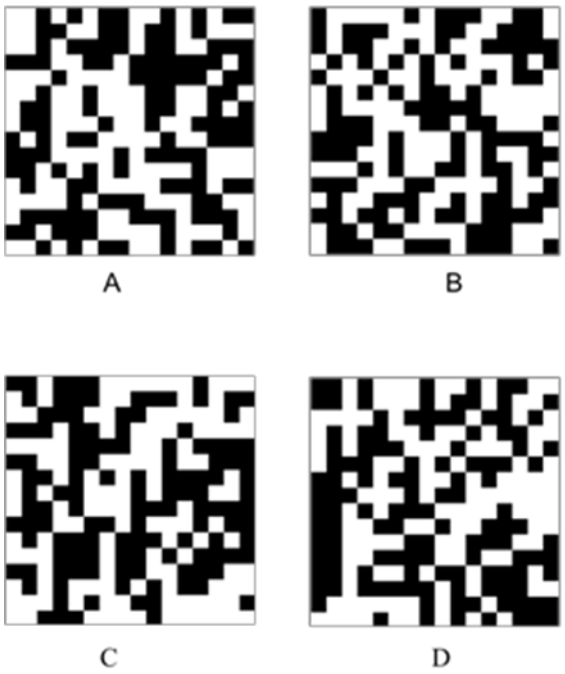

Figure 1: The prototypes for categories A, B, C and D. 48 of the squares were randomly changed to generate each exemplar, thus, on average, 24 squares would be expected to alter from black to white or white to black.

Figure

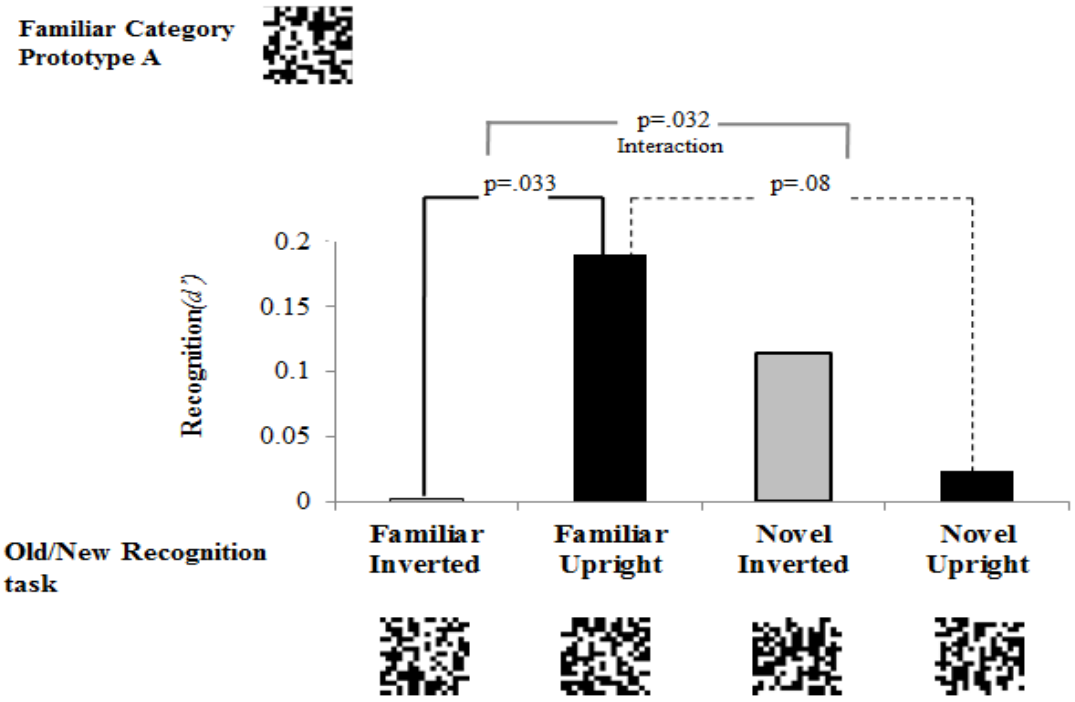

Figure 2: The x-axis represents the four different stimulus conditions, and the y-axis gives the mean d' for the old/new recognition task of Experiment 1a.

