# The effect of target-flanker congruency in visual crowding with complex scenes

Carly Sullivan

*This thesis is submitted in partial fulfilment of the Honours degree of Bachelor of Psychological Science*

The School of Psychology

University of Adelaide

October 2019

Word count: 9,477

**Table of Contents**

List of figures

Abstract

In visual crowding, the presence of surrounding clutter impairs recognition of a peripherally presented target object. Crowding is thought to be the consequence of processes that inappropriately integrate information from neighbouring objects. For simple stimuli (e.g., orientation bars) that have single features (e.g., colour), random or ‘incongruent’ clutter can have an even greater effect on recognition when the surrounding items are similar to the target item. Less is known about how crowding affects recognition when the surrounding items share a natural and global similarity with the target. The current study investigated the extent to which recognition is enhanced when the surrounding clutter is ‘congruent’ with the category of the target. Two experiments were developed to test the effect of target-flanker congruency on visual crowding. Using a recognition task and a discrimination task, it was found that when a target painting or scene is surrounded by other images of the same artistic style or natural category, visual crowding can be markedly reduced (Experiment 1) or even eliminated (Experiment 2). Visual crowding is widely thought to place a fundamental limit on perception, but the results presented here demonstrate that visual context may not always be a detriment to recognition in peripheral vision.

## Declaration

This thesis contains no material which has been accepted for the award of any other degree of diploma in any University, and, to the best of my knowledge, this thesis contains no material previously published except where due reference is made. I give permission for the digital version of this thesis to be made available on the web, via the University of Adelaide's digital thesis repository, the Library Search and through web search engines, unless permission has been granted by the School to restrict access for a period of time.

## Acknowledgements

Heartfelt thanks to my supervisor Dr Rachel Searston for guiding me through the research process. I am beyond grateful for the time that she has taken to help me expand my knowledge. No matter how unsure I was feeling at times, I always walked away from our meetings with renewed positivity, which is a credit to the kind and positive person that she is.

Looking back, I don't know how I would have got through this year without my lab family and the friends that have been made along the way. We had a shared understanding about what this year required of us, and we built a solid support network for each other.

I would also like to thank my friends, family and work colleagues for the interest that they have shown in this research project. Special thanks to JB, who helped to keep me motivated and focused on the big picture. This is an achievement that I am so proud of, and I couldn't have done it without you all.

# CHAPTER 1

## 1.1 Background

Central vision is what allows us to see things clearly; to find a familiar face in a crowd of faces, or locate the red jelly beans in a jar of multicoloured jelly beans. As you shift your gaze to each word of this sentence, notice how they come into sharper focus as they fall directly onto your fovea. Notice that while you are focusing on an individual word, the other words in the sentence outside of your central vision are less clear and more difficult to read. This is because these other words fall into your peripheral vision - an area that occupies as much as 99.9% of the visual field (Rosenholtz, 2016). As distance from the fovea (eccentricity) increases, the density of cone photoreceptors in the retina decreases, resulting in the lower spatial resolution of peripheral vision (Anstis, 1998). This reduction in acuity results in information loss, making it more difficult to resolve and recognise the finer details in a scene.

The surrounding elements or ‘clutter’ in a visual scene also affects recognition, particularly in peripheral vision. For example, a single letter is relatively easy to identify when presented alone in peripheral vision, but when that same letter is surrounded by other letters it becomes more difficult to discern (Figure 1) (Huckauf, Heller & Nazir, 1999). This negative effect of clutter on recognition is called visual crowding, and it is thought to set a fundamental limit on visual perception (Pelli & Tillman, 2008; Whitney & Levi, 2011). Although recognition can be substantially impaired by visual crowding, not all meaningful information is lost in clutter (Rosenholtz, 2016). This thesis will investigate the extent to which visual crowding can be reduced, thus recognition enhanced, when the surrounding clutter is matched with the object in question.


*Figure 1.* Letter crowding*:* Focus your gaze on the red cross and notice how the letter on the left side is easy to identify when it is presented alone in peripheral vision. Keeping your eyes fixed on the cross, notice how that same letter on the right side is more difficult to identify when it is surrounded by other letters. This is visual crowding.

Visual crowding was initially observed with letters (Bouma 1970; Korte, 1923), and has subsequently been shown to occur with a wide variety of target stimuli ranging from relatively simple images such as Gabor patches and orientation bars (Greenwood, Bex, & Dakin, 2010; Yeotikar, Khuu, Asper & Suttle, 2011), to more complex images such as faces (Farzin, Rivera & Whitney, 2009; Louie, Bressler, & Whitney, 2007), everyday objects (Wallace, & Tjan, 2011) and natural scenes (Gong, Xuan, Smart, & Olzak, 2018). When a peripheral target stimulus is crowded by surrounding or 'flanking' stimuli, its presence is still detected, but it is often perceived as indistinct, jumbled or textural in appearance (Levi, Hariharan & Klein, 2002; Pelli, Palomares & Majaj, 2004). Such distorted representations highlight the difficulty the visual system has in demarcating the boundaries and features of the target stimulus from those of the flanking stimuli (Greenwood et al., 2010). Visual crowding has been described as a process where features of the target and its flankers become excessively integrated, 'pooled', or averaged over an inappropriately large integration field (Levi, 2011; Parkes, Lund, Angelucci, Solomon & Morgan, 2001). This 'faulty integration' occurs when the centre-to-centre distance between target and flanking stimuli is less than half the target's eccentricity (Bouma, 1970; Pelli & Tillman, 2008; Strasburger, & Malania, 2013). Recognition can be restored by increasing the spacing between the target object and its flankers to a critical distance that scales with eccentricity, as described by 'Bouma's Law' (Figure 2) (Bouma, 1970).

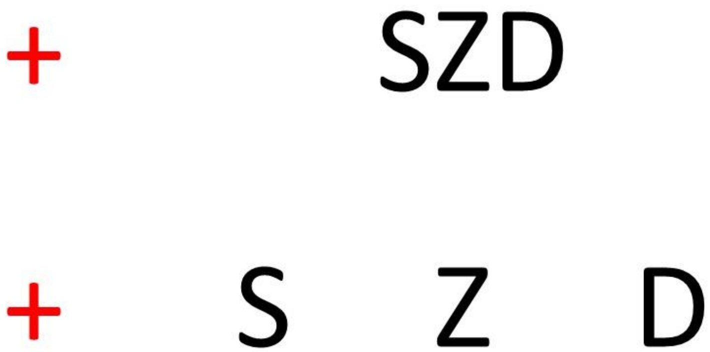



*Figure 2.* Critical distance: When fixating on the red cross at the top, notice how the target letter 'Z' is difficult to recognise when it is closely surrounded by other letters. Now, focus on the cross below and notice how that same target letter 'Z' is easier to recognise when the distance between the letters is increased. This is an example of the critical spacing that is required for recognition (Bouma's Law).

Previous behavioural studies suggest that visual crowding occurs at various stages in the visual hierarchy (Kimchi and Pirkner, 2015; Reuther & Chakravarthi, 2014; Yeotikar et al., 2011). 'Crowding' interferes with low-level feature dimensions such as orientation (Greenwood et al. 2010; Parkes et al., 2001), colour, and motion (Bex & Dakin, 2005; Van den Berg, Roerdink & Cornelissen, 2007), and it also affects high-level recognition of more complex stimuli such as words (Yu, Akau, & Chung, 2012), faces (Farzin et al., 2009; Louie et.al, 2007), and scenes (Gong et al., 2018). Interestingly, some information seems to survive crowding, as people can still recognise facial expressions (Fischer & Whitney, 2011; Kouider, Berthet & Faivre, 2011) and word meaning (Peng, Zhang, Chen, & Zhang, 2013) under cluttered conditions. Neuroimaging studies have indicated that the neural origin of crowding occurs as early as the primary visual cortex (V1), though further research is required to ascertain how higher cortical areas contribute to the crowding effect (Millin, Arman, Chung, & Tjan, 2013).

Crowding has critical implication for everyday navigation and functioning as features and objects are not viewed in isolation, but rather within a rich visual context. Crowding impairs the ability to identify and respond to stimuli within clutter, as it slows saccadic visual search (Vlaskamp & Hooge, 2006) and interferes with the precision of visually guided actions, such as the grasping of objects (Bulakowski, Post, & Whitney, 2009). A better understanding of how the recognition bottleneck caused by visual crowding functions, could provide insight into how we integrate information to form a rich visual representation of our environment (Cohen, Dennett & Kanwisher, 2016).

**1.2 Target-Flanker Relations**

The magnitude of visual crowding can be modulated by target-flanker similarity (Bernard & Chung, 2011; Kooi, Toet, Tripathy & Levi, 1994), and the effect increases when the target and flanking stimuli are similar in shape (Kooi et al., 1994, Nazir, 1992), colour (Põder, 2007), orientation (Andriessen & Bouma, 1976) or size (Van den Berg et al., 2007). The spatial configuration of flankers also influences crowding, for instance, collinear circle configurations of Gabor patches induce little or no crowding on a central target Gabor, while circular arrangements that lack collinearity induce a stronger crowding effect (Livne & Sagi, 2007). Feature and configurational accounts of crowding propose that crowding is weak when the target stands out from the stimulus array and strong when it perceptually groups with the surrounding flankers (Livne & Sagi, 2007; Saarela, Sayim, Westheimer, & Herzog, 2009; Yeotikar et al., 2011).

Many studies have tested target-flanker similarity using simple stimuli (e.g., orientation bars) with single and known features (e.g., colour), but few studies have manipulated such target-flanker relationships with naturally varying stimuli akin to what people see every day. One reason for the limited use of natural stimuli is that similarity can be difficult to quantify and features can be difficult to define (Yu, et al., 2012). Bernard and Chung (2011) tested the effects of target-flanker similarity with letters, devising a psychometric similarity matrix that quantified how similar letters are to one another. Using a range of target-flanker manipulations, they found that crowding is stronger when the target and flanker letters exhibit higher similarity. Alpha-numerical stimuli have also been used to test the effect of target-flanker category membership on visual crowding. Letters and numbers crowd more when they are surrounded by flankers from the same category as the target, (e.g., letters impair recognition of other letters more than numbers) (Huckauf et al.,1999; Reuther & Chakravarthi, 2014). This categorical similarity effect was found to persist even when featural differences between the letter and number categories were controlled, suggesting that similarity is not simply due to low-level feature similarity, but is also influenced by higher-level processing of meaning (Reuther & Chakravarthi, 2014).

There are some notable exceptions to the target-flanker similarity effects observed in visual crowding studies. For instance, upright faces crowd more when flanked with other upright faces, however the effect does not hold when the target face and it's flankers are inverted (Louie, et al., 2007). Instead, inverted target faces crowd more when the flankers are presented in a *dissimilar* orientation to the target. If similarity were responsible for increased crowding across the board, then crowding should be stronger when the inverted targets are surrounded by similarly inverted flankers. The explanation given for these results was that inversion interferes with the holistic processing of faces, and therefore, flanker orientation has no effect when the targets are meaningless (Louie et al., 2007). Ester, Klee and Awh (2014) also found crowding to be stronger with dissimilar targets and flankers in a study using clock face stimuli. When the orientation angle of the flankers was dissimilar to that of the target, crowding was strongest. Taken together, these findings suggest that dissimilar flankers do not preclude strong crowding effects, and target-flanker similarity is not only influenced by crowding at the elementary feature level but also at higher levels of processing. Various models of visual crowding have been developed to explain some of the mechanisms that are thought to give rise to these target-flanker relationships.

**1.3 Models of Visual Crowding**

When asked to identify a crowded object, observers make several types of response errors. Analysis of these mistakes has revealed that perceptual errors are not random, but biased by the identities of adjacent flankers (Harrison & Bex, 2017; Poder & Wagemans, 2007). Perhaps the most parsimonious explanation of visual crowding to arise from these analyses is that of the substitution model, whereby an observer mistakenly reports the identity of a flanker instead of the target (Ester et al., 2014; Freeman, Chakravarthi, & Pelli, 2012; Harrison & Bex 2017; Poder & Wagemans, 2007; Strasburger, 2005). When substitution occurs, information is extracted from and reported about one object only, however it remains unclear as to whether these errors result from positional uncertainty, unfocused spatial attention, or the observer simply failing to see the target and reporting a flanker instead (Freeman et al., 2012; Huckauf, & Heller, 2002; Strasburger, 2005; Whitney & Levi, 2011).

An alternative explanation of crowding is given by simple pooling models. Pooling is thought to occur when an observer reports a summary of the stimulus set, such as the average orientation of Gabor patches despite that particular orientation not matching that of the crowded target (Parkes et al., 2001). People can also detect the average emotion from a set of faces under crowded conditions, suggesting that complex information can still be extracted from a cluttered visual display (Fischer & Whitney, 2011). Simple pooling models hypothesise that features from multiple nearby flankers are pooled over some local region, resulting in an average or summary statistical representation of the stimulus array (Keshvari, & Rosenholtz, 2016; Parkes et al., 2001). Indeed, information about an individual stimulus is lost and replaced with information about distributions of features (Balas, Nakano, & Rosenholtz, 2009). Featural pooling can also lead to incorrect reporting of a flanker, due to the flanker sharing similar features with the sum of the pooled information (Greenwood et al., 2010; Parkes et al., 2001).

Pooling and substitution are two of several models of visual crowding that describe some of the perceptual errors that occur when stimuli are surrounded by clutter, though current models of crowding are yet to give a complete or unifying account of the phenomenon. The utility of reducing a complex and multivariate process into a simple univariate model has been questioned (Agaoglu & Chung, 2016), as a common problem to all models of crowding is that they fail to generalise beyond the types of stimuli used in testing. Studies that support pooling tend to use gratings or orientation bars which are easier to perceive and report as an average, whereas studies that favour substitution models typically use alphanumeric stimuli which are more difficult to meaningfully perceive and report as an average (Strasburger 2005, Strasburger, & Malania, 2013).

A demonstration given by Pelli et al. (2004) showed that when peripherally viewed letters are surrounded by identical letters, they appear normal. Interestingly, when there is a mixture of different letters in an array they take on a confusing and somewhat hybrid appearance (Figure 3). This demonstration gives rise to the question: Can recognition be enhanced when the target is crowded by matching, or congruent, visual context rather than incongruent visual clutter?

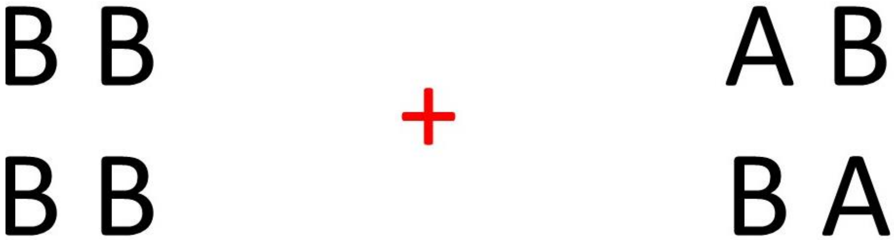



*Figure 3.* Congruent clutter*:* Keep your eyes fixed on the red cross and notice how each of the letters on the left side appear normal when they are surrounded by the same, or 'congruent' letters. However the letters on the right side appear less defined when surrounded by different, or 'incongruent' letters.

**1.4 Perceptual Gist**

People have a remarkable ability to extract information distributed across the visual field, they can glean the global properties or 'gist' of a novel scene within a fraction of a second (Greene & Oliva, 2009), and reliably discriminate between different naturalistic image sets at resolutions as low as four pixels (Searston, Thompson, Vokey, French & Tangen, 2019). Gerhard, & Bethge (2014) found that non-experts are able to discriminate the artistic style of drawings based on textural and image statistics alone, indicating that people are sensitive to stylistic variations in artwork. Artwork offers a particularly salient example of how naturally varying visual stimuli can be categorised based on their gist or general appearance. Although subject matter may vary between paintings from the same artist, genre or era, there is a global quality conveyed within an artistic style that remains constant, though difficult to describe (Abramov, Farkas, & Ochsenschlager, 2006; Gerhard, Wichmann, & Bethge 2013). These studies illustrate how people can process the gist of a visual category, even when the semantic content and finer detail of the stimulus is obscured.

The process of extracting gist information from a scene is analogous to the process of pooling or extracting an average from crowded items in peripheral vision. Indeed, studies on scene perception show that peripheral vision is particularly useful for extracting such global properties (Larson & Loschky, 2009). A recent investigation found that people could identify the gist of a scene even when crowded by a random or cluttered array of other scenes (Gong et al., 2018), although scene recognition was significantly better when presented in isolation, the crowding conditions always consisted of random scenes that were not congruent with the target category. These findings are consistent with other classic examples of visual crowding, where an averaging or pooling process is operating against successful recognition of the target.

