Can emotionally-retrieved semantic information influence visual perception?

Victor Cornejo

3037066554

University of California Berkeley

December 2021

Introduction

Hume once wrote that the subjective experience of color and that of emotion were impressions derived directly from the senses. Under his view, thoughts about color or emotion are based on previous subject experiences of these concepts. Therefore, he implied that a blind person could never *picture* colors, that a deaf person could not imagine sounds, and that a reasonable person would never think of *cruelty.*

Imagine now, 300 years later, his awe when realizing that congenitally blind people understand subjective experiences of color through basic descriptions of object properties; that the subjectiveness of emotions is dependent on *language*; and that we can learn about how emotions affect vision processing by terrifying people who stare at muffins and chihuahuas as this Continuous Flash Suppression (CFS) experiment proposes. This experiment is inspired by a dataset meant to train image-recognition convolutional neural networks, which are pertinent to the topic as they are based on the structure of the neural layers in the retina.

Current discussions about *penetrating* the early visual system tend to refer to the argument posed by Pylyshyn (Pylyshyn 1999). Under this argument, he proposed that the early visual system was separate from that of cognition, attention, and memory as if the visual system is a separate complex system rather than a function of the brain, and regarding the visual system more as an organ, like the heart. This view implies that the early visual system is just the structure in which only *object shape* recognition takes place (amidst contrast and borders). That is the most basal sense, out of any abstract *meaning* or correlation, as these are functions that lie on cognition.

Given the current state in the Cognitive Sciences, this is a constantly challenging argument that is piling up evidence against it. As Pinker said (Pinker 1994), "No matter how influential language might be, it would seem preposterous to a physiologist that it could reach down into the retina and rewire the ganglion cells." Moreover, what would it mean for fear-response stimuli to facilitate access into an object's semantic properties without the need for *attention* and beyond an *image index,* as Pylyshyn (Pylyshyn 1999) proposes through Kahneman? What if our visual working memory can store semantic knowledge and not just *procedural*? Is the visual system's processing in *cascade, parallel distributed,* or simply *sequential* and ultimately impenetrable? This research proposal could provide us with some insight into these matters.

Nonetheless, contemporary understanding of how the early visual system works implies that no structure capable of processing semantic information could access the system, given that the complexity of its layers limits the properties of this one. In other words: the early visual system is too *bottom-up*. More specifically, for the early visual system, its function consists of filtering light through linear operations (dot product) concerning detecting edges, lines, and shapes (v1). As opposed to the primary visual system, in the subsequent stages (selection and encoding), the complexity of their function shifts to a non-linear (spatiotemporal, where and what) processing of information which translates into fire rate responses.

Then, what does it mean for semantic knowledge to affect visual perceptions? It is contemporarily understood that the *where* and *what* processing locations within the visual system are particular to their function, yet it is debated the degree of their dependence. By posing this question, what is the extent to which the knowledge we have about an *object's* properties can influence the *process* in which we visually perceive such objects at the basal level? Is it even possible? This experiment suggests that this is the case by conditioning participants to fear *unrecognizable* object-entities (puppy/muffin) and then observing the correlation between the response time under continuous flash suppression.

Key terms: semantic knowledge, serial-sequential, parallel distributed processing, apperceptive level, associative level, bottom-up, top-down, CFS.

In response to Pylyshyn (Pylyshyn 1999), Yu (Yu 1999) supposed that the early visual system may not be as impossible to penetrate. Under her view, a car could be perceived as moving forward if the front was shown, and a banana would always be imaged as yellow because the *semantic properties* of an object could influence the way we process them. Under this view, the visual system could benefit from this penetrability by establishing a connection with semantic processing systems as this would increase its prediction accuracy, although at the cost of a few errors here and then.

A recent study from Portugal (Gomes, Silva, Silva, Soares 2016) considers the ecological and evolutionary pressures of fear. Their findings support the theory that the early visual system is penetrable since, with the aid of a **CFS** test, the images shown to the participants were suppressed at the level of the fovea. By assailing participants with images of snakes, spiders, and birds, they could spot a correlation between the participant's response time in object recognition. In this way, the participants would not be *aware* of the images projected on the screen unless, as their hypothesis suggested, the association of *fear* within poisonous predators such as spiders and snakes sped up the response time in which the participants became aware of their shown *referents*.

Nonetheless, the study only deals with stimuli which response one could say is already coded in our 'fight or flight' DNA. What is meant by this is that the objects/categories used belong to those of creatures that instinctually elicit a fear response shaped by evolution across millennia and without consideration to constantly changing aversive stimuli that we can adapt. Such that anyone who has never seen a snake frights or jumps when in contact with one for the first time, but not everyone fears a *final exam* unless some conditioning has taught them to do so. On the other hand, this type of automatic response is relevant as a follow-up to what Yu (Yu 1999) had mentioned about the trade between *speed* and *accuracy* when predicting.

