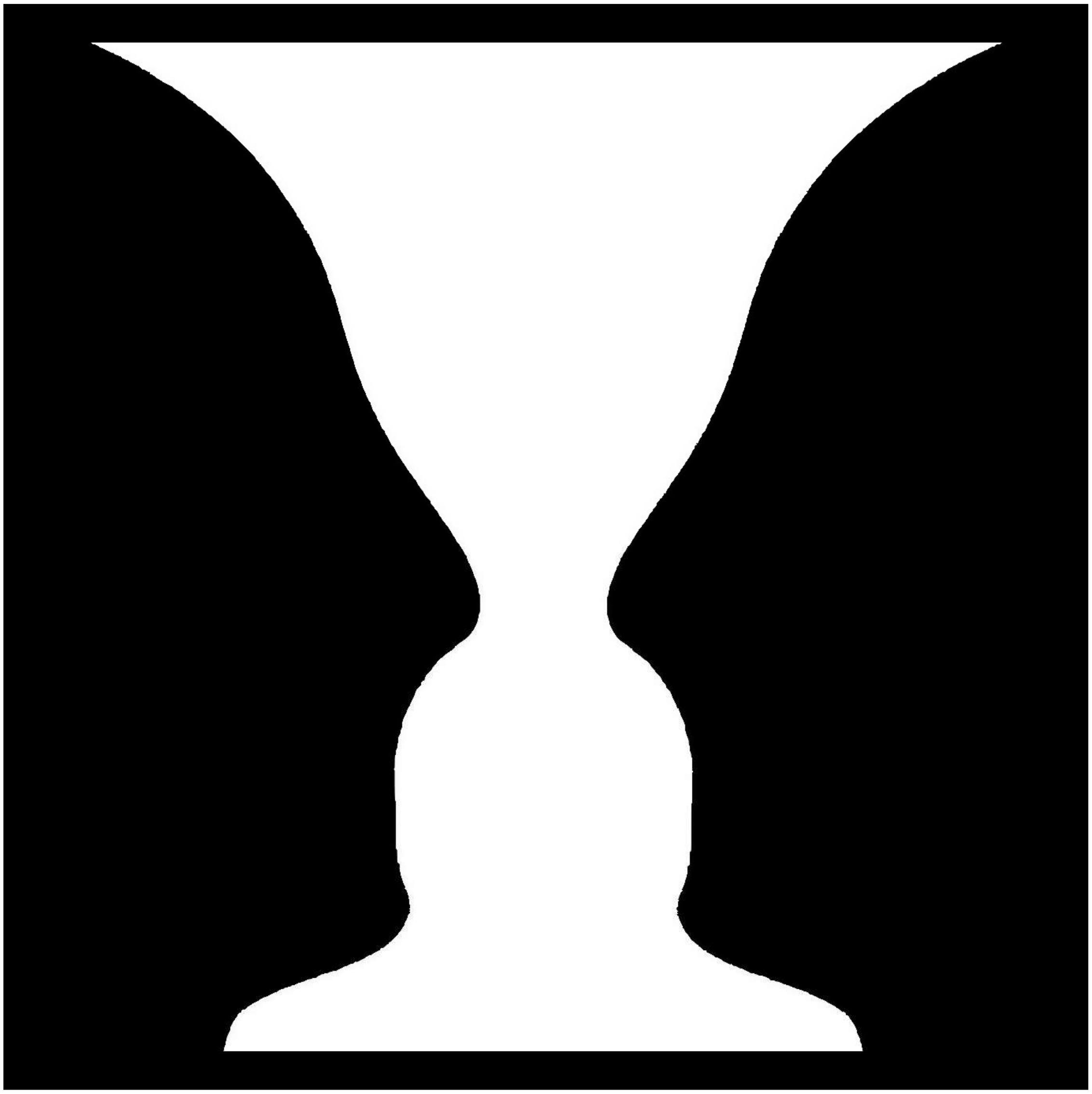


Vase or Face?