Figure

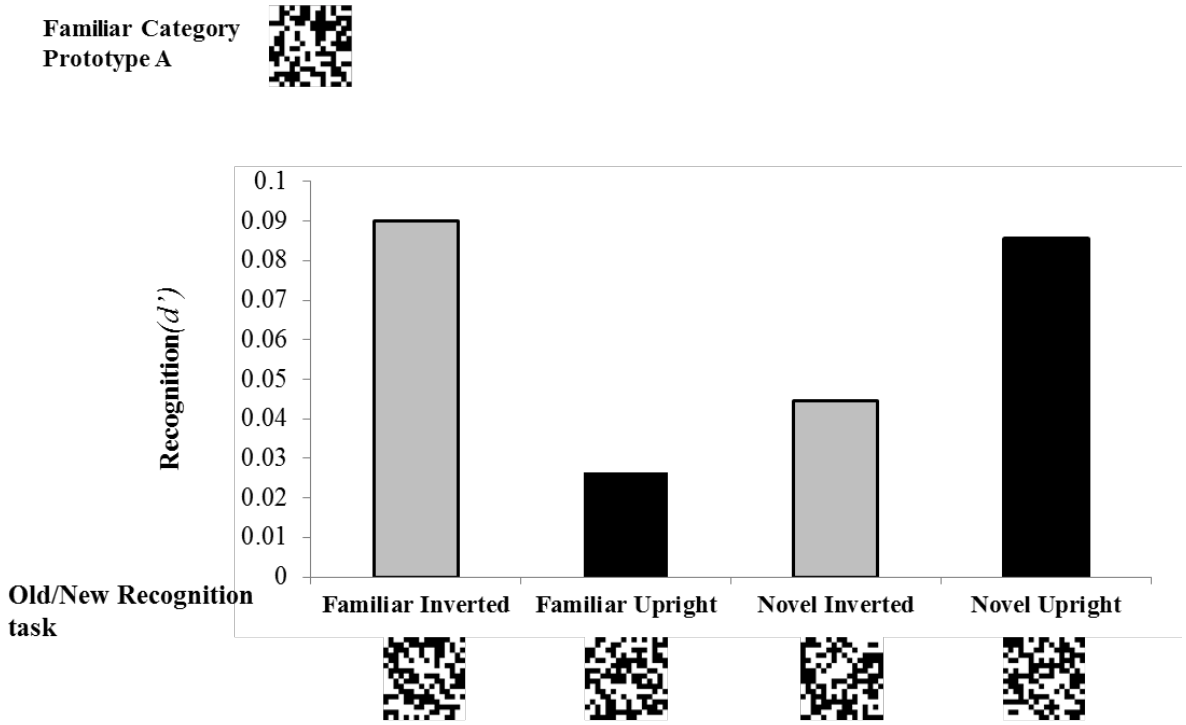

Figure 3: The x-axis shows the four different stimulus conditions, and the y-axis shows the d' scores for the test phase of Experiment 1b.

Figure

Figure 4: The prototypes and some example exemplars for category A in Experiment 1a and for category A in Experiment 2.

Figure

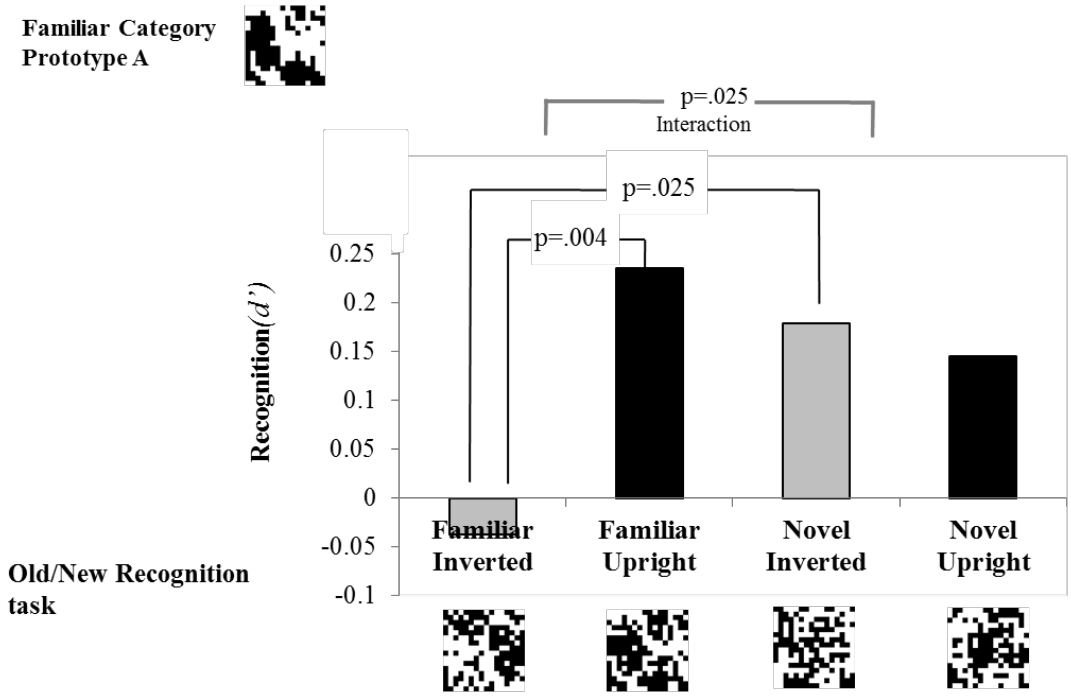

Figure 5: The x-axis gives the four different stimulus' conditions, and the y-axis shows the mean d-prime scores for the old/new recognition phase in Experiment 2.