Visual crowding may make it difficult to identify an image of a beach as a natural scene when it is flanked by urban images of tall buildings and peak hour traffic, but what if that same beach scene is flanked by images of mountains, forests and other natural scenes? Likewise, what happens when you surround Monet's Water Lilies with other Impressionist paintings? In this thesis, two experiments are proposed and designed to exploit the pooling or positional uncertainty processes as described in the visual crowding literature, but for the benefit, rather than to the detriment of extracting the gist of the target. In other words, can visual crowding be reduced when the clutter is congruent with the gist of the target category?

### 1.5 Current project

The current project tests participants ability to distinguish visual categories under conditions of crowding. In particular, it will be examined how flanking a target Impressionist painting, or natural scene in the peripheral vision with other images of the same artistic style or natural category affects the participants ability to correctly classify them. It is of primary interest to know if the gist of the target image is better detected when the flankers are congruent (e.g., a natural scene surrounded by other natural scenes), versus when they are incongruent (e.g., a natural scene surrounded by urban scenes). Participants will classify and discriminate random samples of natural and urban scenes, as well as random samples of Cubist and Impressionist paintings. It should be noted that the intention of the study is not based on natural scene recognition or the detection of artist style per se, instead participant responses will be averaged across the two image sets, creating a more generalisable measure of visual crowding using naturally varying stimuli.

In Experiment 1, participants classify paintings as “Cubist” or “Impressionist” with flanking paintings of the same (congruent) or different (incongruent) artistic style. As an uncrowded control, the same paintings are presented alone, that is, without flankers. Similarly, participants will also classify scenes as “natural” or “urban” under both uncrowded and crowded conditions with congruent and incongruent flankers. The image sets are not treated as an explicit factor in this experiment but are instead included to assess the generalisability of any observed congruency × crowding interaction effects. In Experiment 2, participants will be asked to discriminate whether two paintings are the same artistic style, or different. The same task will also be completed for images of scenes, again while manipulating crowding and congruency conditions. The two experiments offer a test of the same predictions across two different types of tasks.

A guiding hypothesis across the two experiments is that target discriminability will be enhanced when the flankers are from the same category, sharing a similar gist. From this premise, the following four predictions are made:

**Prediction 1:** It is expected that participants will be able to discriminate paintings and scenes in their peripheral vision to a greater extent than if these choices were randomly made (i.e. greater than chance).

**Prediction 2:** The presence of flanking images that are incongruent with the target will reduce the discriminability of the target, as demonstrated elsewhere in classic visual crowding literature.

**Prediction 3:** Discriminability will be higher when the images are presented in isolation than when crowded.

**Prediction 4:** The classic visual crowding effect will be reduced when the flankers are congruent with the target image.

# CHAPTER 2: Experiment 1

This first experiment was designed to test people's ability to correctly classify a target painting as 'Cubist' or 'Impressionist' when it is presented in peripheral vision. The target is either presented alone (uncrowded condition) or flanked by four other paintings (crowded condition). The flanking paintings were either from the same category (congruent flankers) or a different category (incongruent flankers) to the target image. Participants also repeat the task with images of scenes in a separate block of trials, classifying them as 'natural' or 'urban'. The aim of this study is to see if the images are easier to classify when they are surrounded by items from the same category, than when they are surrounded by images from a different category.

## 2.1 Method

It is important to the scientific endeavour that research practices embrace the value of openness, honesty and transparency, therefore this project has been preregistered on the Open Science Framework (OSF), which can be viewed at the following short link: http://tiny.cc/mwafcz.

**2.1.1 Participants.** Twenty four members of the general public self-selected into this study. The sample comprises 4 (16.7%) males and 20 females (83.3%), within the age range of 20 and 57 years ($M$ = 32.7, $SD$ = 9.6). The study was advertised using social media and paper flyers that were placed on community billboards (Appendix A). Each participant was offered a $20 gift voucher as an incentive to complete the experiments. All participants met the minimum age requirement of 18 years, reported normal or corrected to normal visual acuity and spoke fluent English. Ethics approval was granted by the University of Adelaide's Human Research Ethics Committee (19/68) and informed consent was collected from all participants prior to undertaking the study (Appendix B). Data collection took place over eight weeks at the University of Adelaide North Terrace campus, South Australia.

An a priori power analysis indicated that one-hundred and ninety-two observations from 24 participants per crowded and uncrowded conditions provided sufficient power (power = .985) to detect a smallest crowding main effect of interest of $d = 0.45$. This effect size is a conservative estimate based on previously reported crowding effect sizes (e.g., Gong et.al, 2018), accounting for possible inflation due to publication bias (Munafò et al., 2017), and given an $\alpha = 0.05$ (for a rational and discussion on determining a smallest effect sizes of interest, see Lakens, Scheel & Isager, 2018).

**2.1.2 Stimuli.**

***Paintings.*** The paintings are a subsample from the 'How Low Can You Go?' collection (Searston et al., 2019). The full collection contains 5,184 paintings, made up of 18 different paintings by 72 different artists, in each of four different artistic styles (Cubism, Impressionism, Realism, and Renaissance; 288 artists in total, available at https://osf.io/kuja8/). The current experiments used a total of 2,592 paintings that included 1,296 Cubist and 1,296 Impressionist paintings. All of the paintings were originally cropped to the centre of the shortest dimension using a square 1:1 aspect ratio. The images were resized using nearest neighbour scaling to 256×256 pixels and then converted to jpeg format. All signatures were removed using the "Content Aware" fill tool in Photoshop.

***Scenes***. The scenes were a subsample of urban and natural scenes from the LabelMe dataset (Oliva & Torralba, 2001; available at http://cvcl.mit.edu/database.htm). The 'natural' category contained 1,472 scenes that comprised 359 beach, 328 forest, 374 mountain, and 411 open country scenes. The 'urban' category consisted of 1,216 scenes that included 260 highway, 308 inside city, 292 street, and 356 tall building scenes. All of the images were in colour, jpeg format, and 256×256 pixels.

Pilot data were collected from two members of the research lab. Analysis of this data revealed a mean proportion correct of 86% for the crowded condition and 93% for the uncrowded condition. To increase the difficulty of the task and the sensitivity of the experiment, stimulus exposure time was decreased from 100 milliseconds to 34 milliseconds and the image resolution was reduced to 32×32 pixels (upscaled to 256×256) (Figure 4). The decision to reduce the presentation time to 34 milliseconds was based on a previous study by Greene and Oliva (2009) who found that the time required to recognise properties of scenes ranged from 19 milliseconds to 47 milliseconds. The average threshold for classifying global properties such as 'naturalness' was 34 milliseconds.

During the experiment, the stimuli were presented on a white background and viewed from a distance of 57cm. Each of the target and flanker images subtended 4°×4°of visual angle, the target stimuli were presented at 11° eccentricity and the centre-to-centre spacing between target and flanker was 4.5°, in accordance with Bouma's Law (Bouma, 1970).

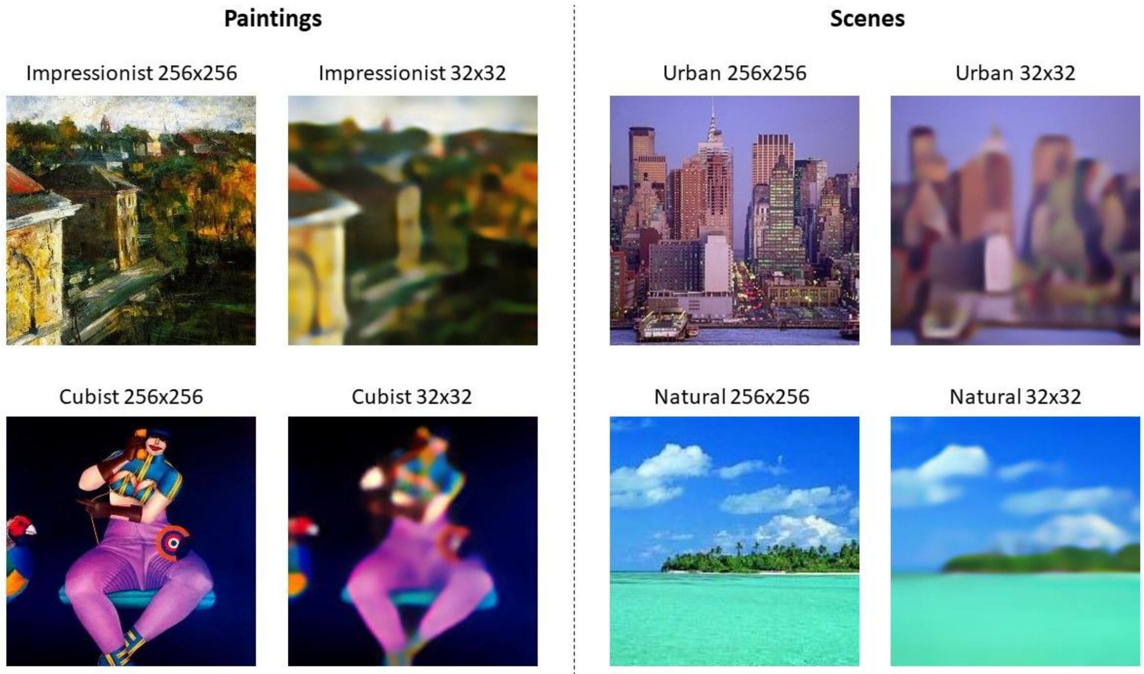



*Figure 4.* Image resolution: Example images from each of the categories for both stimulus sets displaying the original image resolution of 256×256 pixels, and the reduced image resolution 32×32 pixels.

**2.1.3 Apparatus.** The video instructions and task were presented to participants on an elevated 15-inch Apple Macintosh laptop screen with over-ear headphones. The images were viewed binocularly at a distance of 57cm and head position was stabilised using a forehead and chin rest. Responses were entered using a keyboard. The software used to generate the trial sequences, present stimuli to participants, and record their responses were developed in LiveCode (version 9.0.4; the open source 'Community Edition').

**2.1.4 Experimental design.** The experiment uses a 2 (crowding: uncrowded, crowded) × 2 (congruency: congruent, incongruent) fully repeated measures design to test the effect of congruency on peoples' classification of paintings and scenes. Each participant classified 48 paintings and 48 scenes in each of the two crowding conditions and the two congruency conditions. The stimulus sets were not treated as a factor in the design as the plan was to collapse across them during data analysis to increase the generality of the findings. Accuracy was measured by comparing the participants judgements about the category membership of the target stimulus to the true category membership of the target. Discrimination performance was also assessed based on the 'hits' and 'false alarms' in each condition.

To check the integrity of the experiment code and custom software, 24 simulated participants completed the experiment. These simulated participants were programmed to provide a random response on each trial at a random response time between 0 and 3000ms and they represent chance baseline responding. Each simulation was paired with a human participant on trial sequence and stimuli, therefore the plan was to treat them as a set of paired random observations for the purpose of the analyses.

**2.1.5 Sequencing.** Unique trial sequences were pre-generated for 24 participants, with paintings and scenes being randomly sampled for each participant. Target and flanker images were 'yoked' across crowding blocks to minimise noise, but the images were never repeated within each crowding block. A different random sample of images (sampled without replacement) are presented in the congruent and incongruent trials to reflect the randomisation within the crowded and uncrowded blocks. The paintings and scenes were counterbalanced across participants, such that a random half classified the paintings first, and the other half classified the scenes first. Impressionist and Cubist target trials were presented in a different random order to each participant within the paintings block, as were the natural and urban target trials within the scenes block. The crowded and uncrowded trials for each stimulus set were presented in two separate blocks and counterbalanced across participants. Within these blocks the congruent and incongruent trials were also presented to each participant in a different random order.

**2.1.6 Procedure.** Participants first read an information sheet about the experiments (Appendix C) and then watched an instructional video that detailed the nature of the materials, the image classification task and method of responding (http://tiny.cc/qhtzcz). Example images of paintings and scenes were included to familiarise participants with the two category distinctions for each stimulus set. Participants completed a total of 348 trials that were divided into four even blocks. Half of the trials contained images of Cubist and Impressionist paintings and the other half of trials contained images of natural and urban scenes. The stimulus sets and target image conditions were specified at the beginning of each block of trials so that the participants knew whether to expect paintings or scenes, and if they would appear alone or flanked by four other images. Using a two alternative forced choice task, participants had to indicate which category the target image belonged to (e.g., Natural/Urban or Cubist/Impressionist).

On each trial, a fixation point was displayed in the centre of the screen for 500 milliseconds, followed by a single target image or a target image surrounded by four flanking images (Figure 5). The flankers in the crowded condition were either congruent with the category of the target image or incongruent with the category of the target image. The stimuli appeared randomly on the left or right side of the fixation point for 34 milliseconds, a time frame that precludes a saccade to an unexpected stimulus (Carpenter, 1988). The fixation point remained on the screen with the images and participants were instructed to keep their gaze on it until it disappeared. Immediately following the disappearance of the target, a screen was presented prompting participants to indicate their response by pressing ‘C’ or ‘M’ on the keyboard as quickly and accurately as possible (Figure 6). If the response time exceeded 3000ms, a reminder was given to respond within this time frame. Participants were urged to take breaks during the block intervals to minimise the effects of fatigue, and the duration of the experiment was about 20 minutes. No feedback was given about the accuracy of their responses.

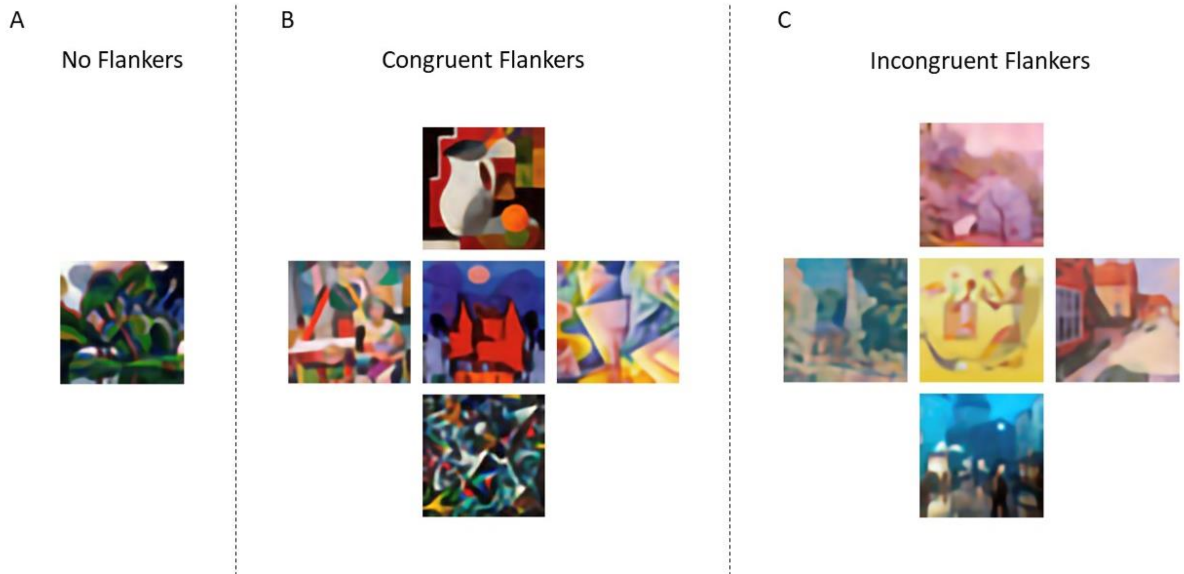



*Figure 5.* Experiment 1 stimulus configuration examples**: Panel A)** A target Cubist painting presented alone (uncrowded condition). **Panel B)** A Target Cubist painting that is surrounded with other Cubist paintings (crowded with congruent flankers). **Panel C)** A target Cubist painting that is surrounded by Impressionist paintings (crowded with incongruent flankers).