Exposing the reasonable correlation between language and perception and how semantic information could be processed in the brain. In a different study, Norman et al. (Noorman, Neville, and Simanova 2018) showed that linguistic labels (which are a form of semantic knowledge) facilitated object recognition in an experiment that asked participants to match images of fruits and vegetables in the categories of ‘fruit’ or ‘vegetable’... They concluded that "visual perception can be affected by the *top-down* guidance of words," which means that hearing a word cue activated a representation of its referent’s shape, which would then affect the shape processing of subsequent visual stimuli. However, one of their significant limitations was the nature of the objects they used in the experiment; since they were fruits and vegetables, no correlation was found between categorical associations but only shape-based associations. This contingency falls on Pylyshyn's argument about *basic objects* and *natural categories* (Pylyshyn 1999), where the referent's categorical mental representations and physical shapes were too similar.

Then, as shown by Boroditsky et al. (Lupyan, Rahman, Boroditsky and Clark 2020), language *can shape* what we perceive as an automatic process at the *apperceptive level,* implying that to recognize something is to categorize it. In other words, to recognize an object, it is necessary to compare the category of the object to these categories that the object does not pertain to. Quite a Hegelian take and possibility that are rooted in semantic knowledge and the *penetrability* of early vision.

In one of their findings, using a *Mooney image* and similarly to the study by Noorman et al. (Noorman, Neville and Simanova 2018), participants were able to identify the *hidden* object once they were given a list of possible names that could match the entity. In another of their tests, participants shortened their time responses in an ‘*odd-one-out’* visual search task once they were told that the objects shown were inversed numbers. In a different test, participants were shown the contoured shape of a bird where after reading passages about movement, they were asked to draw a worm in the bird's peak, a position *correlated* to the directionality that was described in the passage they had just read.

These studies present information supporting this hypothesis, yet their contingencies are the aspects on which this proposal relies. This proposal aims to incorporate the cue word from the object recognition test’s ambiguity and the CFS recognition test where the stimuli presented were non-dynamic and provided no adaptation to changing stimuli. Moreover, adding an electric shock into this experiment is based on simple Pavlovian conditioning whereby exposing participants to a neutral stimulus (amorphic entity) and aversive stimulus (shock) fear can be learned and associated with the neutral stimulus.

However, participants should be asked for consent to be given a shock before the experiment, and the shock level intensity will have to be calibrated to the point where this one is *just* tolerable. It is also essential to clarify that the participants can stop and leave the experiment at any point if they feel it is pertinent or necessary.

Experiment methods

Sample Specifics: Humans, any ethnicity, any education, and preferably right-handed to easily predict ocular dominance. Healthy vision and no medications. No people with epilepsy. No under 18. Because we want to assure that participants are in a perfect state of their visuoperceptual capacities and decision-making, no people with epilepsy because CFS could involuntarily cause an attack, and no under 18 because we do not want to electrocute children. We compare a single group of participants at different time points, Within-Subjects.

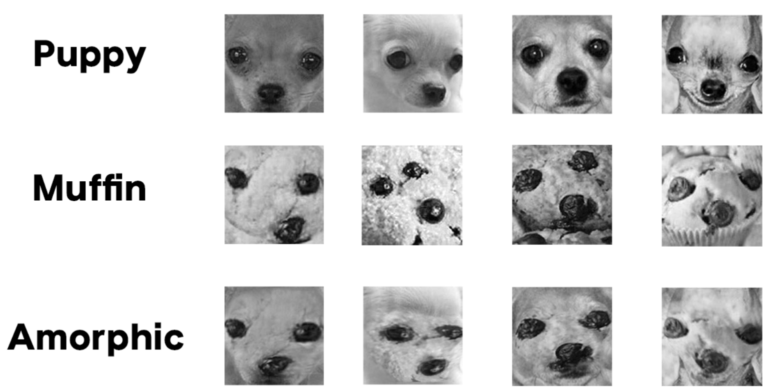
Materials Required: A computer. Specific software capable of: Accurate response-time recording, CFS, and shock administration. Chihuahuas and Blueberry Muffins dataset. CFS optic devices. Shock machine.

Technology Use: Use software to measure the response time of each of the participants before conditioning and after the conditioning.

Experiment Design: Participants are coming for one day, and they will be tested on three different occasions in a time frame of approximately two hours with short breaks. Experimenting one day in a short time window makes it practical since the fear of being conditioned could fade away if not reinforced enough. On the other hand, its length can also make the process a bit too tiresome and even unpleasant enough for some participants to quit.

