Complexity Gates the Emergence of Visual Expertise

Yizhen Shen (ys619@exeter.ac.uk)
Department of Psychology, Faculty of Health & Life Sciences,
University of Exeter, UK

Ciro Civile (c.civile@exeter.ac.uk)
Department of Psychology, Faculty of Health & Life Sciences,
University of Exeter, UK

Abstract

The face inversion effect (FIE)—the disproportionate impairment in recognising inverted faces—has been central to debates between domain-specific and perceptual expertise accounts of visual processing. Although expertise-based explanations predict that inversion effects should emerge for non-face stimuli following learning, most supporting evidence derives from memory-based recognition tasks, leaving open the possibility that such effects reflect retrieval strategies rather than perceptual encoding. The present study tested this claim using a delayed matching-to-sample task with prototype-defined checkerboards, isolating perceptual encoding from long-term memory retrieval while manipulating stimulus complexity. Participants completed a categorisation training phase followed by a matching task involving familiar and novel stimuli presented upright or inverted. Results revealed a robust inversion effect for familiar stimuli only in the high-complexity condition, whereas no such effect emerged for simpler stimuli. These findings support the McLaren–Kaye–Mackintosh theory of perceptual learning by demonstrating that expertise modifies online perceptual encoding and that the engagement of holistic-like processing depends on stimulus complexity. We conclude that inversion effects reflect general expertise-based learning mechanisms operating under conditions of high informational load rather than face-specific perceptual architecture.

Keywords: Perceptual expertise; Face inversion effect; Stimulus complexity; Perceptual learning; Visual categorisation; Holistic processing

Introduction

The human visual system can rapidly discriminate between highly similar exemplars within familiar categories, reflecting a remarkable level of cognitive efficiency. This capacity is most striking in face recognition, where observers can identify a specific individual among thousands in a fraction of a second. However, this ability is disproportionately disrupted by inversion: identifying inverted faces is far more difficult than identifying inverted objects, a phenomenon known as the Face Inversion Effect (Yin, 1969). For decades, the FIE has served as a focal point for two competing theoretical frameworks concerning the organisation of the visual cortex. The domain-specific hypothesis argues that the FIE reflects the operation of a specialised neural module, potentially localised in the fusiform face area, that evolved to process the unique geometry of faces (Kanwisher, 2000; Yovel & Kanwisher, 2004). In contrast, the perceptual expertise hypothesis proposes that the FIE arises not from stimulus class per se, but from extensive experience discriminating among mono-oriented stimuli that share a common configuration (Diamond & Carey, 1986; Gauthier & Tarr, 1997).

To adjudicate between these accounts while avoiding the confounds of social significance and lifelong exposure associated with faces, researchers have increasingly employed artificial stimuli. McLaren (1997) and colleagues established prototype-defined checkerboards as a rigorous analogue for studying perceptual expertise and learning. Unlike Greebles (Gauthier & Tarr, 1997), which share structural similarities with faces, checkerboards are abstract, non-biological, and initially novel to all participants. It is now well established that following categorisation training, participants exhibit a robust inversion effect for familiar checkerboards that is behaviourally and electrophysiologically comparable to the FIE (Civile et al., 2014; McLaren & Civile, 2011). These findings suggest that the "special" processing afforded to faces may reflect a general learning mechanism tuned to stimulus structure rather than face-specific architecture.

The computational basis for this mechanism is articulated by the McLaren–Kaye–Mackintosh (MKM) theory of perceptual learning (McLaren & Mackintosh, 2000; McLaren et al., 1989). MKM theory posits that stimulus representations evolve through error-driven associative learning rather than remaining static. With repeated exposure to a category defined by a central prototype, features common to that category (e.g., the general configuration of facial features or the base grid structure of checkerboards) become highly predicted by context. Through a salience modulation algorithm, well-predicted features undergo a reduction in salience via differential latent inhibition, whereas unique, diagnostic features remain highly salient. As a result, expert representations of upright stimuli suppress redundant configuration information, allowing distinctive

In G. Boleda, I. Dautriche, M. Maldonado, C. Saldana and T. Torrent (Eds.), *Proceedings of the 48th Annual Conference of the Cognitive Science Society*

features to "pop out." Inversion disrupts this process by altering the spatial context such that learned inhibitory weights no longer apply, flooding the system with high-salience prototype noise and impairing fine discrimination—behaviourally manifesting as the inversion effect.