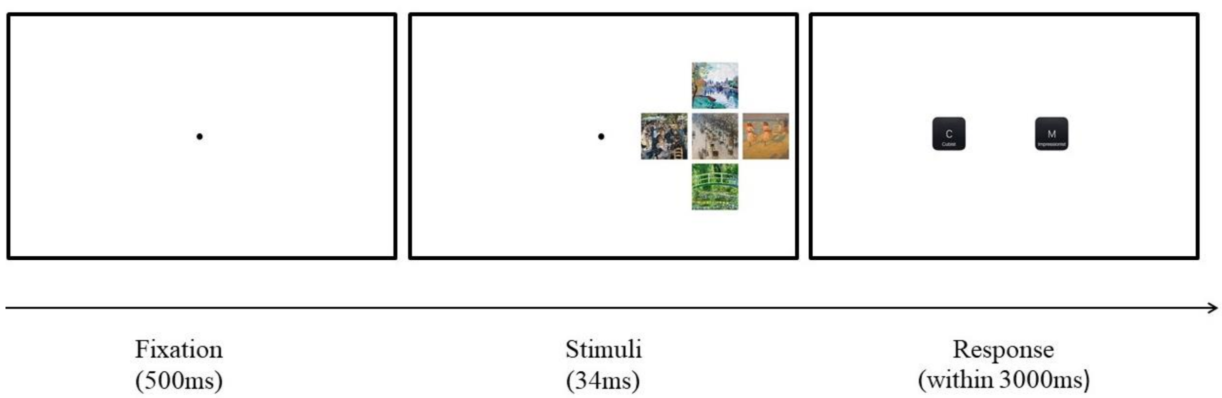



*Figure 6.* Experiment 1 trial sequence with an example of the stimulus configuration for the crowded condition. The target is the centre image and the four surrounding images are the flankers.

## 2.2 Results

Prior to analysis, participants data was screened to ensure that it did not meet any of the exclusion criteria that was stipulated at pre-registration. Univariate density plots and boxplots were visually inspected to assess the distributional nature of the data and to identify outliers for all variables (Appendix D). The distributions for the incongruent flanker condition and the crowded condition were negatively skewed, but all outliers were retained as none met the pre-specified criteria for exclusion, and it was impossible to determine if they were caused by an experimental artefact or genuine variation in responding. As a result of the outliers noted in some conditions (e.g., see the crowded incongruent condition in Figure 8), all preregistered analyses were conducted as planned, with an additional set of non-parametric analyses performed to check the robustness of the results to different distributional assumptions. Given that the pattern of results was identical with both sets of analyses, only the preregistered analyses are reported in the main text for brevity and clarity. The additional non-parametric analyses are provided in Appendix G. All data was analysed using R (version 3.6.1), and the R code and output for the plots and analyses are located in Appendices D-F.

**2.2.1 Accuracy.** Each participant's proportion of correct responses was calculated for each of the conditions, collapsing across the two image sets. Their mean proportion correct was higher when the images were uncrowded ($M$ = 0.80 , $SD$ = 0.07) compared to when they were crowded, ($M$ = 0.65, $SD$ = 0.14), and in the crowded conditions mean proportion correct was higher when the flankers were congruent with the target ($M$ = 0.73, $SD$ = 0.07) , compared with when they were incongruent ($M$= 0.58, $SD$ = 0.15), see figure 7.

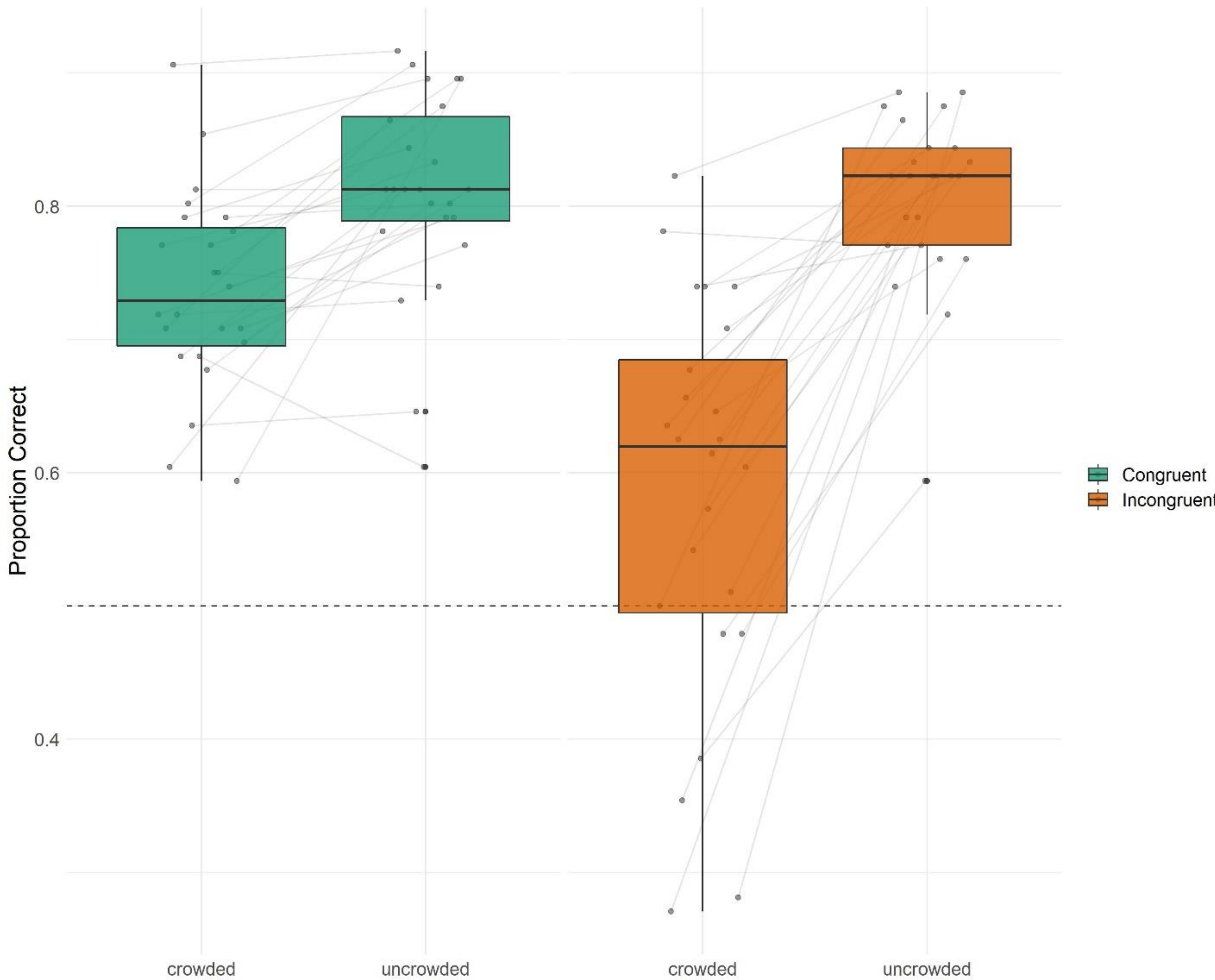



*Figure 7*. Proportion correct boxplots showing the proportion correct for the uncrowded and crowded conditions with congruent and incongruent flankers. Chance performance is indicated by the dotted line (.5).

**2.2.2 Sensitivity.** To compute participants' sensitivity to the target stimuli, participants responses in each condition (collapsing over stimulus sets) were coded as hits and false alarms. A hit in this case is when the participant says "Cubist" and the target is a Cubist painting, or when the participant says "urban" and the target is an urban scene. A false alarm is when the participant says "Cubist" and the target is an Impressionist painting or when the participants says "urban" and the target is a natural scene. Participants' hit and false alarm rates were then used to compute their sensitivity for each condition using a non-parametric model that averages the minimum and maximum proper Receiver Operating Characteristic (ROC) curves through a point

(*A*). An *A* value of 1 indicates perfect sensitivity (e.g., Maximum number of hits and zero false alarms on signal present trials), and a value of .5 indicates chance performance (e.g., half hits, half false alarms). The boxplots in figure 8 offer a visual representation of participants sensitivity for each of the conditions.

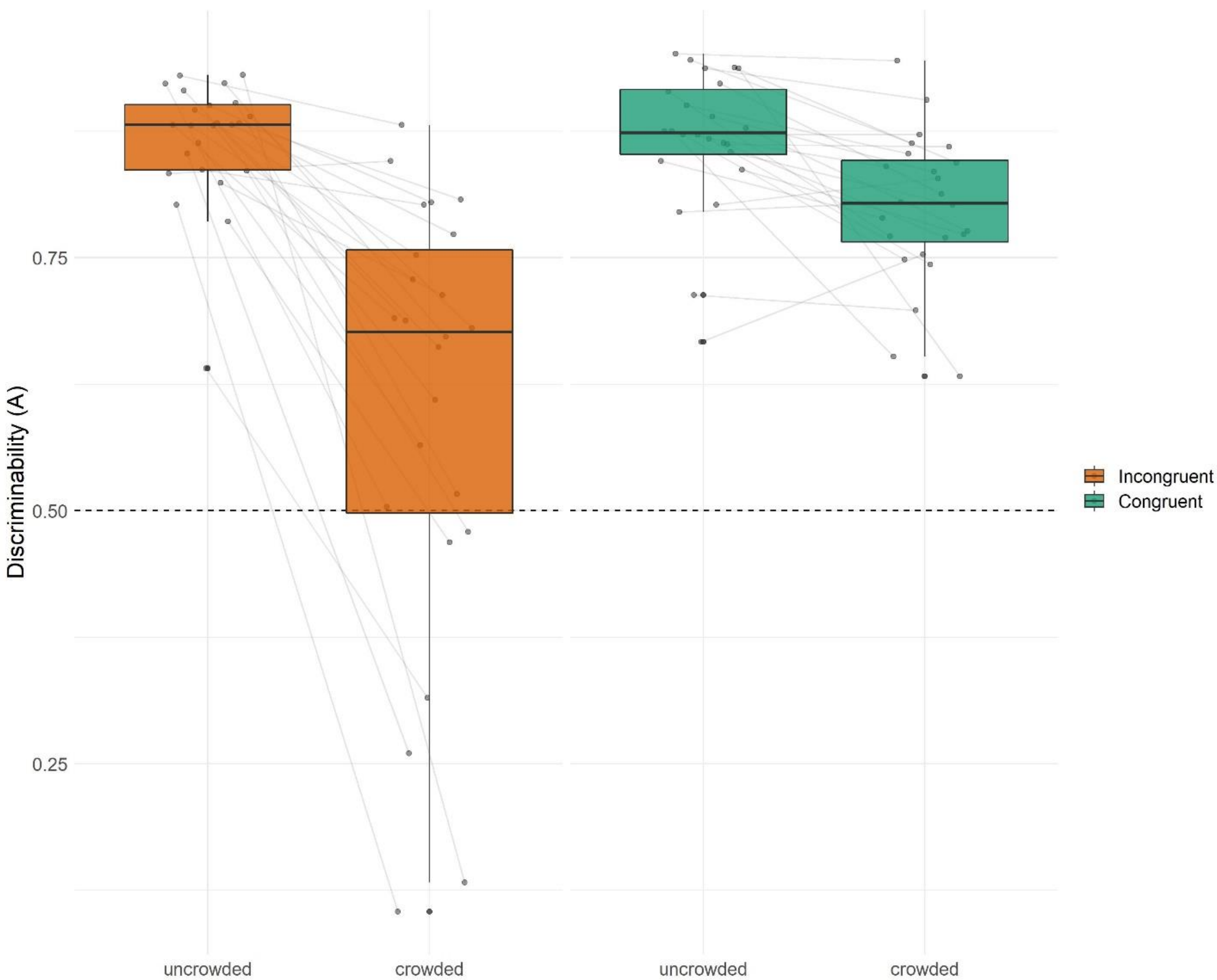



*Figure 8.* Discriminability boxplots showing discriminability performance for the uncrowded and crowded conditions with congruent and incongruent flankers. Chance performance is indicated by the dotted line (.5).

**2.2.3 Comparison to chance.** A paired samples t-test was conducted to compare participants *A* scores to those of simulated participants across each condition. The simulations were programmed to respond randomly in the experiment and reflect chance performance (.5). There was a statistically significant difference between the *A* scores, $t(95) = -15.70$, $p < 0.001$, 95% CI [-0.31, -0.24], $d = -2.18$, supporting prediction 1, that participants can discriminate paintings and scenes in their peripheral vision better than chance.

**2.2.4 Two-way factorial ANOVA.** As planned, participants *A* scores were submitted to a two-way repeated measures ANOVA, with crowding (crowded, uncrowded) and congruency (congruent, incongruent) as within-subjects factors. The analysis showed a significant main effect of crowding on participants' sensitivity, $F(1, 23) = 37.21$, $p < 0.001$, $\eta_G^2 = 0.31$). This finding is consistent with prediction 3, that discriminability of the target image is higher when it is uncrowded ($M = 0.86$, $SD = 0.06$) compared to when it is crowded ($M = 0.70$, $SD = 0.18$). A significant main effect was also found for congruency, $F(1,23) = 31.19$, $p < 0.001$, $\eta_G^2 = 0.14$, such that participants discriminability was higher when the flankers were congruent with the target image ($M = 0.83$ , $SD = 0.07$), compared to when they were incongruent with the target image ($M = 0.73$ , $SD = 0.20$), thus supporting the prediction 2. Importantly, there was also significant interaction between crowding and congruency, $F(1,23) = 34.85$, $p < 0.001$, $\eta_G^2 = 0.13$. Crowding effects were greater when the flankers were incongruent with the target category and this resulted in reduced sensitivity ($M = 0.60$, $SD = 0.21$), compared to when the flankers were congruent with the target category ($M = 0.79$, $SD = 0.07$) and this significant interaction effect supports prediction 4.

In order to determine the extent to which congruent flankers reduce visual crowding, a post hoc paired samples t-test was conducted comparing participants *A* scores under crowded versus uncrowded conditions with congruent flankers, $t(23) = -4.10$, $p$ <0.001, 95% CI [-0.10, -0.03]. A second paired samples $t$-test was also conducted to compare participants $A$ scores under crowded and uncrowded conditions with incongruent flankers $t(23) = -6.3$, $p$ = <0.001, 95% CI [-0.34, -0.17] with application of the Holm correction. The participants were better at discriminating the target when it was uncrowded compared to when it was crowded, but this difference was smaller with congruent flankers ($d = -0.95$) than incongruent flankers ($d = -1.64$).

# CHAPTER 3: Experiment 2

A second experiment was conducted further probing the generality of crowding × congruency interaction effects using a discrimination task that eliminates the need to know the semantic content or category of the peripheral images. For example, there is no need to recognise a natural scene as "natural" in order to recognise a pair of natural scenes as the "same". The participants, materials, apparatus and design are the same as in Experiment 1, with the exception of the task, stimulus presentation time and number of trials. The experiment order was counterbalanced so that a random half of the participants completed Experiment 1 first and the other half completed Experiment 2 first, and both experiments were completed on the same day. This experiment consisted of 256 trials that were divided into four even blocks for the crowded and uncrowded conditions of each stimulus set. Within each of the crowded blocks a random half of targets were presented with congruent flankers and the other half with incongruent flankers.

## 3.1 Method

**3.1.1 Procedure**. Participants were presented with two peripheral target images, one to the left and one to the right of a central fixation point. Their task was to indicate whether the images were belonged to the 'same' category of style or two 'different' categories. To ensure that this was a purely perceptual discrimination task, participants were not provided with the semantic category labels for either of the stimulus sets during the instructional video (http://tiny.cc/eitzcz). Although it was possible for the participants who completed Experiment 1 first to have learned the target categories without any feedback, this knowledge was not necessary to perform well in Experiment 2, and the counterbalancing ensures that only half of the participants had the opportunity to learn the categories from the first experiment.

Each trial commenced with a central fixation point for 500 milliseconds, followed by the simultaneous presentation of the target images for 64 milliseconds (figure 9). The targets appeared alone or flanked by four other images. In the crowded condition, both targets were either flanked with images from the same category (congruent flankers), or one of the target images was flanked with images of a different category (incongruent flankers) (Figure 10). For example, if participants are presented with two target Impressionist paintings, the correct response is “same”. In this example the left and right flanking paintings are all Impressionist paintings in the congruent condition (i.e., the flankers are congruent with the correct response of “same”). In the incongruent condition, on the other hand, the left flankers are Impressionists and the right flankers are Cubists such that the flankers are in line with an incongruent response of “different”. That is, the left and right flankers are either congruent or incongruent with the correct response, and vice versa for different category trials. The duration of this experiment was about 13 minutes and no feedback was given about the accuracy of responses.