Controlling Perception to Investigate Crowding

Thomas Zaugg

Table of Content

Abstract

This research explores the phenomenon of visual crowding by using the Rubin's vase/face illusion—a bistable image that can be interpreted either as a vase or as two faces in profile. Visual crowding is a phenomenon where objects are rendered unrecognizable in clutter. Crowding happens only when the features of the objects are similar enough.

The three core research questions are:

1. Does Rubin's vase/face illusion elicit visual crowding effects?
2. Can we control what individuals perceive within the illusion?
3. Is the extent of crowding influenced by whether the vase or the faces are perceived?

To answer the first question, we designed and tested novel Vernier stimuli with features that are highly similar to those in Rubin's illusion. For the second, we found that manipulating the background color of the illusion was an effective strategy for controlling perception, with changes in contrast and contour playing a key role in determining whether observers saw the vase or the faces. The experiment to settle the third question is still in preparation during the time of writing of this report.

This research offers a new foundation for studying visual crowding through bistable images, which can enrich our understanding of perceptual processes and visual cognition. For the first time, crowding is being studied using bistable images. By keeping the viewed image the same but changing perception, we gain insights into how our brain creates the world we see with the light that hits the eyes. The findings have the potential to inform theories of high-level visual phenomena, including how our brain resolves ambiguity in complex visual stimuli.

Introduction

Visual crowding has been a significant focus of research in the field of vision science, as it directly impacts how we process visual information, especially in our peripheral vision. This research specifically investigates the phenomenon of visual crowding through the use of Rubin's vase/face illusion, a well-known bistable image that challenges our perceptual system by allowing for two distinct interpretations of the same visual information. We aim to determine whether it is possible to control what individuals perceive in such an illusion and to investigate how this perceptual choice affects the degree of crowding experienced.

Bistable images, like Rubin's illusion, present an intriguing challenge to the visual system because while the image stays the same, the brain alternates between two distinct interpretations. This phenomenon raises fascinating questions about how perception is constructed and whether external factors can influence which interpretation dominates. The main focus on our research is whether visual crowding—typically seen when stimuli are presented in close proximity to other object—behaves differently depending on the perceived interpretation of the illusion.

Visual Crowding

Visual crowding occurs when objects are close to each other in the visual field, particularly in peripheral vision, and the viewer experiences difficulty in distinguishing individual objects. For example, reading a single letter becomes harder when that letter is surrounded by others. A more classical example is a Vernier - a pair of misaligned lines seen in Figure 1 - surrounded by Flankers. In such cases, it is often the similarity between the features of the objects (such as shape, size, color, or orientation) that leads to crowding. This phenomenon is of particular interest because it reveals important aspects of how our visual system organizes and processes complex scenes, particularly when we are not directly focusing on the objects in question.

Figure 1: The offset of the upper unflanked Vernier is a lot easier to recognize than the lower flanked one.

Visual crowding is typically studied using a control condition, in which a Vernier stimulus is presented in isolation against a neutral background, and an experimental condition, in which the Vernier is surrounded by "flankers" or embedded within an image. In most experiment, the Vernier stimulus alternates randomly between two configurations (misaligned either to the left or the right), and participants report which configuration they perceive. Importantly, participants are instructed not to look directly at the stimulus but to perceive it peripherally. The percentage of correct responses provides an index of crowding; fewer correct answers indicate greater crowding, suggesting that the presence of surrounding stimuli significantly disrupts perception.

Previous studies have shown that visual crowding typically occurs in peripheral vision, where our visual system has a harder time resolving fine details. By studying crowding within Rubin's vase/face illusion, we aim to further investigate how the perceptual shift between vase and faces might alter crowding effects, particularly given that the interpretations involve inverted figure/ground configurations and different boarder ownership.

Bistable Images

Bistable images, such as Rubin's vase/face illusion, are fascinating examples of how the brain can interpret the same set of visual information in two different ways. Rubin's illusion, in particular, allows for the perception of either a central vase or two faces in profile, depending on which features the observer's brain emphasizes. Bistable perception occurs when the brain spontaneously shifts between these two interpretations, even though the visual stimulus itself remains unchanged. This phenomenon provides a unique window into the workings of the brain, as it highlights how perception is not merely a reflection of objective reality but is influenced by cognitive processes that select one interpretation over another.