Despite these advances, a critical theoretical ambiguity remains regarding whether these effects reflect changes in perceptual encoding or downstream memory processes. Most prior demonstrations of checkerboard inversion effects have relied on Old/New recognition paradigms (Civile et al., 2016; Civile et al., 2014), in which participants study items and later retrieve them from memory. This reliance on memory tasks leaves the expertise hypothesis vulnerable to the critique that the FIE reflects memory retrieval strategies rather than alterations to perceptual encoding. Indeed, Megreya and Burton (2006) showed that performance in simultaneous matching tasks—where memory demands are minimal—is surprisingly poor for unfamiliar faces and often lacks typical inversion effects. They suggested that holistic or configural processing requires stable memory representations to engage. However, MKM theory predicts that perceptual learning reshapes front-end representations, effectively altering the contents of the "visual sketchpad" at the moment of encoding (Baddeley, 2000; McLaren et al., 1989). If expertise is fundamentally perceptual, then inversion effects should emerge even in pure matching tasks, provided observers possess sufficient category familiarity.

Recent work has begun to address this issue by adapting matching paradigms to index perceptual learning (Civile et al., 2021). However, to fully evaluate the MKM account against domain-specific alternatives, it is also necessary to consider the boundary conditions under which expertise mechanisms operate, particularly with respect to stimulus complexity. Domain-specific accounts often imply that faces require holistic processing because their complexity exceeds the capacity of part-based analysis. MKM theory offers a precise prediction: the latent inhibition mechanism is an adaptive response to high informational load. For simple stimuli with few features, observers may succeed using serial, feature-based strategies, making expert filtering helpful but unnecessary. As stimulus complexity increases, however, the abundance of shared elements creates a signal-to-noise problem that overwhelms serial processing capacity. Under such conditions, suppressing redundant prototype features becomes computationally necessary for efficient discrimination.

The present study sought to test these theoretical predictions empirically. We employed a delayed matching-to-sample task using prototype-defined checkerboards to isolate perceptual encoding from long-term memory retrieval, while manipulating stimulus complexity by generating two families of stimuli: "Simple" (clumpy) and "Harder" (fragmented). Following a categorisation training phase, we tested two hypotheses derived from MKM theory. First, we predicted that categorisation training would induce a robust inversion effect in a matching task, thereby confirming that expertise modifies online perceptual encoding and countering the memory-artifact critique. Second, we hypothesised a main effect of complexity: although the MKM mechanism should operate automatically in simple stimuli, the magnitude of the inversion effect should scale with complexity. Such a pattern would indicate that holistic-like processing is not an intrinsic property of faces, but rather the outcome of a domain-general learning system optimising signal-to-noise ratios under conditions of high-dimensional visual input.

Methods

Participants

The final sample comprised 53 participants (N = 53) recruited from the University of Exeter. All participants were over 18 years of age and reported normal or corrected-to-normal vision. The study was designed as a pilot behavioural experiment to establish baseline parameters for a subsequent project involving transcranial direct current stimulation (tDCS). Participants were randomly assigned to one of two between-subjects conditions based on stimulus complexity: 29 participants were assigned to the Simple checkerboard condition (N = 29), and 24 to the Harder checkerboard condition (N = 24). No additional demographic variables were collected.

Stimuli

The stimuli were artificial, prototype-defined checkerboards generated using established methods (Civile et al., 2014). All stimuli consisted of 16 × 16 grids with equal proportions of black and white squares (50% luminance). To manipulate complexity, two distinct stimulus families were created: a Harder ("classic") condition and a Simple ("clumpy") condition.

In the Harder condition, four prototypes were generated as fully random binary matrices, each containing 50% black and 50% white squares and constrained to share exactly 50% of their squares with one another to ensure moderate inter-category similarity. Individual exemplars were created by randomly flipping 48 squares in each prototype (black ↔ white), with flips applied independently and uniformly across the grid. This procedure produced fine-grained, fragmented patterns that remained similar to their prototypes while requiring detailed discrimination, yielding visually complex stimuli.