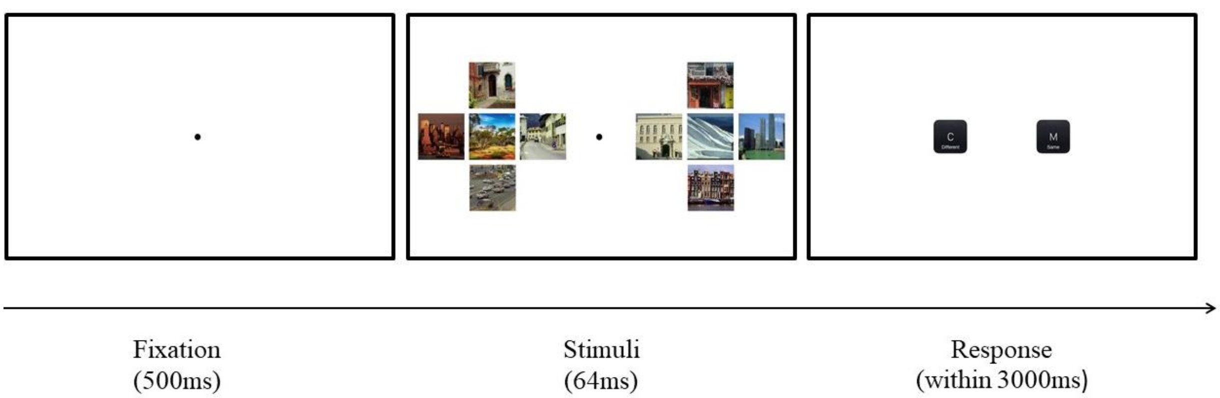



*Figure 9.* Experiment 2 trial sequence with an example of the stimulus configuration for the crowded condition. The targets are the centre images on each side and the surrounding images are the flankers.


*Figure 10.* Experiment 2 stimulus configurations. The image set on the top row are paintings. In the no flanker condition, the target images are both Impressionist paintings. For the crowded condition with incongruent flankers, the two targets are Impressionist paintings, the flankers on the left are all Cubist paintings and the flankers on the right are all Impressionist paintings. For the congruent condition, the two targets are Impressionist paintings and the flankers on each side are all Impressionist paintings. The image set on the bottom row are scenes. In the no flanker condition, the target on the left is a natural scene and on the right is an urban scene. For the incongruent flankers, the target on the left is an urban scene that is flanked with urban scenes, whereas the target on the right is a natural scene flanked with urban scenes. In the congruent condition, the target on the left is an urban scene flanked with urban scenes and the target on the right is a natural scene flanked with natural scenes. The flankers are either congruent or incongruent with the correct response.

### 3.2 Results

Prior to analysis, participants data was screened to ensure that it did not meet any of the exclusion criteria outlined in pre-registration. Univariate density plots and boxplots were visually inspected to assess the distributional nature of the data and identify any extreme outliers across conditions (Appendix H). The distributions appeared approximately normal across conditions, except the crowded condition appeared negatively skewed. All outliers were retained as no participants showed patterns of responding that met the prespecified criteria for exclusion. The preregistered parametric analyses were performed as planned, with an additional set of non-parametric analyses performed to check the robustness of the results to different distributional assumptions. The pattern of results was identical across tests, therefore, the primary preregistered analyses are reported in the main text. The non-parametric analyses are included in Appendix K. The R code and output for the analyses are featured in Appendices H-J.

**3.2.1 Accuracy**. The proportion of correct responses were calculated for each participant in each of the conditions, collapsing across the two image sets. Participants' mean proportion correct was higher when the images were uncrowded (*M* = 0.68, *SD* = 0.08) compared to when they were crowded (*M* = 0.63, *SD* = 0.10). Proportion correct was also higher when the flankers were congruent with the target (*M* = 0.68, *SD* = 0.08), compared to when they were incongruent (*M*= 0.63, *SD* = 0.10). When the flankers were incongruent with the target, as in classic visual crowding experiments, the difference in participants' mean proportion correct for crowded (*M* = 0.58, *SD* = 0.10) and uncrowded (*M* = 0.67, *SD* = 0.08) conditions was substantial. When the flanker were congruent with the target, on the other hand, the difference in participants' mean proportion correct for crowded (*M* = 0.68, *SD* = 0.07) and uncrowded (*M*= 0.69, *SD* = 0.08) conditions was minimal. See the boxplots in Figure 11.

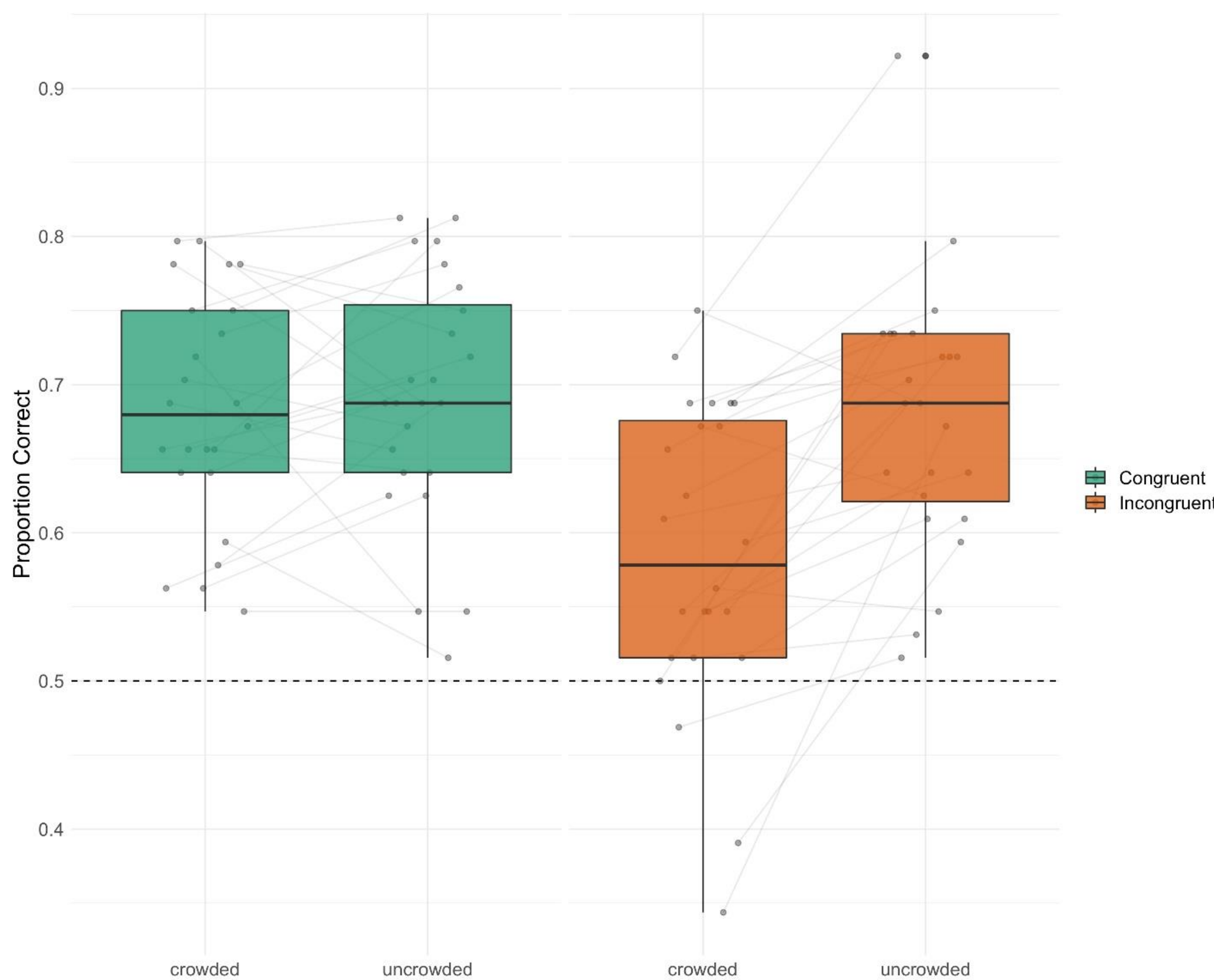



*Figure 11.* Proportion correct boxplots showing the proportion correct for the uncrowded and crowded conditions with congruent and incongruent flankers. Chance performance is indicated by the dotted line (.5).

**3.2.2 Sensitivity**. As in Experiment 1, participants' sensitivity or discriminability was computed based on their hits and false alarms. A hit in this experiment refers to correctly identifying two target images from the same category as the "same" (i.e., 'signal') and a false alarm refers to incorrectly identifying two target images from the same category as "different". Each participants' discriminability was computed using the same nonparametric model of sensitivity used in Experiment 1 (*A*). The boxplots in Figure 12 offer a visual representation of participants discriminability for each condition, collapsing over image sets.

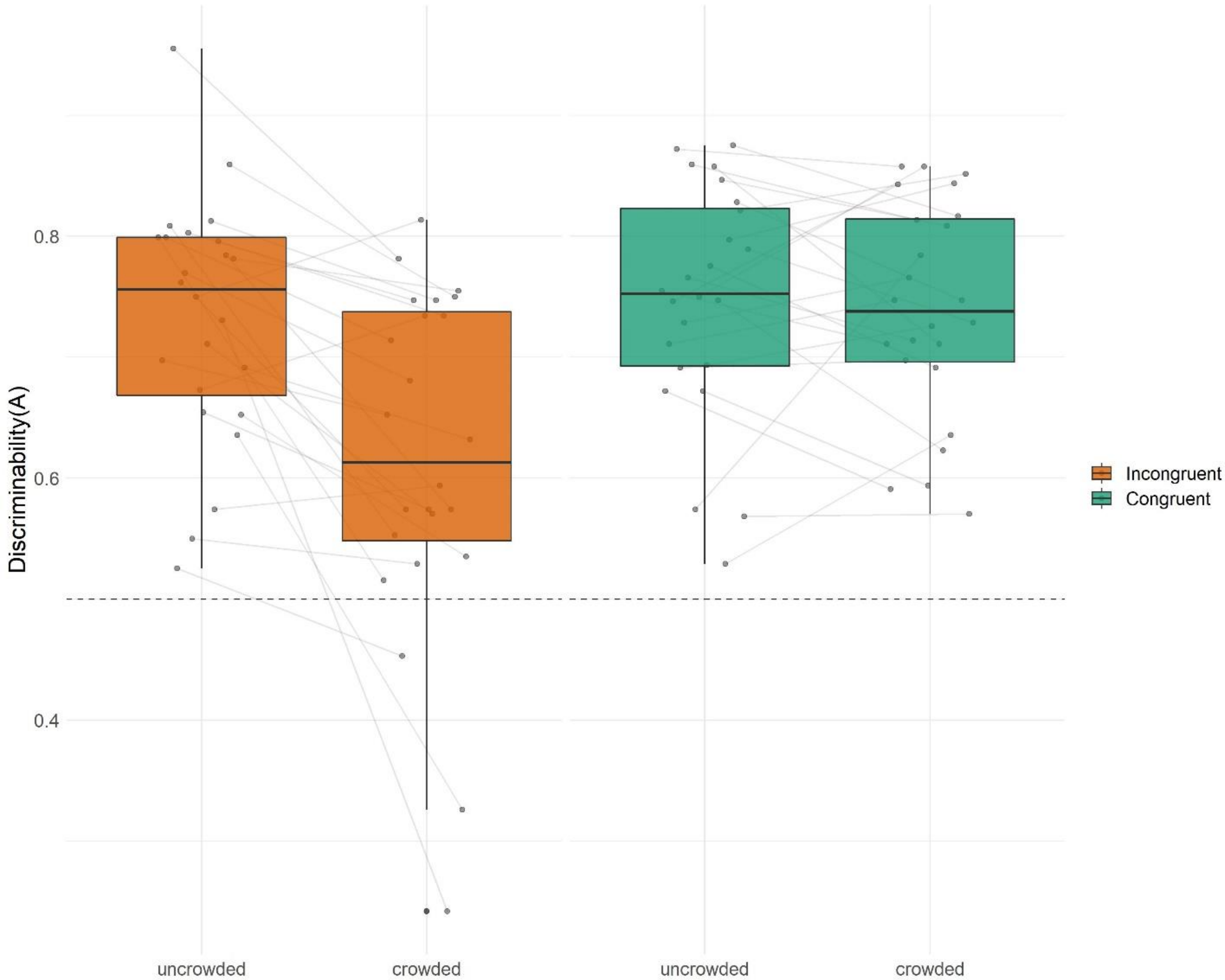



*Figure 12.* Discriminability Boxplots showing discriminability performance for the uncrowded and crowded conditions with congruent and incongruent flankers. Chance performance is indicated by the dotted line (.5).

**3.2.3 Comparison to chance.** A paired samples t-test was used to compare participants mean discriminability to that of simulated participants across each of the conditions. The simulations were programmed to respond randomly and they reflect chance performance (.5) with random variability in responding. There was a statistically significant difference between the *A* scores, $t(95) = -13.97$, $p < 0.001$, 95% CI [-0.24, -0.18], $d = -1.94$, thus supporting prediction 1, that participants can discriminate paintings and scenes in their peripheral vision better than chance.

**3.2.4 Two-way factorial ANOVA.** As planned, participants *A* scores were submitted to a 2×2 repeated measures ANOVA, with crowding (crowded, uncrowded) and congruency (congruent, incongruent) as within-subjects factors. There was a significant main effect of crowding, $F(1, 23) = 22.94$, $p < 0.001$, $\eta_G^2 = 0.07$. Consistent with prediction 3, participants showed better discriminability of the target images when they were uncrowded ($M = 0.74$ , $SD = 0.09$) compared to when they were crowded ($M = 0.67$, $SD = 0.13$). A significant main effect was also found for congruency, $F(1,23) = 14.08$, $p < 0.001$, $\eta_G^2 = 0.09$, such that participants discriminability was higher when the flankers were congruent with the target images ($M = 0.74$, $SD = 0.09$), compared to when they were incongruent with the target images ($M = 0.67$, $SD = 0.09$), thus supporting prediction 2. Importantly, there was a significant interaction between crowding and congruency, $F(1,23) = 9.57$ , $p < 0.001$, $\eta_G^2 = 0.06$. This interaction supports prediction 4 as greater crowding effects occurred when the flankers were incongruent with the correct response (e.g., flankers from mismatching categories but targets from matching categories or vice versa) and this resulted in decreased discriminability ($M = 0.61$ , $SD = 0.14$) compared with when the flankers were congruent with the correct response (e.g., flankers from matching categories and targets from matching categories; $M = 0.73$, $SD = 0.08$).

Post hoc paired samples t-test with Holm correction were conducted to determine whether there was a significant difference between participants *A* scores for the crowded condition with congruent flankers and the uncrowded control condition, $t(23) = -0.48$, $p = 0.63$, 95% CI [-0.04, 0.02], $d = -0.08$. This analysis showed that there was no significant difference in discrimination performance when the target that is surrounded by congruent flankers compared to when it is uncrowded. A post hoc samples t-test was also conducted with the *A* scores under crowded versus uncrowded conditions with incongruent flankers, $t(23) = -4.53$, p = <0.001 , 95%

CI[-0.16, -0.06], $d = -0.94$, which showed a significant difference between discrimination performance.

# CHAPTER 4: General Discussion

The current project investigated the extent to which visual crowding can be reduced and recognition enhanced, when the surrounding clutter is congruent with the gist of the target category. Participants classified images of scenes and paintings presented to the left or right of a fixed point (Experiment 1) and they also discriminated images of paintings and scenes presented to the left and right of a fixed point (Experiment 2). Some of these images were surrounded or crowded by images that were congruent (or incongruent) with the correct response, and others were presented alone. The guiding hypothesis was that target discriminability would be enhanced when the flankers are from the same category as the target. From these two experiments it was shown that congruent flankers can markedly reduce, and even eliminate visual crowding.

The presence of flanking images significantly reduced discriminability, suggesting that a crowding effect impaired participants' ability to correctly discriminate or categorise the target images. Despite this crowding effect, participants could still identify the gist of the target category even when the images were presented for 34 milliseconds with an image resolution of 32×32 pixels. This is consistent with a previous finding by Gong et al. (2018), who demonstrated that the gist of natural scenes can be extracted under conditions of crowding. However, the current study extends this to also include images of paintings and a much lower image resolution. Second, the presence of flanking images that were incongruent with the target reduced participants' ability to correctly classify them, suggesting that the incongruent flankers were producing strong crowding effects. Third, it was found that visual crowding was significantly reduced when the flankers were congruent with the category of the target images, compared to when they were incongruent. In Experiment 1, congruency significantly diluted visual crowding, however in Experiment 2, congruency released the target images from visual crowding altogether as no crowding was observed on congruent trials.