Figure

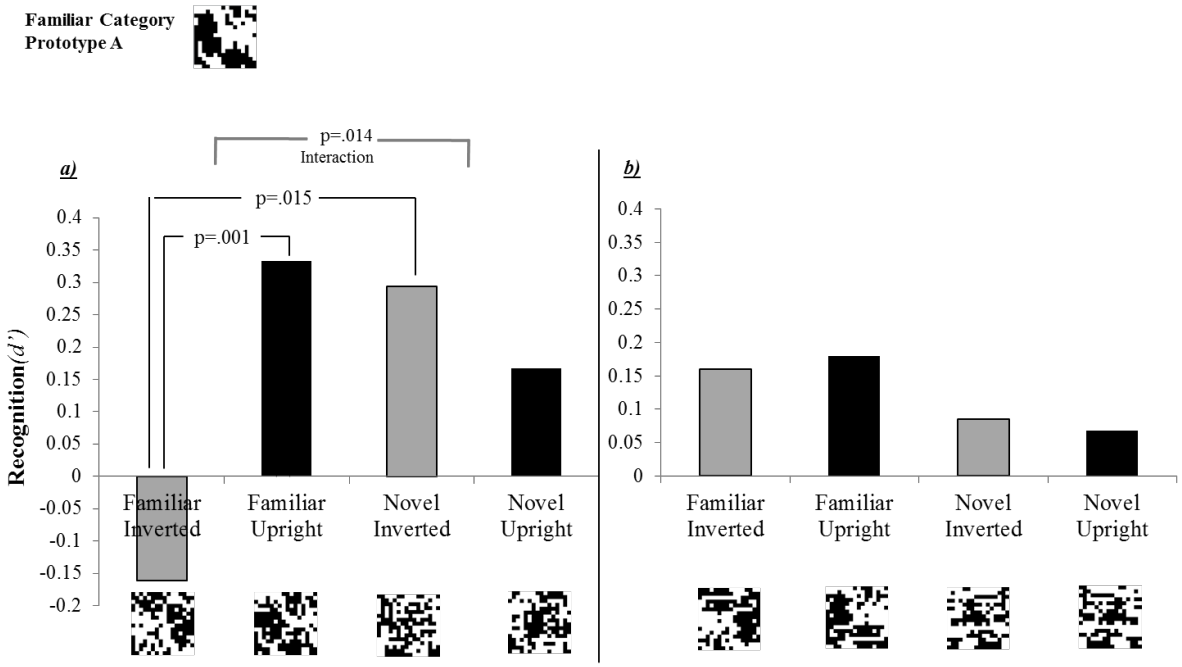

Figure 6: Panel *a)* represents the results obtained in Experiment 3*a*. Panel *b)* represents the results obtained in Experiment 3*b*. Thus in both panels the x-axis gives the four different stimulus' conditions, and the y-axis shows the mean d-prime scores for the old/new recognition phase in the experiment.