In the Simple condition, four prototypes were generated using a neighbourhood-dependent algorithm in which the probability of a square being black or white depended on the colours of its neighbours, producing large contiguous regions ("clumps") of the same colour. These prototypes also contained 50% black and 50% white squares and shared 50% overlap with one another. Exemplars were generated by randomly flipping 96 squares in each prototype (double the

mutation rate of the Harder condition). Because the underlying prototypes were spatially coherent, these changes preserved large-scale structure while increasing within-category variability, resulting in stimuli that were easier to categorise and recognise.

Thus, although the Simple condition employed a higher mutation rate, perceptual difficulty was primarily determined by spatial organisation rather than physical distortion, allowing complexity to be manipulated independently of overall stimulus similarity.

Figure 1: Exemplars from harder and simpler categories.

Procedure

The experiment was implemented on the Gorilla online platform and conducted in a quiet testing environment using an iMac computer, with participants seated approximately 70 cm from the screen. The design adapted a same–different delayed matching-to-sample paradigm from previous studies (Civile et al., 2024; Civile et al., 2021), and consisted of three sequential phases: categorisation, practice, and matching.

The categorisation phase served as a pre-exposure training period to familiarise participants with two of the checkerboard categories (A/B or C/D). Participants completed 128 trials (64 per category). Each trial began with a 1-s fixation cross, followed by a single checkerboard stimulus. Participants classified each stimulus as Category A or B within a 4-s response window and received immediate feedback to facilitate error-driven learning.

Participants then completed a practice phase to learn the response keys for the matching task. This phase consisted of 48 trials (24 same, 24 different), each comprising a 1-s fixation cross followed by a 1-s stimulus presentation. Participants pressed 'x' or '.' to indicate their response, with key mappings counterbalanced across participants. Feedback was provided on every trial.

The final matching task followed a 2×2×2 mixed factorial design with Orientation (Upright vs. Inverted) and Familiarity (Familiar vs. Novel) as within-subjects factors and Complexity (Simple vs. Harder) as a between-subjects factor. Participants completed 128 trials with equal numbers of same and different pairs. Each trial began with a fixation cross (1-s), followed by a Target stimulus (1-s), a blank interstimulus interval (ISI) of 1.5 s, and then a Test stimulus (2-s). Participants judged whether the Test stimulus matched the Target using the response keys learned during practice.

Orientation was manipulated by rotating stimuli 180° on inverted trials; within any trial, both Target and Test stimuli appeared in the same orientation. Familiarity was manipulated by drawing matching pairs from either the familiar prototype categories learned during categorisation (e.g., A/B) or from novel categories not encountered during training (e.g., C/D). Prototype familiarity was counterbalanced across participants to ensure that any performance differences reflected category exposure rather than intrinsic stimulus properties. Accuracy and reaction times were recorded for all trials.

Results

Performance on the matching task was converted into perceptual sensitivity (d') for each participant and condition, calculated from hit rates (correct "same" responses) and false alarm rates (incorrect "same" responses to different pairs). Sensitivity data were analysed using both a 2 × 2 × 2 mixed-design ANOVA and a linear mixed-effects model (LMM) to account for the hierarchical structure of the data. Fixed effects included Complexity (Simple vs. Harder), Familiarity (Familiar vs. Novel), Orientation (Upright vs. Inverted), and their interactions.

The mixed-design ANOVA revealed a significant main effect of Complexity, $F_{(1, 51)} = 13.65$, $p < .001$, $\eta_p^2 = 0.21$, indicating that recognition sensitivity differed across stimulus complexity levels. A significant main effect of Orientation was also observed, $F_{(1, 51)} = 11.20$, $p = .002$, $\eta_p^2 = 0.17$, demonstrating a robust inversion effect following training. The main effect of Familiarity was not significant, $F_{(1, 51)} = 0.72$, $p = .399$, consistent with MKM theory, which predicts that familiarity benefits should be orientation-specific rather than generalised.