**4.1 Target-flanker relations**

The finding that congruent flankers decrease the effects of crowding seem to be at odds with prior research that has assessed the effects of target-flanker similarity on recognition. The general conclusion drawn from these previous studies is that crowding effects are stronger when the targets and flankers exhibit higher similarity. Similarity is often defined by dichotomising some physical characteristic of the target and flankers such as orientation (Greenwood et al., 2010), colour (Põder, 2007), or shape (Nazir, 1992), and it has been proposed that crowding is reduced when the features of the target differ from that of the flankers. Most of the studies that have tested the influence of target-flanker similarity have used simple stimuli with known features, as it allows researchers to disentangle the contributions of various processes that influence recognition at a fine grain level (Agaoglu & Chung, 2016). This thesis, by contrast, uses natural images with multidimensional features as it offers an opportunity to study visual crowding with stimuli that more closely resemble visual displays encountered by people every day.

Although this study did not directly manipulate or measure similarity, images of the same category do share a certain natural similarity. For example, the contents of specific paintings may change, but the global characteristics of an Impressionist are distributed across the entire category and remain detectable. Congruency was defined by such category membership in the present experiments, and can be thought of as a natural manipulation of similarity at a coarse grained level. Previous evidence suggests that people rely on this information to distinguish visual categories (Oliva, 2005). In Experiment 2, participants were not provided with the category labels for the images that they were comparing, yet they were able to discriminate between low resolution images that belonged to the same or different category with remarkable accuracy. Furthermore, the effect of visual crowding disappeared when surrounding images were congruent with the gist of the target image. This result suggests that the structure distributed across images of the same category may facilitate recognition in peripheral vision.

While the pattern of results are consistent with accounts of gist perception, the findings do conflict with some previous research on visual crowding. For instance, it has been found that flanking stimuli produce stronger crowding effects when they are from the same category as the target (Huckauf et al., 1999; Reuther & Chakravarthi, 2014). Indeed, these conclusions have been drawn from only a few studies that have used alphanumeric stimuli to assess the impact of category membership on recognition. It has been suggested that similarity at the category level is not simply due to low-level feature similarity, but also influenced by higher-level processing of meaning (Reuther & Chakravarthi, 2014).

In experiment 1 of this study, participants were provided with the semantic category labels for each of the categories. In order to be able to classify the images, they had to have an adequate understanding of the meaning associated with each visual category. In the case of congruent flankers, all of the images belong to the same category, sharing the same meaning. Congruency doesn't only refer to the perceptual information but also the conceptual meaning that is associated with the visual category. It could therefore be possible that high-level cognitive processes such as concept formation could also modulate the strength of crowding, as described in gist perception research (Oliva, 2005). Flankers that share similar category membership (i.e., semantic meaning) to the target could perhaps facilitate recognition, despite the requisite conditions for visual crowding. When images belong to a different category to the target, on the other hand, there is conflicting semantic meaning from the two different categories and this may create confusion as to which category the target belongs to when perceiving the gist of the whole display in peripheral vision. Experiment 2 provides some evidence that shared semantic meaning is perhaps not the main driver of the crowding × congruency interaction, because people showed reduced visual crowding with congruent flankers, despite needing to know the categories they were discriminating.

**4.2 Pooling**

Another possible explanation for these results is given by pooling models. Pooling is considered a regularisation process of the visual system that creates consistency in appearance among items of a stimulus array. Instead of representing each item individually, information is compressed to form an average or statistical summary of the array (Balas et al., 2009). Gist perception shares similarities with descriptions of this pooling process in visual crowding in that it is thought to be a form of texture recognition (Renninger & Malik, 2004). When viewing a scene, the summary statistics or texture information can provide enough description for an observer to identify its category (Renninger & Malik, 2004).

The results of this study show that flankers which belong to the same category as the target, produce less crowding. If pooling was occurring, information from the target and flankers would be pooled together to form a texture and because the congruent flankers already share common statistical regularities with the target image, the resulting percept may be more representative of the categorical identity of the image. Perhaps these shared statistical regularities combine together more cohesively when pooled, resulting in an average or more accurate perceptual representation of the target image. It therefore might be the case that when information is pooled between target and flanking images that are incongruent, the resulting texture becomes a mix of the two different categories, making it more difficult for the participants to determine the identity of the target.

Pooling can have a detrimental effect on recognition for things such as letters (Freeman et al., 2012), as the ability to recognise them is reliant on an observer being able to discern its unique boundaries and curves. When letters are flanked by random other letters, the distinctiveness of their boundaries becomes blurred through the formation of perceptual texture, and a textural representation does not suffice for object recognition with these type of stimuli. The paintings and scenes used in this study were all presented in a square aperture meaning that there was consistency in their borders and boundaries. Scene and painting recognition isn't as dependent on the observer being able to discern its boundaries, but rather what resides within those boundaries. Regardless of their categorical affiliation, all of the images used in this study shared a consistent border, and this may have allowed the visual system to pool or summarise the statistical information that resides within those borders in a more seamless way. Interestingly, when a target letter is surround with more congruent flanking letters, it is easier to distinguish. For example, as it was shown earlier, flanking the letter 'B' with different instances of the same letter appears to create less crowding than if you were to surround it with incongruent letters. This may be due to the consistency in its borders and the way that it occupies space. Perhaps it may be that the structure and consistency of congruent surroundings better survives the disruption from pooling processes.

**4.3 Substitution**

Another possible explanation for the crowding effects observed across the two experiments is substitution, whereby the flankers are reported instead of the targets. On the crowded trials with congruent flankers in experiment 1, target images were flanked by four other images, all belonging to the same category as the target. Although participants showed better discriminability for crowded target images with congruent flankers, it is difficult to determine if a correct classification is the result of substitution, as opposed to pooling, because the flanker and target share the same categorical identity, and reporting either of them would result in a correct classification.

A better indication of substitution can be gleaned by looking at the responses that were made in the incongruent flanker conditions, because substitution errors and correct classifications produce different responses. On visual inspection of the data for Experiment 1, note that the average discriminability for seven of the participants for the crowded condition with incongruent flankers was far below chance level. That is, they were possibly confusing the categories consistently across trials. Although a qualitative observation, these outlying observations were isolated to the incongruent crowding condition and possibly reflect evidence for substitution. Furthermore, the discriminability for these same observers was above chance level in the uncrowded control condition, indicating that their performance in the crowded condition are not likely to be explained by participants' failure to understand the categories.

Similarly in Experiment 2, recognition was enhanced when the flankers were congruent with the target pair, and this increased performance could also be influenced by substitution errors. The task in Experiment 2 required participants to judge whether two images were from the same category or a different category. When the flankers were congruent with the correct response of "same," all of the images on the screen were from the same category. Therefore, if the participants mistakenly judged a flanker for the target, then this would result in a correct response. This is also the case when the flankers were congruent with the correct response of "different." For example. a target natural scene that is presented on the left side would be surrounded with other natural scenes, and a target urban scene that is presented on the right side would be surrounded by urban scenes. If people were unable to see the target images on both sides and they were substituting (or pooling) the targets with flankers, then the responses that they would be giving would be correct, as the flankers on the right are also different to the flankers on the left. When assessing the boxplots for this experiment, it was again found that performance for four of the participants was below chance level for the incongruent flanker condition, demonstrating that they were incorrectly discriminating the images. However, performance was above chance for the uncrowded condition and this pattern of results demonstrates that their decreased performance in the crowded condition is not likely to be explained by a failure to recognise the differences between the stimuli.

Research on visual crowding has shown evidence for substitution and pooling, but no one model has been shown to account for all of the perceptual errors that are made under conditions of visual crowding (Harrison & Bex, 2017). Although these different accounts of visual crowding offer possible explanations for the observed crowding effects, the experiments reported here don't provide evidence that would discriminate among them and it is possible that different individuals were relying on multiple crowding mechanisms to different degrees across the two tasks. The main goal of this thesis was to test the effect of congruency on visual crowding but future experiments might look to using experimental designs which allow pooling and substitution to be distinguished.

**4.4 Strengths**

This thesis uses two experimental tasks and two sets of stimuli to test a single hypothesis and its corresponding predictions. The rationale for using two different experiments was to test the robustness of the findings and ensure that any effects that were observed weren't due to the specific task, as it has previously been suggested that the difficulty of a task can modulate crowding (Kalpadakis-Smith, Goffaux & Greenwood, 2018). A strength of these experiments is the use of a within-subjects design, as it removes the possibility that the pattern of results were due to uncontrolled individual differences among participants that may potentially confound the results. Within-subjects analyses also offer increased statistical power for detecting effects and reduce the number of participants and trials required to detect an effect. Counterbalancing and randomisation procedures were used at each step to control for any order effects that are associated with a within-subjects design.

The sample size of 24 in this study was also much larger than those typically used in visual crowding studies (e.g., Farzin, et al., 2009; Greenwood et al., 2009; Kimchi & Pirkner, 2015; Poder, 2007; Saarela et al., 2009; Vandenberg et al., 2007), which range between 3 and 12. Furthermore, many of these studies include the authors as participants and this may implicitly influence responding (Rosenthal & Fode, 1963). All participants in the current experiments were naïve to the study goals, predictions, materials and conditions, as familiarity with these elements could have influenced the results (Rosenthal & Fode, 1963).

Another strength was having a large pool of images for each stimulus set as it allowed for unique, randomised trial sequences to be generated for each participant. The choice to use images of paintings and scenes was also a strength of the study as it allowed visual crowding to be assessed using naturally varying stimuli. Many of the studies that have previously assessed visual crowding have used artificial stimuli (e.g. orientation bars, gratings and Gabor patches) and it has been proposed that natural stimuli may engage different neural mechanisms compared to that of artificial stimuli (Simoncelli & Olshausen, 2001; Felsen, & Dan, 2005).

The findings reported in this thesis offer new insights into the role that visual context may play on recognition. The two experiments extend on the large body of previous visual crowding research by providing a novel example where the bottleneck of object recognition may be lifted. As models of crowding are advanced, they may look to account for the congruency of the visual surroundings.

**4.5 Limitations**

In Experiment 2, the participants did not need know the categories that they were classifying in order to successfully distinguish them. A qualification on this point, however, is that half of the participants had already undertaken Experiment 1, where they were provided with the categories and had learnt how to distinguish them. Furthermore, all of the participants watched the instructional video which gave examples of the image categories that would be presented for discrimination. As such, it is also possible that participants implicitly (via the instructions) or explicitly (by completing Experiment 1 first) learned the semantic labels associated with each category.

In both experiments, the participants were required to either report the category that the target stimulus belonged to, or discriminate whether two images were from the same or different category. Given that the participants were not required to report the identity of the stimuli at a more basic level, it is difficult to tease apart the processes that may be facilitating or hampering recognition. If performance is impacted by substitution, then the only way that this could be confirmed is to have the identity of the target image reported. This idea of congruency could be tested in the future using images such as natural scenes as flankers, but the reporting requirements could be at the basic level (e.g., mountain, beach, forest).

**4.6 Conclusions**

The present results show that the strength of crowding is influenced by the information that is shared with the target. The gist information that comprises a particular painting or scene category is better detected when the surroundings are congruent with the target. The presence of congruent flankers significantly reduced (Experiment 1) and even fully released (Experiment 2) the paintings and scene images from visual crowding. This finding sets the groundwork for future research to explore whether and to what extent a more congruent visual context may facilitate recognition with other forms of stimuli. Importantly, it has been shown that perception of high-level object information is enhanced when the surroundings are match the target at hand. Visual crowding is widely thought to place a fundamental limit on perception, but the two experiments presented here demonstrate that visual context may not always be a detriment to visual recognition in peripheral vision.

## Contributions

I was involved in the conceptual design of this study in collaboration with my supervisor. I obtained ethics approval from the University of Adelaide prior to commencing data collection and contributed ideas for refining the scope, aims and particulars of the project. I independently sourced and reviewed the literature that was pertinent to this thesis. My supervisor prepared the stimuli, programmed the experiment and developed a program for me to extract the data from the raw individual participant data files. I created the written and video instructions for both experiments, registered the study on the Open Science Framework, recruited each participant, conducted all data collection, performed the statistical analyses, and generating all plots and figures in the thesis.

Reference List

Abramov, I., Farkas, A., & Ochsenschlager, E. (2006). A study in classification: style and visual perception. *Visual Anthropology*, *19*(3-4), 255-274.

Agaoglu, M. N., & Chung, S. T. (2016). Can (should) theories of crowding be unified?. *Journal of vision*, *16*(15), 10-10.

Andriessen, J. J., & Bouma, H. (1976). Eccentric vision: Adverse interactions between line segments. *Vision research*, *16*(1), 71-78.

Anstis, S. (1998). Picturing peripheral acuity. *Perception*, *27*(7), 817-825.

Balas, B., Nakano, L., & Rosenholtz, R. (2009). A summary-statistic representation in peripheral vision explains visual crowding. *Journal of vision*, *9*(12), 13-13.

Bernard, J. B., & Chung, S. T. (2011). The dependence of crowding on flanker complexity and target–flanker similarity. *Journal of vision*, *11*(8), 1-1.

Bex, P. J., & Dakin, S. C. (2005). Spatial interference among moving targets. *Vision research*, *45*(11), 1385-1398.

Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature*, *226*(5241), 177.

Bulakowski, P. F., Post, R. B., & Whitney, D. (2009). Visuomotor crowding: the resolution of grasping in cluttered scenes. *Frontiers in Behavioral Neuroscience*, *3*, 49.

Carpenter, R. H. (1988). *Movements of the Eyes, 2nd Rev*. Pion Limited.

Cohen, M. A., Dennett, D. C., & Kanwisher, N. (2016). What is the bandwidth of perceptual experience?. *Trends in cognitive sciences*, *20*(5), 324-335.

Ester, E. F., Klee, D., & Awh, E. (2014). Visual crowding cannot be wholly explained by feature pooling. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(3), 1022.

Farzin, F., Rivera, S. M., & Whitney, D. (2009). Holistic crowding of Mooney faces. *Journal of vision*, *9*(6), 18-18.

Felsen, G., & Dan, Y. (2005). A natural approach to studying vision. *Nature neuroscience*, *8*(12), 1643.

Fischer, J., & Whitney, D. (2011). Object-level visual information gets through the bottleneck of crowding. *Journal of Neurophysiology*, *106*(3), 1389-1398.

Freeman, J., Chakravarthi, R., & Pelli, D. G. (2012). Substitution and pooling in crowding. *Attention, Perception, & Psychophysics*, *74*(2), 379-396.

Gerhard, H. E., & Bethge, M. (2014). Towards rigorous study of artistic style: a new psychophysical paradigm. *Art & Perception*, *2*(1-2), 23-44.

Gerhard, H. E., Wichmann, F. A., & Bethge, M. (2013). How sensitive is the human visual system to the local statistics of natural images?. *PLoS computational biology*, *9*(1), e1002873.

Gong, M., Xuan, Y., Smart, L. J., & Olzak, L. A. (2018). The extraction of natural scene gist in visual crowding. *Scientific reports*, *8*(1), 14073.

Greene, M. R., & Oliva, A. (2009). The briefest of glances: The time course of natural scene understanding. *Psychological Science*, *20*(4), 464-472.

Greenwood, J. A., Bex, P. J., & Dakin, S. C. (2010). Crowding changes appearance. *Current Biology*, *20*(6), 496-501.

Harrison, W. J., & Bex, P. J. (2017). Visual crowding is a combination of an increase of positional uncertainty, source confusion, and featural averaging. *Scientific reports*, *7*, 45551.

Huckauf, A., & Heller, D. (2002). What various kinds of errors tell us about lateral masking effects. Visual Cognition, 9(7), 889-910.

Huckauf, A., Heller, D., & Nazir, T. A. (1999). Lateral masking: Limitations of the feature interaction account. *Perception & Psychophysics, 61*(1), 177-189.

Kalpadakis-Smith, A. V., Goffaux, V., & Greenwood, J. A. (2018). Crowding for faces is determined by visual (not holistic) similarity: Evidence from judgements of eye position. *Scientific reports, 8*(1), 12556.

Keshvari, S., & Rosenholtz, R. (2016). Pooling of continuous features provides a unifying account of crowding. *Journal of Vision*, *16*(3), 39-39.

Kimchi, R., & Pirkner, Y. (2015). Multiple level crowding: Crowding at the object parts level and at the object configural level. *Perception*, *44*(11), 1275-1292.