Brain imaging studies have explored the neural basis of bistable perception, revealing that shifts in interpretation may involve higher-order visual areas that exert top-down influence on early visual processing. Modulations in brain activity have been observed as early as in the V1 area of the visual cortex, suggesting that bistable

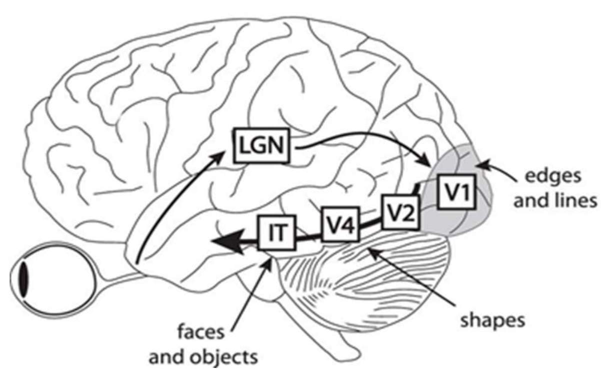


Figure 2: The classical feed-forward model of perception with the different areas involved in visual processing.

perception taps into both low-level sensory processing and higher-level cognitive functions.

Goal of the Research

My research project during the internship was driven by the goal of investigating whether Rubin's vase/face illusion exhibits visual crowding, and whether crowding differs depending on the perceived image (vase or faces). Specifically, the research sought to address the following questions:

1. Does Rubin's vase/face illusion exhibit crowding?
2. Can we influence what individuals perceive in the illusion?
3. Can we measure a difference in crowding depending on whether the observer perceives the vase or the faces?

These questions formed the basis for the successive experiments and modifications that we introduced throughout the course of the internship.

Experimental Design

Stimulus Design

We initially tested how a classic Vernier stimulus behaved when placed within Rubin's vase/face illusion. As expected, the straight, thin lines did not crowd significantly, as they lacked similarity with the curvilinear contours of the faces in Rubin's illusion. This outcome led us to redesign the Vernier stimuli to more closely match the contours of the faces and the vase. After several iterations, we settled on a design in which the Vernier lines were curved, resembling the contours of the faces in Rubin's illusion. This new design allowed us to manipulate crowding effects more effectively, as we could now vary the difficulty by altering the vertical displacement of the curves rather than just the horizontal displacement.

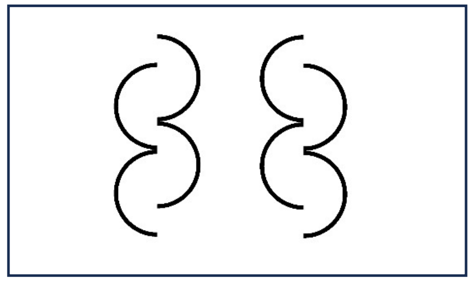


Figure 3: The new stimuli designed for the Rubin's illusion.

Controlling Perception

Attempt 1: Priming

Our first attempt to influence what people perceive in Rubin's illusion involved priming participants by briefly flashing either an image of a vase or two faces before showing them Rubin's illusion. The idea was that priming participants with one image would bias their perception toward that image when presented with the illusion. We tested both realistic and abstract versions of the vase and faces, flashing them for 250 milliseconds, followed by a 400-millisecond blank screen, and then presenting Rubin's illusion for 100 milliseconds.

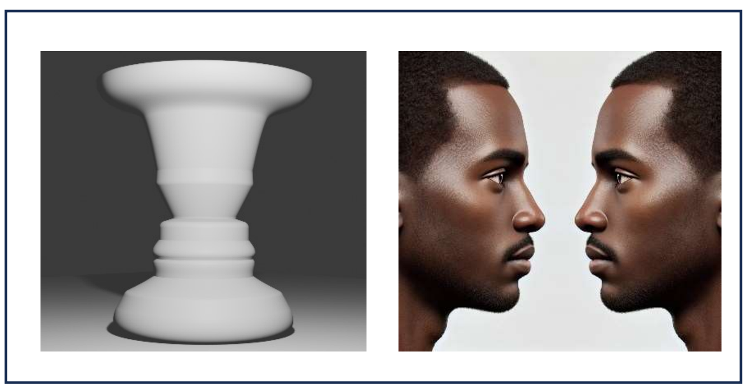


Figure 4: Realistic versions of the two precepts.