No significant two-way interactions were detected between Complexity and Familiarity, $F_{(1, 51)} = 0.03$, $p = .872$; Complexity and Orientation, $F_{(1, 51)} = 2.37$, $p = .13$; or Familiarity and Orientation, $F_{(1, 51)} = 1.99$, $p = .164$. The three-way interaction between Complexity, Familiarity, and Orientation was also not significant, $F_{(1, 51)} = 1.50$, $p = .227$. However, given strong a priori hypotheses grounded in MKM theory, we examined simple effects within each level

of Complexity, as the predicted inversion effect was expected to emerge selectively in the high-complexity condition.

Accordingly, we fitted a linear mixed-effects model examining the Familiarity × Orientation interaction separately for each complexity level. In the Harder checkerboard condition (N = 24), there was no significant main effect of Familiarity, $F_{(1, 153)} = 0.490$, $p = .485$, but there was a significant main effect of Orientation, $F_{(1, 153)} = 7.378$, $p = .007$, $\eta_p^2 = 0.046$. Critically, a significant Familiarity × Orientation interaction was observed, $F_{(1, 153)} = 4.051$, $p = .046$, $\eta_p^2 = 0.026$. Planned comparisons revealed that for familiar stimuli, performance was significantly better when stimuli were upright ($M = 2.35$, $SE = 0.150$) than inverted ($M = 1.87$, $SE = 0.163$), $t_{(153)} = 3.344$, $p = .0062$, $\eta_p^2 = 0.269$. No such advantage was observed for novel stimuli, $t_{(153)} = 0.498$, $p = 1$. This pattern indicates the emergence of orientation-specific processing mechanisms following categorisation exposure for complex stimuli.

Complexity	Familiarity	Orientation	Mean	SE
Harder	Familiar	Upright	2.35	0.15
		Inverted	1.87	0.16
	Novel	Upright	2.08	0.20
		Inverted	2.01	0.20
Simple	Familiar	Upright	2.80	0.12
		Inverted	2.68	0.11
	Novel	Upright	2.73	0.11
		Inverted	2.64	0.14

Table 2: Descriptive statistics for perceptual sensitivity (d') across groups.

In the Simple checkerboard condition (N = 29), upright stimuli ($M = 2.77$, $SE = 0.113$) yielded marginally higher sensitivity than inverted stimuli ($M = 2.66$, $SE = 0.126$), but this difference did not reach statistical significance. This suggests that the lower informational density of these stimuli permits successful discrimination via feature-based strategies that are relatively orientation-invariant.

Finally, although we planned to examine the relationship between categorisation learning rates and inversion magnitude, preliminary analyses revealed no significant correlation ($p > .05$). This suggests that explicit category learning and implicit perceptual tuning may rely on partially dissociable mechanisms. To maintain focus on the primary manipulations of Complexity and Orientation, these correlational analyses are not reported further.

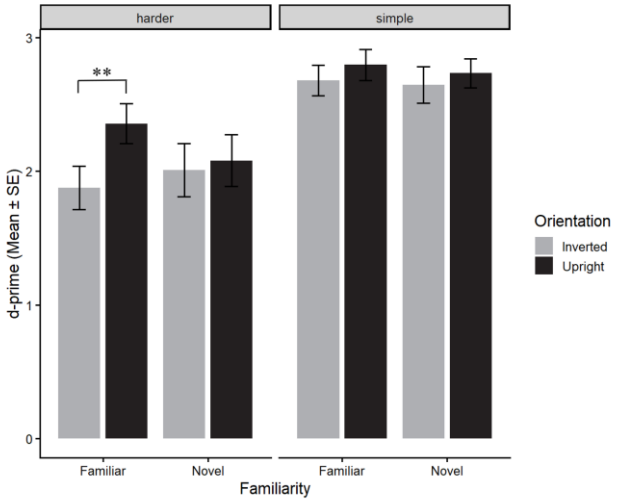

Figure 3: Perceptual sensitivity (d') as a function of Complexity, Familiarity, and Orientation. Mean d' scores are shown for Harder (left) and Simple (right) checkerboard conditions. In the Harder condition, a significant interaction between Familiarity and Orientation was observed, with higher sensitivity for Upright compared to Inverted stimuli only for familiarised patterns (**$p < .01$). In the Simple condition, no significant orientation effects were found, suggesting the use of orientation-invariant feature-based strategies. Error bars represent ±1 standard error of the mean (SEM).