Kooi, F. L., Toet, A., Tripathy, S. P., & Levi, D. (1994). The effect of similarity and duration on spatial interaction in peripheral vision. *Spatial Vision*, *8*, 255–279.

Korte, W. (1923). Über die Gestaltauffassung im indirekten Sehen. *Zeitschrift für Psychologie*, *93*, 17-82.

Kouider, S., Berthet, V., & Faivre, N. (2011). Preference is biased by crowded facial expressions. *Psychological Science*, *22*(2), 184-189.

Lakens, D., Scheel, A. M., & Isager, P. M. (2018). Equivalence testing for psychological research: A tutorial. *Advances in Methods and Practices in Psychological Science*, *1*(2), 259-269.

Larson, A. M., & Loschky, L. C. (2009). The contributions of central versus peripheral vision to scene gist recognition. *Journal of Vision*, *9*(10), 6-6.

Levi, D. M. (2011). Visual crowding. *Current Biology*, *21*(18), R678-R679.

Levi, D. M., Hariharan, S., & Klein, S. A. (2002). Suppressive and facilitatory spatial interactions in peripheral vision: Peripheral crowding is neither size invariant nor simple contrast masking. *Journal of vision*, *2*(2), 3-3.

Livne, T., & Sagi, D. (2007). Configuration influence on crowding. *Journal of vision*, *7*(2), 4-4.

Louie, E. G., Bressler, D. W., & Whitney, D. (2007). Holistic crowding: Selective interference between configural representations of faces in crowded scenes. *Journal of Vision*, *7*(2), 24-24.

Millin, R., Arman, A. C., Chung, S. T., & Tjan, B. S. (2013). Visual crowding in V1. *Cerebral Cortex*, *24*(12), 3107-3115.

Munafò, M. R., Nosek, B. A., Bishop, D. V., Button, K. S., Chambers, C. D., Du Sert, N. P., ... & Ioannidis, J. P. (2017). A manifesto for reproducible science. *Nature human behaviour*, *1*(1), 0021.

Nazir, T. A. (1992). Effects of lateral masking and spatial precueing on gap-resolution in central and peripheral vision. *Vision research*, *32*(4), 771-777.

Oliva, A. (2005). Gist of the scene. In *Neurobiology of attention* (pp. 251-256). Academic Press.

Oliva, A., & Torralba, A. (2001). Modeling the shape of the scene: A holistic representation of the spatial envelope. *International journal of computer vision*, *42*(3), 145-175.

Parkes, L., Lund, J., Angelucci, A., Solomon, J. A., & Morgan, M. (2001). Compulsory averaging of crowded orientation signals in human vision. *Nature neuroscience*, *4*(7), 739.

Pelli, D. G., & Tillman, K. A. (2008). The uncrowded window of object recognition. *Nature neuroscience*, *11*(10), 1129.

Pelli, D. G., Palomares, M., & Majaj, N. J. (2004). Crowding is unlike ordinary masking: Distinguishing feature integration from detection. *Journal of vision*, *4*(12), 12-12.

Peng, C., Zhang, Y., Chen, Y., & Zhang, M. (2013). Crowded words can be processed semantically: Evidence from an ERP study. *International Journal of Psychophysiology*, *88*(1), 91-95.

Põder, E. (2007). Effect of colour pop-out on the recognition of letters in crowding conditions. *Psychological Research*, *71*(6), 641-645.

Poder, E., & Wagemans, J. (2007). Crowding with conjunctions of simple features. *Journal of Vision*, *7*(2), 23-23.

Renninger, L. W., & Malik, J. (2004). When is scene identification just texture recognition?. *Vision research*, *44*(19), 2301-2311.

Reuther, J., & Chakravarthi, R. (2014). Categorical membership modulates crowding: Evidence from characters. *Journal of Vision*, *14*(6), 5-5.

Rosenholtz, R. (2016). Capabilities and limitations of peripheral vision. *Annual Review of Vision Science*, *2*, 437-457.

Rosenthal, R., & Fode, K. L. (1963). The effect of experimenter bias on the performance of the albino rat. *Behavioral Science*, *8*(3), 183-189.

Saarela, T. P., Sayim, B., Westheimer, G., & Herzog, M. H. (2009). Global stimulus configuration modulates crowding. *Journal of Vision*, *9*(2), 5-5.

Searston, R. A., Thompson, M. B., Vokey, J. R., French, L. A., & Tangen, J. M. (2019). How low can you go? Detecting style in extremely low resolution images. *Journal of Experimental Psychology: Human Perception and Performance*, *45*(5), 573.

Simoncelli, E. P., & Olshausen, B. A. (2001). Natural image statistics and neural representation. *Annual review of neuroscience*, *24*(1), 1193-1216.

Strasburger, H. (2005). Unfocussed spatial attention underlies the crowding effect in indirect form vision. *Journal of Vision*, *5*(11), 8-8.

Strasburger, H., & Malania, M. (2013). Source confusion is a major cause of crowding. *Journal of Vision*, *13*(1), 24-24.

Van den Berg, R., Roerdink, J. B., & Cornelissen, F. W. (2007). On the generality of crowding: Visual crowding in size, saturation, and hue compared to orientation. *Journal of Vision*, *7*(2), 14-14.

Vlaskamp, B. N., & Hooge, I. T. C. (2006). Crowding degrades saccadic search performance. *Vision research*, *46*(3), 417-425.

Wallace, J. M., & Tjan, B. S. (2011). Object crowding. *Journal of vision*, *11*(6), 19-19.

Whitney, D., & Levi, D. M. (2011). Visual crowding: A fundamental limit on conscious perception and object recognition. *Trends in cognitive sciences*, *15*(4), 160-168.

Yeotikar, N. S., Khuu, S. K., Asper, L. J., & Suttle, C. M. (2011). Configuration specificity of crowding in peripheral vision. *Vision Research*, *51*(11), 1239-1248.

Yu, D., Akau, M. M., & Chung, S. T. (2012). The mechanism of word crowding. *Vision research*, *52*(1), 61-69.

Appendix A - Poster

# Appendix B – Consent Form

THE UNIVERSITY of ADELAIDE

Human Research Ethics Committee (HREC)

## CONSENT FORM

1. I have read the attached Information Sheet and agree to take part in the following research project:

| | |
|---|---|
| **Title:** | What's your style? Discriminating artistic and naturalistic style in visual crowding. |
| **Ethics Approval Number:** | [redacted] |

2. I have had the project, so far as it affects me, and the potential risks and burdens fully explained to my satisfaction by the research worker. I have had the opportunity to ask any questions I may have about the project and my participation. My consent is given freely.
3. I agree to participate in the activities outlined in the participant information sheet.
4. I understand that as my participation is anonymous, I can withdraw any time up until submission of my data for analysis.
5. I have been informed that the information gained in the project may be published in a thesis, conference presentation, website, report, book or journal.
6. I have been informed that in the published materials I will not be identified and my personal results will not be divulged.
7. I consent to any data gathered from this participation to be used for future research purposes and to the data being stored in an online public repository (e.g., Open Science Framework).
8. I understand my information will only be disclosed according to the consent provided, except where disclosure is required by law.
9. I am aware that I should keep a copy of this Consent Form, when completed, and the attached Information Sheet.

**Participant to complete:**

Name: ______________________ Signature: ________________________ Date: __________

**Researcher/Witness to complete:**

I have described the nature of the research to ________________________________________
*(print name of participant)*

and in my opinion she/he understood the explanation.

Signature: ___________________ Position: _________________________ Date: __________

Appendix C – Participant Information Sheet

## PARTICIPANT INFORMATION SHEET

THE UNIVERSITY of ADELAIDE

**PROJECT TITLE:** What's your style? Discriminating artistic and naturalistic style in visual crowding.
**HUMAN RESEARCH ETHICS COMMITTEE APPROVAL NUMBER:** [redacted]
**PRINCIPAL INVESTIGATOR:** Dr #####
**STUDENT RESEARCHER:** #####
**STUDENT'S DEGREE:** Psychology, Honours

Dear Participant,

We invite you to join us in The Expert Cognition Lab within the University of Adelaide's School of Psychology, to participate in a study on human perception and cognition.

**What is the project about?**
Central vision is what allows us to see things clearly; to find a familiar face in a crowd of faces, or locate the red jelly beans in a jar of multicoloured jelly beans. As you shift your gaze to each word of this sentence, they enter your central vision. Notice that as you focus on each word, you also have an awareness that there are other words in the sentence, however, you can't identify what they are until you look at them directly. This is because these other words fall into your peripheral vision. Peripheral vision makes up most of our visual field and although what we see here is relatively impoverished, we can still extract large amounts of useful information from it.
The aim of this research is to gain an understanding about how we make sense of what our eyes see when information is presented within the peripheral visual field. We already know that peripheral vision has many limitations on our ability to recognise objects, especially if they are surrounded by other objects. Your participation will help us to understand why this occurs and what type of information can be extracted from peripheral vision.

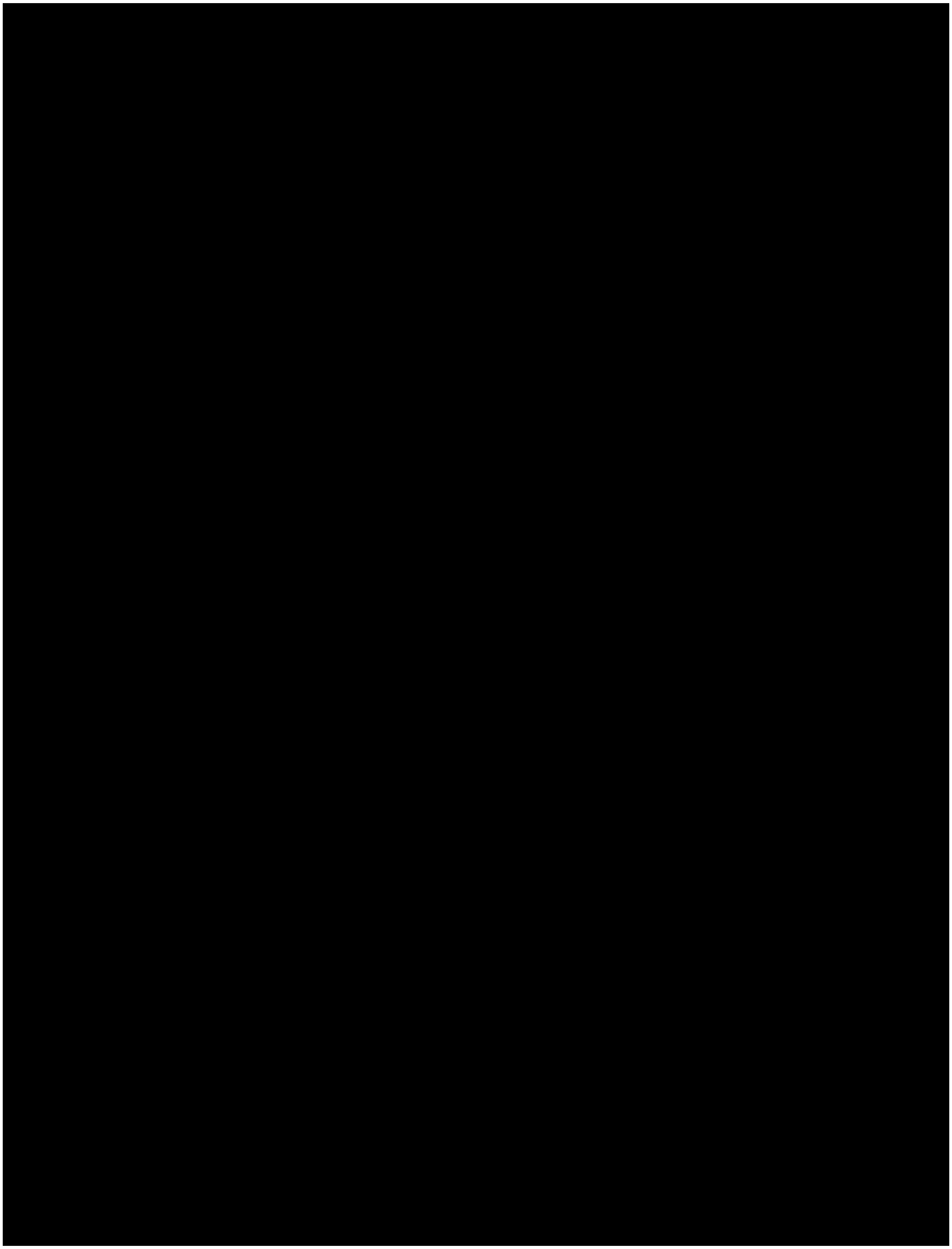
**Who is undertaking the project?**
This project is being conducted by ##### and will form the basis for the Psychology Honours degree at the University of Adelaide under the supervision of #######.

**Am I eligible to participate?**
You are eligible to participate if you:

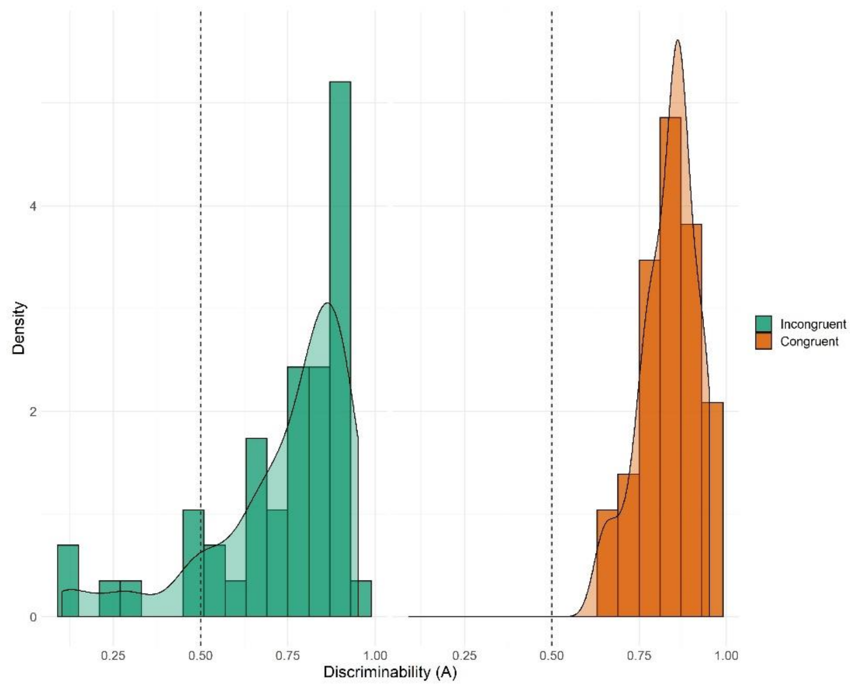
- ❑ Are fluent in English
- ❑ Are 18 years of age and over
- ❑ Have normal or corrected to normal vision – so wear your contact lenses or bring your glasses!
- ❑ Have functional use of your dominant hand (i.e. no broken fingers/arm)

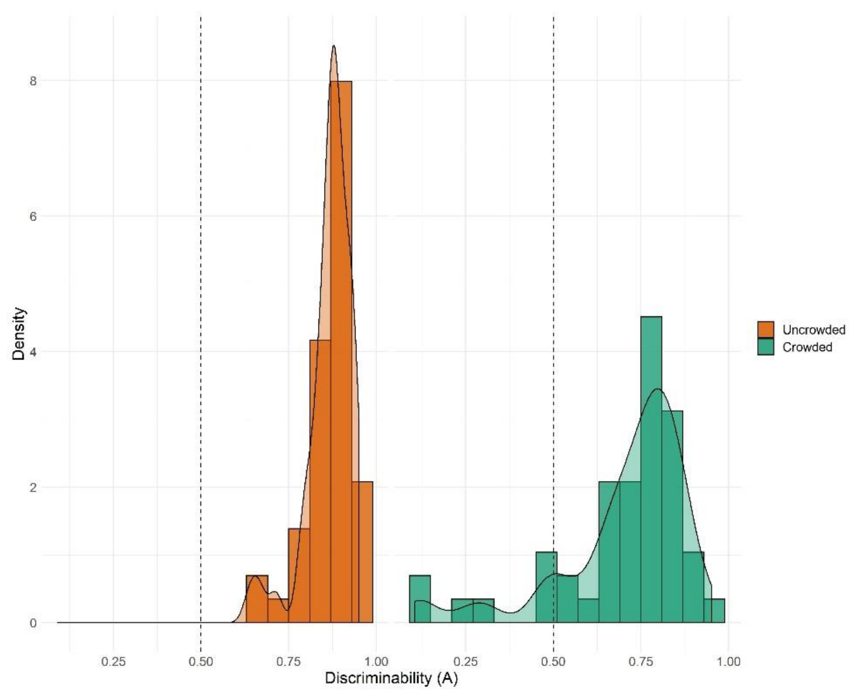
**What am I being invited to do?**
The study will involve you classifying images of paintings and scenes presented in your peripheral vision on a computer screen. No specific prior knowledge is required. After you have completed the experiment, the researcher will discuss the study with you and answer any questions you may have.