However, our results were surprising. Instead of the expected priming effect, we observed the opposite: flashing a vase image led participants to perceive faces more often, while flashing a face image led them to perceive the vase. Although the overall effect could be used, it was too small and reliable. As a result, we abandoned the priming approach and turned to other methods for controlling perception. However, we found that people perceived more often the face, which was why we made the image more vase-like by smoothening the contours.

Attempt 2: Background Color

Our second approach involved altering the background color of Rubin's illusion to control what participants perceived. We hypothesized that if the background color matched the faces, the vase would stand out more, and vice versa. To avoid creating a strobo-disco experience, we made the color changes more subtle by reducing the contrast between the faces and the vase as can be seen in Figure 5:

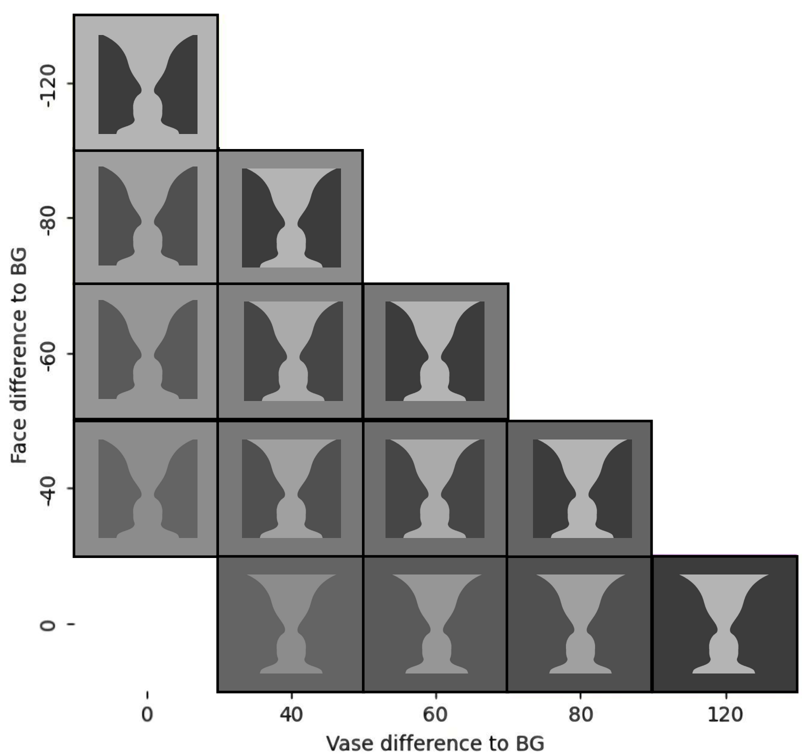


Figure 5: This is a collection of all the different contrast combinations explored placed on different backgrounds.

In MATLAB, we adjusted the grayscale image of Rubin's illusion, where the faces originally had a value of 0 (black) and the vase had a value of 255 (white). We modified these values so that the faces had grayscale values ranging from 60 to 100, while the vase had values ranging from 140 to 180. This manipulation effectively influenced perception: when the background matched the faces, participants were more likely to report seeing the vase, and when the background matched the vase, participants were more likely to report seeing the faces.