Discussion

This study examined whether the inversion effect emerges in a matching-to-sample task using prototype-defined checkerboards and whether this effect is modulated by stimulus complexity. Our findings demonstrate that categorisation exposure is sufficient to induce an expertise-based inversion effect for artificial, non-face stimuli, but that this effect is most pronounced under conditions of high visual complexity. Specifically, participants showed a significant advantage for upright over inverted familiar checkerboards in the Harder condition, whereas this effect was attenuated and non-significant in the Simple condition. These results provide behavioural support for the MKM theory of perceptual learning (McLaren et al., 1989) and contribute to ongoing debates concerning the mechanisms underlying the Face Inversion Effect.

A central contribution of this work is the validation of the perceptual expertise hypothesis within a paradigm that minimises memory demands. Critics have argued that inversion effects in checkerboard studies may reflect memory retrieval strategies inherent to Old/New recognition tasks rather than changes in perceptual encoding. This critique is motivated by findings from Megreya and Burton (2006), who showed that matching performance for unfamiliar faces is often poor and lacks inversion effects, suggesting that holistic processing depends on stable

memory representations. However, our data challenge this interpretation. By demonstrating a robust inversion effect in a delayed matching task with minimal memory load, we show that expertise alters perceptual encoding itself. Thus, the "special" processing associated with expertise reflects changes in front-end representations rather than post-perceptual retrieval strategies, consistent with MKM theory's proposal that learning reshapes the contents of the visual sketchpad (Baddeley, 2000; McLaren et al., 1989).

The mechanism underlying this transformation is well captured by MKM theory's principle of differential latent inhibition. Through repeated exposure, common stimulus features become highly predictable and consequently down-regulated in salience via error-driven learning (McLaren & Mackintosh, 2000). In the present stimuli, this entails suppression of the redundant grid structure, thereby enhancing the diagnostic value of unique deviations. Inversion disrupts the spatial context required for these inhibitory predictions to operate, allowing the salience of common elements to dominate the representation and impairing discrimination (Civile et al., 2016). Our findings imply that this filtering process operates rapidly during perceptual encoding, facilitating matching of upright familiar stimuli while hindering processing of inverted stimuli.

Importantly, our manipulation of complexity reveals boundary conditions on the deployment of expertise-based processing. The inversion effect emerged robustly for complex stimuli but not for simple stimuli, suggesting that expertise mechanisms are engaged when computational necessity demands them. In the Simple condition, the relatively low number of features likely permitted serial, feature-based comparison strategies that are largely orientation-invariant, thereby attenuating inversion effects. In contrast, the dense and fragmented patterns in the Harder condition presumably exceeded the capacity of serial attention, forcing reliance on learned configural associations to filter redundant information. Thus, expertise-based processing appears to be recruited not merely by familiarity, but by the combination of familiarity and high informational load. This parallels face perception, where the inherent complexity of faces may explain why holistic processing emerges as the default mode for this stimulus class.

We observed no significant relationship between categorisation performance during training and the magnitude of the inversion effect during matching. Although this null result should be interpreted cautiously, it aligns with dual-process accounts of category learning (Ashby & Maddox, 2011), suggesting that explicit categorisation strategies may dissociate from the implicit perceptual tuning that drives inversion effects. Participants may learn to categorise checkerboards using explicit heuristics or single-feature cues, while the global salience modulation required for expertise-based encoding accumulates implicitly and independently. This dissociation further supports the view that inversion effects reflect low-level representational changes rather than high-level cognitive strategies.