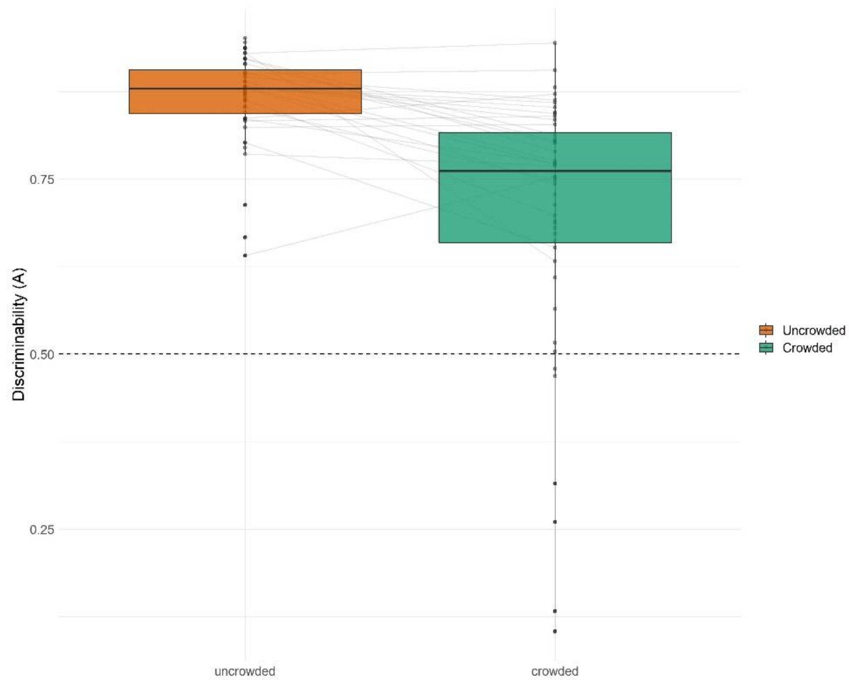
**How much time will my involvement in the project take?**

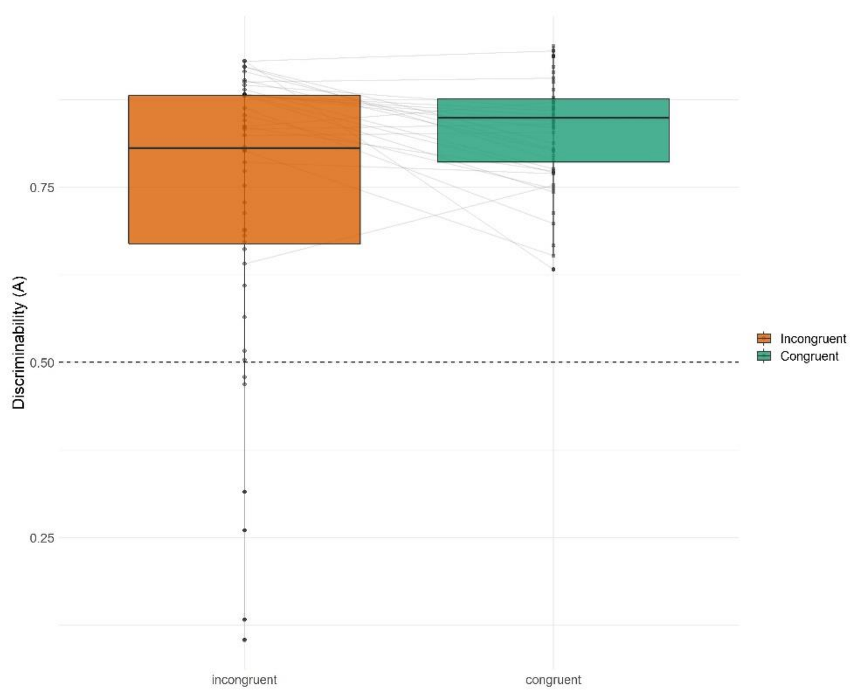


Appendix D – Experiment 1 Density Plots and Boxplots




Discriminability (A)
0.75
0.50
0.25
uncrowded
crowded
Uncrowded
Crowded


Discriminability (A)
0.75
0.50
0.25
incongruent
congruent
Incongruent
Congruent

# Appendix E – Experiment 1 Descriptive Statistics

## Proportion correct – Crowding × congruency conditions

```
my_data_p %>%
  group_by(Crowding, Congruent) %>% summarise(
    mean = mean(mean_PC),
    variance = var(mean_PC),
    SD = sd(mean_PC)
  )
```

| Crowding <fctr> | Congruent <fctr> | mean <dbl> | variance <dbl> | SD <dbl> |
|---|---|---|---|---|
| uncrowded | incongruent | 0.8072917 | 0.004161005 | 0.06450586 |
| uncrowded | congruent | 0.8103299 | 0.005896920 | 0.07679141 |
| crowded | incongruent | 0.5828993 | 0.022932510 | 0.15143484 |
| crowded | congruent | 0.7361111 | 0.005412767 | 0.07357151 |

4 rows

## Proportion correct – Crowding conditions

```
my_data_p %>%
  group_by(Crowding) %>% summarise(
    mean = mean(mean_PC),
    variance = var(mean_PC),
    SD = sd(mean_PC)
  )
```

| Crowding <fctr> | mean <dbl> | variance <dbl> | SD <dbl> |
|---|---|---|---|
| uncrowded | 0.8088108 | 0.00492432 | 0.0701735 |
| crowded | 0.6595052 | 0.01986442 | 0.1409412 |

2 rows

## Proportion correct – Congruency conditions

```
my_data_p %>%
  group_by(Congruent) %>% summarise(
    mean = mean(mean_PC),
    variance = var(mean_PC),
    SD = sd(mean_PC)
  )
```

| Congruent <fctr> | mean <dbl> | variance <dbl> | SD <dbl> |
|---|---|---|---|
| incongruent | 0.6950955 | 0.026114341 | 0.16159932 |
| congruent | 0.7732205 | 0.006940933 | 0.08331226 |

2 rows

## *A* scores – Crowding × congruency conditions

```
my_data_p %>%
  group_by(Crowding, Congruent) %>% summarise(
    mean = mean(A),
    variance = var(A),
    SD = sd(A)
  )
```

| Crowding <fctr> | Congruent <fctr> | mean <dbl> | variance <dbl> | SD <dbl> |
|---|---|---:|---:|---:|
| uncrowded | incongruent | 0.8653791 | 0.003845635 | 0.06201319 |
| uncrowded | congruent | 0.8671152 | 0.004779624 | 0.06913482 |
| crowded | incongruent | 0.6022497 | 0.047325997 | 0.21754539 |
| crowded | congruent | 0.7986654 | 0.005416998 | 0.07360026 |

4 rows

## *A* scores – Congruency conditions

```
my_data_p %>%
  group_by(Congruent) %>% summarise(
    mean = mean(A),
    variance = var(A),
    SD = sd(A)
  )
```

| Congruent <fctr> | mean <dbl> | variance <dbl> | SD <dbl> |
|---|---:|---:|---:|
| incongruent | 0.7338144 | 0.042718982 | 0.20668571 |
| congruent | 0.8328903 | 0.006186102 | 0.07865178 |

2 rows

## *A* scores – Crowding conditions

```
my_data_p %>%
  group_by(Crowding) %>% summarise(
    mean = mean(A),
    variance = var(A),
    SD = sd(A)
  )
```

| Crowding <fctr> | mean <dbl> | variance <dbl> | SD <dbl> |
|---|---:|---:|---:|
| uncrowded | 0.8662471 | 0.004221641 | 0.06497416 |
| crowded | 0.7004575 | 0.035660387 | 0.18883958 |

2 rows

# Appendix F – Experiment 1 Statistical Analyses

## Comparison to chance t-test

```
t.test(
  my_data$A[my_data$Mode == "Sim"],
  my_data$A[my_data$Mode == "Participant"],
  paired = TRUE,
  alternative = "two.sided"
)
```

```

	Paired t-test

data:  my_data$A[my_data$Mode == "Sim"] and my_data$A[my_data$Mode == "Participant"]
t = -15.709, df = 95, p-value < 2.2e-16
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.3151799 -0.2444537
sample estimates:
mean of the differences
             -0.2798168
```

```
## Cohen's d
sampleMean1<-mean(my_data$A[my_data$Mode == "Sim"], na.rm = TRUE)
sampleMean2<-mean(my_data$A[my_data$Mode == "Participant"], na.rm = TRUE)
sampleSD1<-sd(my_data$A[my_data$Mode == "Sim"], na.rm = TRUE)
sampleSD2<-sd(my_data$A[my_data$Mode == "Participant"], na.rm = TRUE)
sampleSDpooled<-sqrt((sampleSD1*sampleSD1+sampleSD2*sampleSD2)/2)
d<-(sampleMean1-sampleMean2)/sampleSDpooled
d
```

```
[1] -2.185558
```

## Post hoc t-test (uncrowded incongruent, crowded incongruent)

```
t.test(
  incongruent_1$A[incongruent_1$Crowding == "crowded"],
  incongruent_1$A[incongruent_1$Crowding == "uncrowded"],
  paired = TRUE,
  alternative = "two.sided"
)
```

```

	Paired t-test

data:  incongruent_1$A[incongruent_1$Crowding == "crowded"] and incongruent_1$A[incongruent_1$Crowding == "uncrow
ded"]
t = -6.3034, df = 23, p-value = 1.971e-06
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.3494839 -0.1767748
sample estimates:
mean of the differences
             -0.2631293
```

```
## Cohen's d
sampleMean1<-mean(incongruent_1$A[incongruent_1$Crowding == "crowded"], na.rm = TRUE)
sampleMean2<-mean(incongruent_1$A[incongruent_1$Crowding == "uncrowded"], na.rm = TRUE)
sampleSD1<-sd(incongruent_1$A[incongruent_1$Crowding == "crowded"], na.rm = TRUE)
sampleSD2<-sd(incongruent_1$A[incongruent_1$Crowding == "uncrowded"], na.rm = TRUE)
sampleSDpooled<-sqrt((sampleSD1*sampleSD1+sampleSD2*sampleSD2)/2)
d<-(sampleMean1-sampleMean2)/sampleSDpooled
d
```

```
[1] -1.645014
```

Post hoc t-test (uncrowded congruent, crowded congruent)

```
congruent_1 <- my_data_p  %>%
 filter(Congruent == "congruent")

t.test(
  congruent_1$A[congruent_1$Crowding == "crowded"],
  congruent_1$A[congruent_1$Crowding == "uncrowded"],
  paired = TRUE,
  alternative = "two.sided"
)
```

```

    Paired t-test

data:  congruent_1$A[congruent_1$Crowding == "crowded"] and congruent_1$A[congruent_1$Crowding == "uncrowded"]
t = -4.1043, df = 23, p-value = 0.0004341
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.10295024 -0.03394935
sample estimates:
mean of the differences
             -0.0684498
```

```
## Cohen's d
sampleMean1<-mean(congruent_1$A[congruent_1$Crowding == "crowded"], na.rm = TRUE)
sampleMean2<-mean(congruent_1$A[congruent_1$Crowding == "uncrowded"], na.rm = TRUE)
sampleSD1<-sd(congruent_1$A[congruent_1$Crowding == "crowded"], na.rm = TRUE)
sampleSD2<-sd(congruent_1$A[congruent_1$Crowding == "uncrowded"], na.rm = TRUE)
sampleSDpooled<-sqrt((sampleSD1*sampleSD1+sampleSD2*sampleSD2)/2)
d<-(sampleMean1-sampleMean2)/sampleSDpooled
d
```

```
[1] -0.9586476
```

## Holm correction for post hoc t-tests

```
## p values
pvals <- c(0.001, 0.0004341)
## Adjust p values
bonf  <- p.adjust(pvals, "bonferroni")
holm  <- p.adjust(pvals, "holm")
hoch  <- p.adjust(pvals, "hochberg")
hom   <- p.adjust(pvals, "hommel")
BH    <- p.adjust(pvals, "BH")
BY    <- p.adjust(pvals, "BY")
## Print p values
bonf
```

```
[1] 0.0020000 0.0008682
```

Hide

```
holm
```

```
[1] 0.0010000 0.0008682
```

## Two-way ANOVA

```
options(contrasts=c("contr.sum","contr.poly"))
ezANOVA(data=my_data_p, dv=A, wid= Participant, within= .(Crowding, Congruent))
```

```
$ANOVA
```

| | Effect <chr> | DFn <dbl> | DFd <dbl> | F <dbl> | p <dbl> | p<.05 <chr> | ges <dbl> |
|---|---|---|---|---|---|---|---|
| 2 | Crowding | 1 | 23 | 37.21962 | 3.188140e-06 | * | 0.3185052 |
| 3 | Congruent | 1 | 23 | 31.19197 | 1.106321e-05 | * | 0.1430340 |
| 4 | Crowding:Congruent | 1 | 23 | 34.85038 | 5.115791e-06 | * | 0.1387545 |

3 rows

# Appendix G - Experiment 1 Nonparametric Tests

## Friedman rank sum test (crowding × congruency)

```
	Friedman rank sum test

data:  A and congcrowd and Participant
Friedman chi-squared = 48.113, df = 3, p-value = 2.015e-10
```

## Wilcoxon signed rank test (congruent crowded, congruent uncrowded)

```
wilcox.test(
   congruent_1$A[congruent_1$Crowding == "crowded"],
   congruent_1$A[congruent_1$Crowding == "uncrowded"],
   paired = TRUE,
   alternative = "two.sided"
 )
```

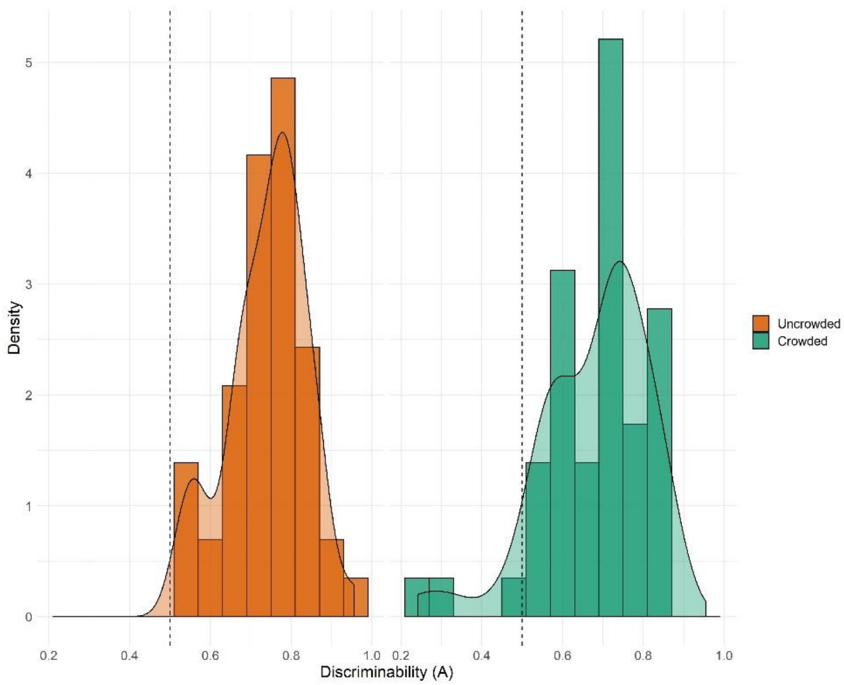
```
	Wilcoxon signed rank test with continuity correction

data:  congruent_1$A[congruent_1$Crowding == "crowded"] and congruent_1$A[congruent_1$Crowding == "uncrowded"]
V = 22, p-value = 0.0004429
alternative hypothesis: true location shift is not equal to 0
```