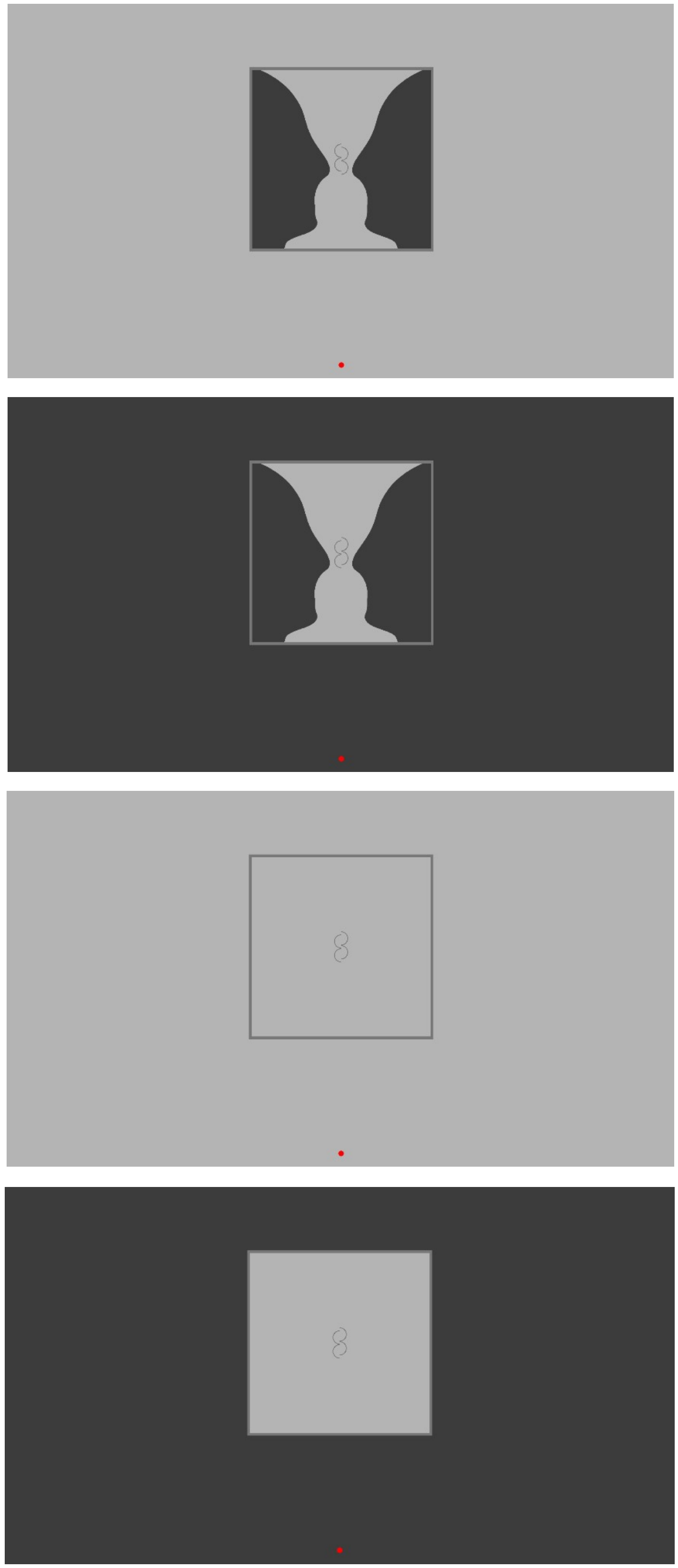
Final Adjustments

After developing a method to control perception and creating an appropriate Vernier stimulus, we made additional refinements to the experiment. These included adjusting the stimulus presentation duration and the spacing between the Vernier lines. We also observed that participants often focused exclusively on the Vernier offset without paying attention to the vase or faces. To address this, we introduced an additional question asking participants what they perceived—whether they saw the vase or the faces—after each trial. We also added a neutral grey border around the stimulus to maintain consistent contrast and added two control conditions consisting of a square with the different backgrounds.

Figure 6: Me in experiment room.

Final Experiment Design

The experimental setup involved participants being seated in a dimly lit room in front of a Samsung SyncMaster 957DF CRT monitor (31 cm × 23 cm, 1024 × 768 pixels, 100 Hz). A headrest was positioned 75 cm from the screen to ensure a consistent viewing distance. Participants began with a practice trial, during which they received feedback in the form of an auditory beep for incorrect answers. A researcher was present during the practice phase to answer any questions and ensure that participants understood the task. Once participants demonstrated proficiency with the task, the practice trial ended, and the main experiment began. At this point, the researcher left the room, and data collection began. On the following page, the four conditions can be examined.