Experiment Procedures: In the first part of the experiment, placed under a CFR Device, participants will be asked to correctly grade the relationship between the categorical cue words of 'puppy' or 'muffin' with the shown image of either a 'puppy,' a

Figure 1
Categorical Representation



In order to *recognize* the correct category of each item, the early visual system should retrieve *semantic* information from the categories ‘Puppy’ or ‘Muffin’ in order to return that the amorphic figure is none. Later visual processing stages are required for this task.

'muffin' or a amorphic composite of both. For example, the specifically designed software shouts a word, and subsequentially, an image is displayed for less than a second, then the participant has to press a specific key in the computer for the relation between the word cue and the shown image correct or incorrect.

In the second part, we remove the CFS device and make participants follow the same process as the first step, although another grading choice is introduced: incorrect or N/A (when the *amorphic entity* appears). Therefore every wrong answer proceeds with an electric shock.

For the third part, everything in the first part is repeated by introducing the electric shock possibility as an incorrect answer (when the *amorphic entity* appears).

Data plan

To support the alternative hypothesis that *emotionally-retrieved semantic information affects visual perception,* we need to observe that fear (independent variable) affects the response time of the participants (dependent variable). The change can be in any directionality, either towards the alternative hypothesis or in support of the null. Also, we could conclude that we do not have enough information. Moreover, at the moment of this experiment, we would not measure any other event that can cause a false co-relation, such as health conditions, exercise, what they eat before the experiment, and more confounding variables. Finally, the participants are intended to be a random sample from the population.

If the conditioned fear does have a positive effect, we would expect the average response times of all the participants in the sample to decrease after the second step, where fear was conditioned. It is hypothesized that under a fearful situation that *evolutionarily* meant danger, the response time in *recognizing* a threat in the environment was crucial for survival.

What is meant by this is that by exposing participants to a learned fear stimulus besides observing the change in the reaction times, we can also observe how the brain processes visual input in the context of our current understanding of information processing in the early visual system. This experiment can prove that the early visual system can directly retrieve semantic information instead of the held notion that it cannot.

Conclusion and further directions

The presence of a threat has been shown to decrease the response times of danger recognition, yet the findings of this experiment are transcendental in that it can show an existing *penetrability* of the early visual system as previous attempts have not. Because observing a decrease in the average response time means that semantic information processing had to occur for the visual system to recognize the stimuli shown on screen.

The shown images belonged to muffins, puppies, or a amorphic figure of both, the latter form of stimulus has no category to be referred to as an *image index* corresponding to the shape of the object, as mentioned earlier: “to recognize an object it is necessary to compare the category of the object to these categories that the object does not pertain to." If the early visual system was impenetrable and only able to compute shapes and shades, *making sense* of the object seen would not be possible at the v1 level as the v4, and IT regions of the visual cortex are in charge of object and face recognition. Therefore, it can be concluded that under a *threatening* situation, the visual system is penetrable and that emotionally-retrieved semantic information influences visual perception in *top-down* processing and not just *bottom-up*.

Further considerations are to implement EEG into this area so we can see what regions of the visual cortex are active when the aversive stimuli are shown, as this would help us predict not just 'if' but 'how' the penetrability of early vision happens. Can our visual working memory store semantic information? Would it make sense for a system to compartmentalize its memory to access it quicker in danger? Or does the visual system compute information parallelly in a way we cannot yet understand?

References:

Pylyshyn, Z. (1999). Is vision continuous with cognition?: The case for cognitive impenetrability of visual perception. *Behavioral and Brain Sciences*, *22*(3), 341–365. https://doi.org/10.1017/s0140525x99002022

Yu, K. (1999). Penetrating the impenetrable? *Behavioral and Brain Sciences*, *22*(3), 401–414. https://doi.org/10.1017/s0140525x9961202x

Noorman, S., Neville, D. A., & Simanova, I. (2018). Words affect visual perception by activating object shape representations. *Scientific Reports*, *8*(1). https://doi.org/10.1038/s41598-018-32483-2

Gomes, N., Silva, S., Silva, C. F., & Soares, S. C. (2017). Beware the serpent: the advantage of ecologically-relevant stimuli in accessing visual awareness. *Evolution and Human Behavior*, *38*(2), 227–234. https://doi.org/10.1016/j.evolhumbehav.2016.10.004

Lupyan, G., Abdel Rahman, R., Boroditsky, L., & Clark, A. (2020). Effects of Language on Visual Perception. *Trends in Cognitive Sciences*, *24*(11), 930–944. https://doi.org/10.1016/j.tics.2020.08.005