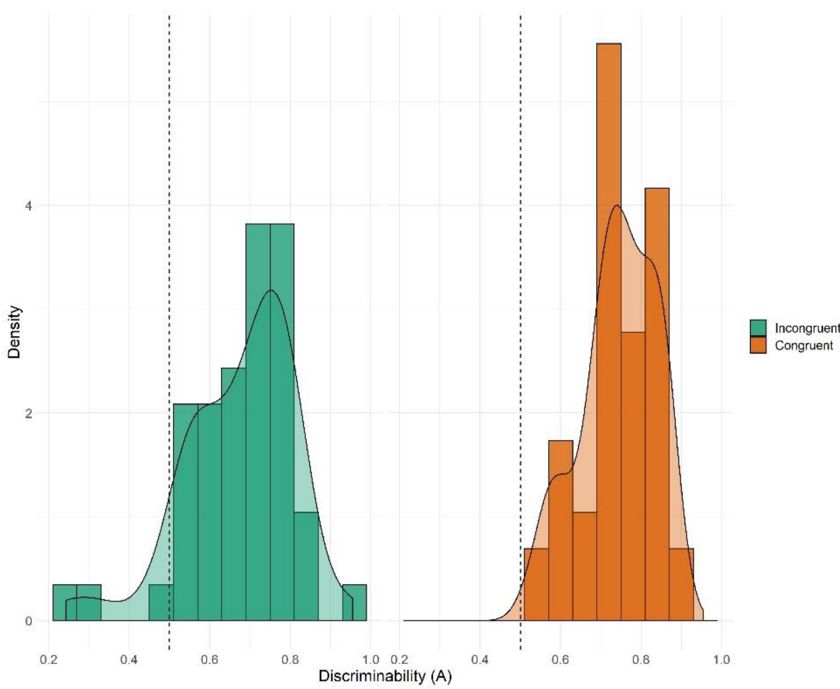
## Wilcoxon signed rank test (comparison to chance)

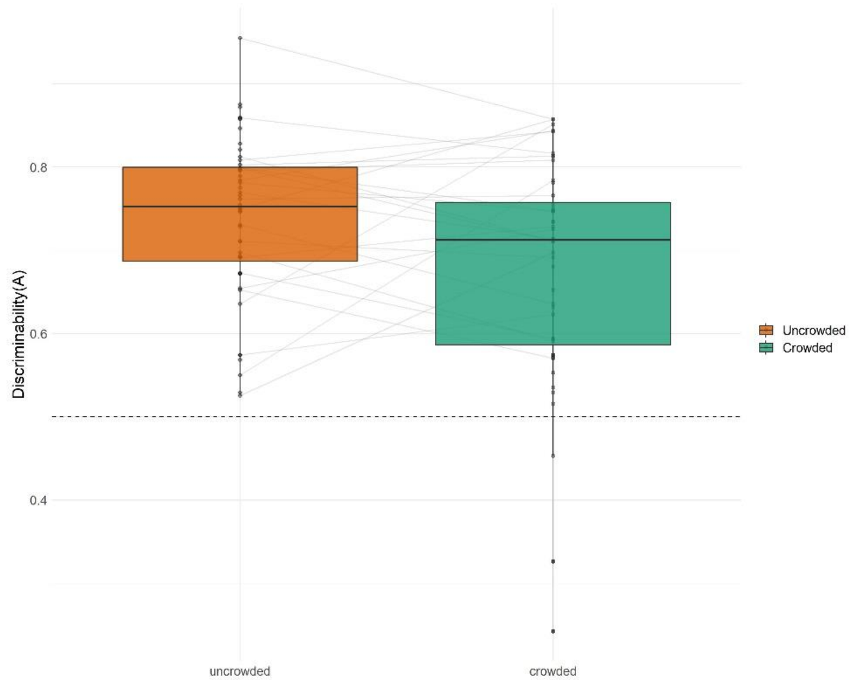
```
wilcox.test(
   my_data$A[my_data$Mode == "Sim"],
   my_data$A[my_data$Mode == "Participant"],
   paired = TRUE,
   alternative = "two.sided"
 )
```

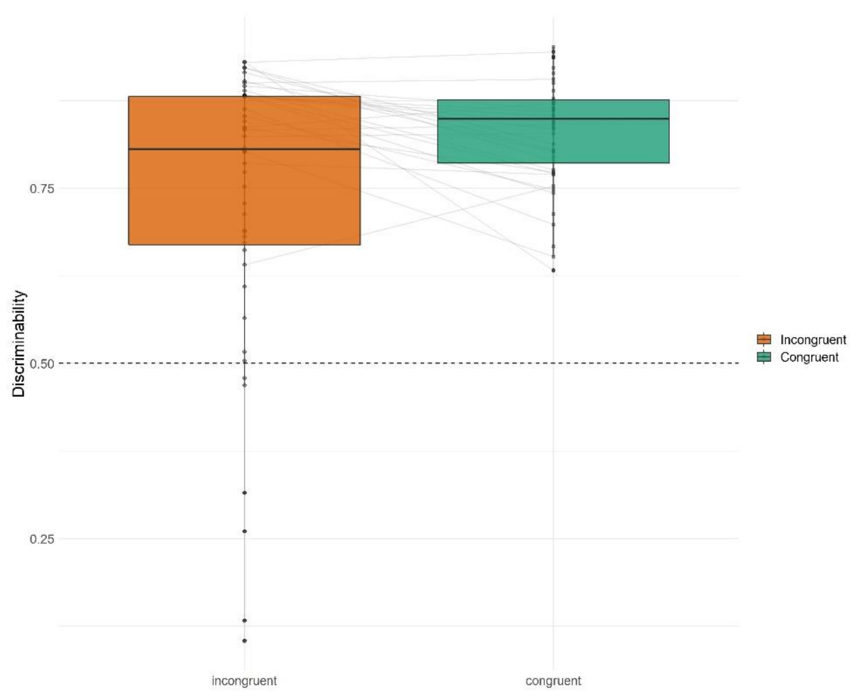
```
	Wilcoxon signed rank test with continuity correction

data:  my_data$A[my_data$Mode == "Sim"] and my_data$A[my_data$Mode == "Participant"]
V = 188, p-value = 5.35e-15
alternative hypothesis: true location shift is not equal to 0
```

Appendix H – Experiment 2 Density Plots and Boxplots




Discriminability(A)
0.8
0.6
0.4
uncrowded
crowded
Uncrowded
Crowded


Discriminability
0.75
0.50
0.25
incongruent
congruent
Incongruent
Congruent

## Appendix I – Experiment 2 Descriptive statistics

### *A* scores - Crowding × congruency conditions

```r
my_data_2_p %>%
  group_by(Crowding, Congruent) %>% summarise(
    mean = mean(A),
    variance = var(A),
    SD = sd(A)
  )
```

| Crowding <fctr> | Congruent <fctr> | mean <dbl> | variance <dbl> | SD <dbl> |
|---|---|---|---|---|
| uncrowded | incongruent | 0.7322184 | 0.010145269 | 0.10072372 |
| uncrowded | congruent | 0.7468669 | 0.009235795 | 0.09610304 |
| crowded | incongruent | 0.6159261 | 0.020325161 | 0.14256634 |
| crowded | congruent | 0.7386475 | 0.007993019 | 0.08940369 |

4 rows

### *A* scores - Congruency conditions

```r
my_data_2_p %>%
  group_by(Congruent) %>% summarise(
    mean = mean(A),
    variance = var(A),
    SD = sd(A)
  )
```

| Congruent <fctr> | mean <dbl> | variance <dbl> | SD <dbl> |
|---|---|---|---|
| incongruent | 0.6740723 | 0.018363973 | 0.13551374 |
| congruent | 0.7427572 | 0.008448371 | 0.09191502 |

2 rows

### *A* scores - Crowding conditions

```r
my_data_2_p %>%
  group_by(Crowding) %>% summarise(
    mean = mean(A),
    variance = var(A),
    SD = sd(A)
  )
```

| Crowding <fctr> | mean <dbl> | variance <dbl> | SD <dbl> |
|---|---|---|---|
| uncrowded | 0.7395426 | 0.009539136 | 0.0976685 |
| crowded | 0.6772868 | 0.017703075 | 0.1330529 |

2 rows

Proportion correct – Crowding × congruency conditions

```
my_data_2_p %>%
  group_by(Crowding, Congruent) %>% summarise(
    mean = mean(mean_PC),
    variance = var(mean_PC),
    SD = sd(mean_PC)
  )
```

| Crowding <fctr> | Congruent <fctr> | mean <dbl> | variance <dbl> | SD <dbl> |
|---|---|---|---|---|
| uncrowded | incongruent | 0.6777344 | 0.008084505 | 0.08991387 |
| uncrowded | congruent | 0.6914062 | 0.006947393 | 0.08335102 |
| crowded | incongruent | 0.5852865 | 0.010741745 | 0.10364239 |
| crowded | congruent | 0.6829427 | 0.006145533 | 0.07839345 |

4 rows

Proportion correct – Crowding conditions

```
my_data_2_p %>%
  group_by(Crowding) %>% summarise(
    mean = mean(mean_PC),
    variance = var(mean_PC),
    SD = sd(mean_PC)
  )
```

| Crowding <fctr> | mean <dbl> | variance <dbl> | SD <dbl> |
|---|---|---|---|
| uncrowded | 0.6845703 | 0.007403759 | 0.0860451 |
| crowded | 0.6341146 | 0.010698900 | 0.1034355 |

2 rows

Proportion correct – Congruency conditions

```
my_data_2_p %>%
  group_by(Congruent) %>% summarise(
    mean = mean(mean_PC),
    variance = var(mean_PC),
    SD = sd(mean_PC)
  )
```

| Congruent <fctr> | mean <dbl> | variance <dbl> | SD <dbl> |
|---|---|---|---|
| incongruent | 0.6315104 | 0.011394961 | 0.1067472 |
| congruent | 0.6871745 | 0.006425465 | 0.0801590 |

2 rows

# Appendix J – Experiment 2 Statistical Analyses

## Comparison to chance t-test

```
t.test(
  my_data_2$A[my_data_2$Mode == "Sim"],
  my_data_2$A[my_data_2$Mode == "Participant"],
  paired = TRUE,
  alternative = "two.sided"
)
```

```

    Paired t-test

data:  my_data_2$A[my_data_2$Mode == "Sim"] and my_data_2$A[my_data_2$Mode == "Participant"]
t = -13.972, df = 95, p-value < 2.2e-16
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.2481118 -0.1863772
sample estimates:
mean of the differences
             -0.2172445
```

```
## Cohen's d
sampleMean1<-mean(my_data_2$A[my_data_2$Mode == "Sim"], na.rm = TRUE)
sampleMean2<-mean(my_data_2$A[my_data_2$Mode == "Participant"], na.rm = TRUE)
sampleSD1<-sd(my_data_2$A[my_data_2$Mode == "Sim"], na.rm = TRUE)
sampleSD2<-sd(my_data_2$A[my_data_2$Mode == "Participant"], na.rm = TRUE)
sampleSDpooled<-sqrt((sampleSD1*sampleSD1+sampleSD2*sampleSD2)/2)
d<-(sampleMean1-sampleMean2)/sampleSDpooled
d
```

```
[1] -1.947587
```

## Two way ANOVA

```
options(contrasts=c("contr.sum","contr.poly"))
ezANOVA(data=my_data_2_p, dv=A, wid= Participant, within= .(Crowding, Congruent))
```

```
$ANOVA
```

| | Effect <chr> | DFn <dbl> | DFd <dbl> | F <dbl> | p <dbl> | p<.05 <chr> | ges <dbl> |
|---|---|---|---|---|---|---|---|
| 2 | Crowding | 1 | 23 | 22.948709 | 7.84878e-05 | * | 0.07816056 |
| 3 | Congruent | 1 | 23 | 14.087979 | 1.03572e-03 | * | 0.09354892 |
| 4 | Crowding:Congruent | 1 | 23 | 9.571776 | 5.12192e-03 | * | 0.06004186 |

3 rows

```
NA
```

Post hoc t-test (congruent uncrowded, congruent crowded)

```
 t.test(
  congruent_2$A[congruent_2$Crowding == "crowded"],
  congruent_2$A[congruent_2$Crowding == "uncrowded"],
  paired = TRUE,
  alternative = "two.sided"
)
```

```

    Paired t-test

data:  congruent_2$A[congruent_2$Crowding == "crowded"] and congruent_2$A[congruent_2$Crowding == "uncrowded"]
t = -0.48253, df = 23, p-value = 0.634
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.04345721  0.02701841
sample estimates:
mean of the differences
           -0.008219401
```

```
## Cohen's d
sampleMean1<-mean(congruent_2$A[congruent_2$Crowding == "crowded"], na.rm = TRUE)
sampleMean2<-mean(congruent_2$A[congruent_2$Crowding == "uncrowded"], na.rm = TRUE)
sampleSD1<-sd(congruent_2$A[congruent_2$Crowding == "crowded"], na.rm = TRUE)
sampleSD2<-sd(congruent_2$A[congruent_2$Crowding == "uncrowded"], na.rm = TRUE)
sampleSDpooled<-sqrt((sampleSD1*sampleSD1+sampleSD2*sampleSD2)/2)
d<-(sampleMean1-sampleMean2)/sampleSDpooled
d
```

```
[1] -0.08855793
```

Post hoc t-test (incongruent uncrowded, incongruent crowded)

```
incongruent_2 <- my_data_2_p  %>%
  filter(Congruent == "incongruent")

 t.test(
   incongruent_2$A[incongruent_2$Crowding == "crowded"],
   incongruent_2$A[incongruent_2$Crowding == "uncrowded"],
   paired = TRUE,
   alternative = "two.sided"
 )
```

```

    Paired t-test

data:  incongruent_2$A[incongruent_2$Crowding == "crowded"] and incongruent_2$A[incongruent_2$Crowding == "unc
rowded"]
t = -4.5345, df = 23, p-value = 0.0001485
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.16934566 -0.06323898
sample estimates:
mean of the differences
             -0.1162923
```

```
## Cohen's d
sampleMean1<-mean(incongruent_2$A[incongruent_2$Crowding == "crowded"], na.rm = TRUE)
sampleMean2<-mean(incongruent_2$A[incongruent_2$Crowding == "uncrowded"], na.rm = TRUE)
sampleSD1<-sd(incongruent_2$A[incongruent_2$Crowding == "crowded"], na.rm = TRUE)
sampleSD2<-sd(incongruent_2$A[incongruent_2$Crowding == "uncrowded"], na.rm = TRUE)
sampleSDpooled<-sqrt((sampleSD1*sampleSD1+sampleSD2*sampleSD2)/2)
d<-(sampleMean1-sampleMean2)/sampleSDpooled
d
```

```
[1] -0.9421645
```

## Holm correction

```
## p values
pvals2 <- c(0.000148, 0.634)
## Adjust p values
bonf  <- p.adjust(pvals2, "bonferroni")
holm  <- p.adjust(pvals2, "holm")
hoch  <- p.adjust(pvals2, "hochberg")
hom   <- p.adjust(pvals2, "hommel")
BH    <- p.adjust(pvals2, "BH")
BY    <- p.adjust(pvals2, "BY")
```

```
holm
```

```
[1] 0.000296 0.634000
```

# Appendix K – Experiment 2 Nonparametric Tests

## Wilcoxon signed rank test (congruent crowded, congruent uncrowded)

```r
wilcox.test(
   congruent_2$A[congruent_2$Crowding == "crowded"],
   congruent_2$A[congruent_2$Crowding == "uncrowded"],
   paired = TRUE,
   alternative = "two.sided"
 )
```

```
    Wilcoxon signed rank test with continuity correction

data:  congruent_2$A[congruent_2$Crowding == "crowded"] and congruent_2$A[congruent_2$Crowding == "uncrowded"]
V = 120, p-value = 0.3993
alternative hypothesis: true location shift is not equal to 0
```

## Friedman rank sum test (crowding × congruency)

```
    Friedman rank sum test

data:  A and congcrowd and Participant
Friedman chi-squared = 22.933, df = 3, p-value = 4.17e-05
```

## Wilcoxon signed rank test (comparison to chance)

```r
wilcox.test(
   my_data_2$A[my_data_2$Mode == "Sim"],
   my_data_2$A[my_data_2$Mode == "Participant"],
   paired = TRUE,
   alternative = "two.sided"
 )
```

```
    Wilcoxon signed rank test with continuity correction

data:  my_data_2$A[my_data_2$Mode == "Sim"] and my_data_2$A[my_data_2$Mode == "Participant"]
V = 126, p-value = 8.62e-16
alternative hypothesis: true location shift is not equal to 0
```