Further Steps

While considerable progress was made during the eight-week internship, the project is still ongoing. Fortunately, I have been offered a position as a student assistant in the lab, allowing me to continue working on this research. One of the primary challenges we currently face involves the participants' attention. When asked whether they saw the vase or the faces, their perception aligned with our expectations based on the background manipulation. However, when asked about the offset of the Vernier, participants seemed to focus solely on the Vernier and not on the vase or faces. To address this, we implemented a dual-task approach, requiring participants to answer both questions (about the vase/face perception and the Vernier offset). However, this approach proved too demanding for most participants, as it exceeded the capacity of their working memory.

Moving forward, the key challenge will be to find an experimental setting that allows participants to simultaneously perceive both the vase/face illusion and the Vernier offset without overwhelming their cognitive resources. This may involve simplifying the task or altering the stimuli in ways that reduce cognitive load.

Conclusion

This is the first time that crowding has been studied using bistable images. By keeping the viewed image the same but changing perception, we can gain valuable insights into how the brain constructs visual reality from the sensory information it receives. During this project, we successfully designed a new Vernier stimulus that is compatible with Rubin's vase/face illusion and developed a method for controlling perception using background color. The next major challenge is to find an experimental setup that allows participants to perceive both the illusion and the Vernier offset simultaneously. Regardless of the specific outcomes, this research opens the door to further studies on crowding using other bistable images, such as the Necker cube or the optical illusion known as #theDress.

Reflection

This internship was a gift. I gained new skills—MATLAB, statistical analysis, and insights into crowding research. But more than that, I worked with a team of researchers, all focused on pushing the boundaries of knowledge. I saw the highs and lows of the scientific process, felt the frustration and uncertainty, but also curiosity and those rare moments of clarity. It made clear to me that I want to continue pursuing research. And then there were the moments outside the lab—dinners, drinks by Lake Geneva—which made the experience richer. I am excited to continue working at the lab.

Acknowledgments

I would like to express my sincere gratitude to the Laidlaw Foundation for supporting this research opportunity. I’m also very thankful to Prof. Michael Herzog and the entire lab for their warm welcome and invaluable support throughout the project. Special thanks go to Lisa Schwetlick for her continuous guidance and mentorship.

References

Kornmeier, Jürgen, and Michael Bach. "Early Neural Activity in Necker-cube Reversal: Evidence for Low-level Processing of a Gestalt Phenomenon." *Psychophysiology* 41, no. 1 (January 2004): 1–8. https://doi.org/10.1046/j.1469-8986.2003.00126.x.

Kornmeier, Jürgen, and Michael Bach. "Ambiguous Figures – What Happens in the Brain When Perception Changes But Not the Stimulus." *Frontiers in Human Neuroscience* 6 (2012). https://doi.org/10.3389/fnhum.2012.00051.

Leopold, David A., and Nikos K. Logothetis. "Multistable Phenomena: Changing Views in Perception." *Trends in Cognitive Sciences* 3, no. 7 (July 1999): 254–64. https://doi.org/10.1016/S1364-6613(99)01332-7.

Lupyan, Gary. "Changing What You See by Changing What You Know: The Role of Attention." *Frontiers in Psychology* 8 (May 1, 2017). https://doi.org/10.3389/fpsyg.2017.00553.

Parkkonen, Lauri, Jesper Andersson, Matti Hämäläinen, and Riitta Hari. "Early Visual Brain Areas Reflect the Percept of an Ambiguous Scene." *Proceedings of the National Academy of Sciences* 105, no. 51 (December 23, 2008): 20500–504. https://doi.org/10.1073/pnas.0810966105.

Tse, Peter U., Eric A. Reavis, Peter J. Kohler, Gideon P. Caplovitz, and Thalia Wheatley. "How Attention Can Alter Appearances." In *Handbook of Experimental Phenomenology*, edited by Liliana Albertazzi, 1st ed., 291–315. Wiley, 2013. https://doi.org/10.1002/9781118329016.ch12.

Images

All the images and figures were created by the authors with the exception of:

Figure 2: Manassi, Mauro, Bilge Sayim, and Michael H. Herzog. “When Crowding of Crowding Leads to Uncrowding.” *Journal of Vision* 13, no. 13 (November 8, 2013): 10. https://doi.org/10.1167/13.13.10.