Figure

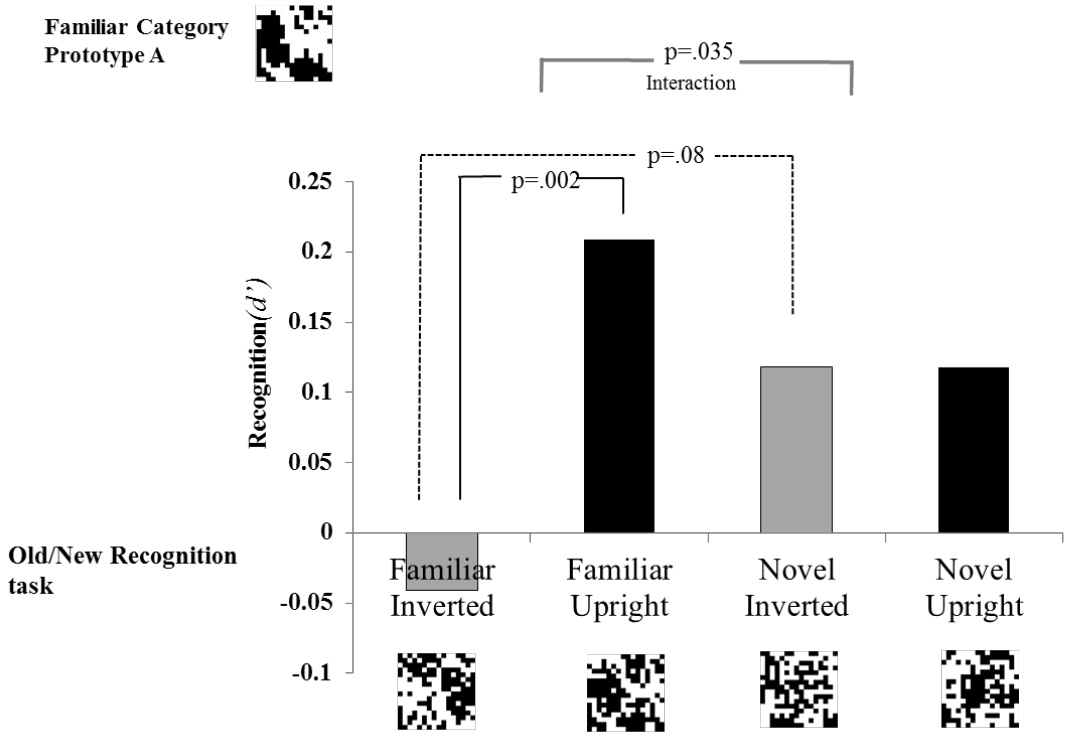

Figure 7: The x-axis gives the four different stimulus' conditions, and the y-axis shows the mean d-prime scores for the old/new recognition phase in Experiment 4.

Figure

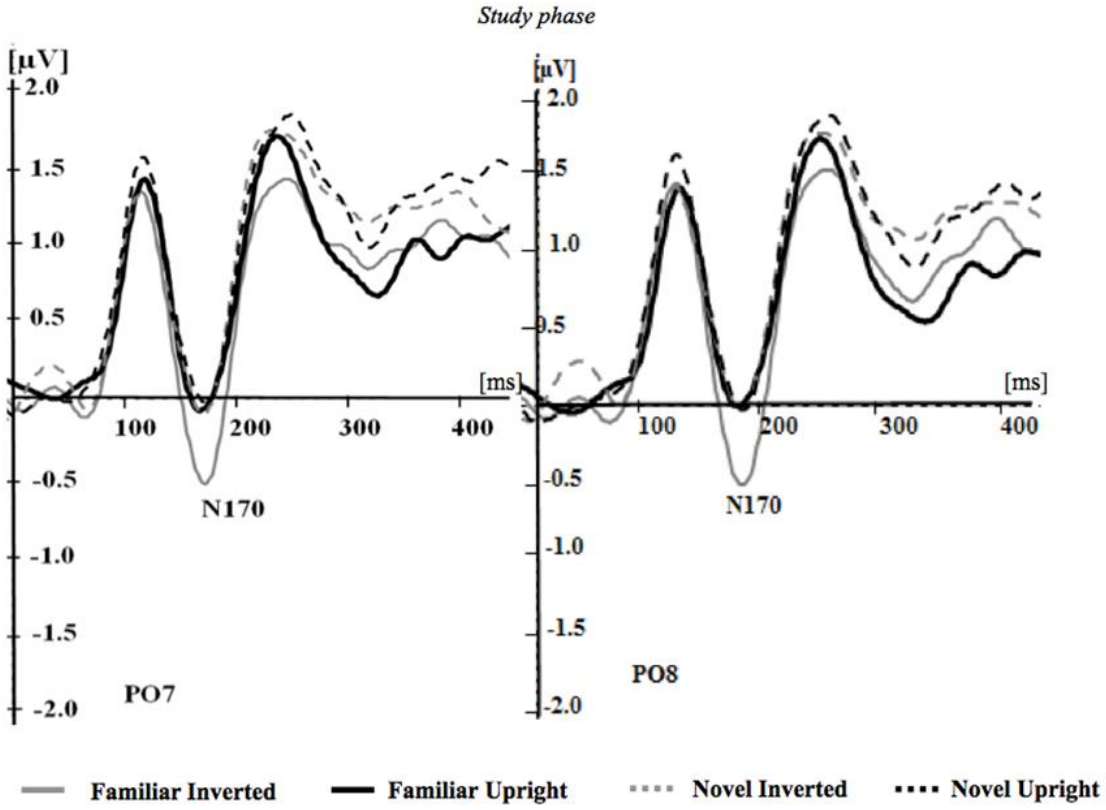

Figure 8: Waveforms obtained at electrode PO7 and PO8 during the study phase of Experiment 4.