Several limitations warrant consideration. The absence of a significant three-way interaction in the omnibus analysis suggests that the effect of complexity is graded rather than categorical. Future work could parametrically vary complexity to identify the precise threshold at which serial processing gives way to configural expertise. Moreover, while MKM theory predicts the observed behavioural patterns, direct neural evidence is needed to confirm the proposed salience-reduction mechanism. Future studies could employ EEG to measure N170 amplitudes during matching tasks; if our account is correct, upright familiar checkerboards should elicit reduced N170 amplitudes relative to inverted stimuli, paralleling established face-processing effects (Civile et al., 2014; Rossion et al., 1999). Additionally, transcranial direct current stimulation (tDCS) could be used to causally disrupt perceptual learning, as previous work has shown that anodal stimulation to left dorsolateral prefrontal cortex abolishes inversion effects in recognition memory (Civile et al., 2016). Whether similar disruption would occur in online perceptual encoding remains an open question.

In conclusion, this study demonstrates that inversion effects can be elicited in a pure matching task using artificial, complex stimuli, provided observers acquire category-specific expertise. These findings refute the notion that inversion effects are solely memory artefacts and instead support a domain-general account of visual expertise. We propose that what renders faces "special" is not their social relevance or biological geometry, but their statistical complexity, which necessitates error-driven perceptual learning mechanisms to manage high-dimensional visual input efficiently.

Reference

Ashby, F. G., & Maddox, W. T. (2011). Human category learning 2.0. *Annals of the New York Academy of Sciences*, *1224*(1), 147-161.

Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends in cognitive sciences*, *4*(11), 417-423.

Civile, C., McCourt, S., & McLaren, R. (2024). Modulate Perceptual Learning: Transcranial Direct Current Stimulation Reduces the Inversion Effect for Faces and Checkerboards. *The American Journal of Psychology*, *137*(3), 233-247.

Civile, C., Quaglia, S., Waguri, E., Ward, M., McLaren, R., & McLaren, I. (2021). Using transcranial Direct Current Stimulation (tDCS) to investigate why faces are and are not special. *Scientific reports*, *11*(1), 1-11.

Civile, C., Verbruggen, F., McLaren, R., Zhao, D., Ku, Y., & McLaren, I. (2016). Switching off perceptual learning: Anodal transcranial direct current stimulation (tDCS) at Fp3 eliminates perceptual learning in humans. *Journal of Experimental Psychology: Animal Learning and*

Cognition, *42*(3), 290.
Civile, C., Zhao, D., Ku, Y., Elchlepp, H., Lavric, A., & McLaren, I. P. (2014). Perceptual learning and inversion effects: Recognition of prototype-defined familiar checkerboards. *Journal of Experimental Psychology: Animal Learning and Cognition*, *40*(2), 144.
Diamond, R., & Carey, S. (1986). Why faces are and are not special: an effect of expertise. *Journal of Experimental Psychology: General*, *115*(2), 107.
Gauthier, I., & Tarr, M. J. (1997). Becoming a "Greeble" expert: Exploring mechanisms for face recognition. *Vision research*, *37*(12), 1673-1682.
Kanwisher, N. (2000). Domain specificity in face perception. *Nature Neuroscience*, *3*(8), 759-763.
McLaren, I. (1997). Categorization and perceptual learning: An analogue of the face inversion effect. *The Quarterly Journal of Experimental Psychology Section A*, *50*(2), 257-273.
McLaren, I., & Mackintosh, N. (2000). An elemental model of associative learning: I. Latent inhibition and perceptual learning. *Animal Learning & Behavior*, *28*(3), 211-246.
McLaren, I. P., & Civile, C. (2011). Perceptual learning for a familiar category under inversion: An analogue of face inversion?
McLaren, I. P., Kaye, H., & Mackintosh, N. J. (1989). An associative theory of the representation of stimuli: Applications to perceptual learning and latent inhibition.
Megreya, A. M., & Burton, A. M. (2006). Unfamiliar faces are not faces: Evidence from a matching task. *Memory & cognition*, *34*(4), 865-876.
Rossion, B., Delvenne, J.-F., Debatisse, D., Goffaux, V., Bruyer, R., Crommelinck, M., & Guérit, J.-M. (1999). Spatio-temporal localization of the face inversion effect: an event-related potentials study. *Biological psychology*, *50*(3), 173-189.
Yin, R. K. (1969). Looking at upside-down faces. *Journal of experimental psychology*, *81*(1), 141.
Yovel, G., & Kanwisher, N. (2004). Face perception: domain specific, not process specific. *Neuron*, *44*(5), 889-898.