

Effect of Spatial Frequency and Filter Condition on Vividness of Visual Imagery

Adrian Wong

Dr. Frank Tong

PSY-PC 4999 Honors Thesis

Spring 2024

Abstract

Mental imagery has been characterized as the isolation of the feedback process that occurs during normal perception, since stimuli are “perceived” without actual feedforward sensory input. Predictive coding theory has been used to explain, among other phenomena, visual perception and visual imagery, and the theory makes several predictions about the phenomenological experience of visual imagery, which is most often weak and blurry when compared with the crisp and clear experience of visual perception. Specifically, we expected imagery strength to decrease with increasing spatial frequency, and that low-pass filtered stimuli would allow for greater imagery strength than high-pass filtered and non-filtered stimuli. To test these hypotheses, 16 participants completed two experiments using a binocular rivalry paradigm. Participants’ imagery strength was operationalized as the percentage of trials where the imagined stimulus matched the dominant stimulus during rivalry. In Experiment 1, there were no significant differences between the imagery strength of low, medium, and high spatial frequency stimuli. In Experiment 2, imagery strength of low-pass filtered stimuli was not significantly different than that of high-pass filtered and non-filtered stimuli. For both experiments, participants’ overall imagery strength was not correlated with their total score on the Vividness of Visual Imagery Questionnaire 2 (VVIQ2). Taken together, these results show that spatial frequency and filter condition are stimulus properties that cannot affect strength of visual imagery. The conclusions of this study fail to verify the predictions about the nature of mental images made by the predictive coding theory of brain function.

Introduction

Mental imagery is the experience of perceiving mental representations without sensory input. It is our ability to mentally form and report imagined qualities of any possible or conceivable sensory experience, and plays a crucial role in memory, decision-making, and planning for the future. Additionally, understanding the nature of mental imagery and the neural mechanisms that support it are crucial for developing therapies and treatments for mental disorders characterized by hallucinations, such as schizophrenia and post-traumatic stress disorder. Though mental imagery of all modalities exists, this study will focus on visual imagery since it has been the most studied and best understood.

Early work in visual imagery demonstrated that imagery and perception of the same modality share common processing mechanisms. Specifically, researchers showed that imagining a visual pattern would interact with concurrent sensory input to boost detection performance (Ishai & Sagi, 1995). A positron emission tomography (PET) study of brain-damaged patients examined the commonality of activated brain regions during visual perception and imagery, finding that two thirds of all brain areas active in either perception or imagery were active in both cases (Kosslyn et al., 1997). The shared processes between imagery and perception demonstrate a high degree of specificity and specialization as illustrated by O'Craven & Kanwisher (2000). Using functional magnetic resonance imaging (fMRI), the researchers showed that when faces were imagined, the fusiform face area (FFA) was activated, while imagery of indoor or outdoor scenes activated the parahippocampal place area (PPA). There was little to no cross-activation of the FFA by imagining scenes nor of the PPA by imagining faces. These are similar results to brain activation during normal visual perception (O'Craven & Kanwisher, 2000). Also using fMRI, Slotnick et al. (2005) investigated whether the nature of visual imagery representations were symbolic (language-like) or depictive (picture-like). The researchers found that early visual areas were

SPATIAL FREQUENCY, FILTER CONDITION, IMAGERY VIVIDNESS

activated retinotopically by visual imagery, similar to perceptual activation, which supports visual imagery representations being depictive (Slotnick et al., 2005). These foundational studies clearly demonstrate many commonalities between visual perception and visual imagery, but how can they be explained? We argue that predictive coding can provide a parsimonious yet accurate explanation for both abilities under a unified theory of brain function.

Predictive coding

Predictive coding models of visual perception can be used to explain imagery as a phenomenon. These models posit that perception arises from incoming sensory input being checked against our internal models of the world, which are iteratively updated to minimize prediction error (Friston, 2005; Lee & Mumford, 2003; Rao & Ballard, 1999). In Bayesian terms, posteriors are generated from the interaction between priors and likelihoods. Imagery in this account is the result of the priors about some stimuli being perceived in the absence of sensory input.

In line with predictive coding models, recent neuroimaging evidence has shown that visual imagery and perception rely on similar neural mechanisms. Imagined and perceived stimuli share neural representations in the visual, parietal, and frontal cortices, and imagery and perception rely on similar top-down connectivity (Ishai et al., 2000; Dijkstra et al., 2019). However, bottom-up processing is lacking during imagery since there is no sensory input, unlike perception which relies on both top-down and bottom-up processing. In addition, a recent MEG study showed that information flow during imagery is reversed as compared to the perceptual feed-forward cascade (Dijkstra et al., 2020). Representations of the imagined stimulus are activated via feedback processing from high-level to low-level visual areas. Essentially, imagery can be thought of as the isolation of the feedback visual processing stream. Since higher order visual areas have neurons with larger receptive fields (regions of

SPATIAL FREQUENCY, FILTER CONDITION, IMAGERY VIVIDNESS

space that the neuron is most responsive to) (Smith et al., 2001), representations generated in those areas during imagery lack the fine-grained detail of perceptual representations. As per predictive coding, during perception, top-down signals interact with bottom-up inputs to continuously update internal models of the world. In contrast, since there are no inputs to check top-down signals against during imagery, lower-level areas with smaller receptive fields and higher acuity have to infer the sensory causes of the representations being sent downstream from higher-level areas, which have larger receptive areas and lower acuity. This reversal of the feedforward sweep could explain why phenomenologically, imagery appears and feels different from perception despite sharing much common circuitry (Keonig-Robert & Pearson, 2020).

The above account of visual imagery makes several predictions about the nature of mental images. The reduced level of detail in imagery representations, as a result of the larger receptive fields of neurons in high-level areas, suggests that imagery of stimuli with lower spatial frequencies would appear more vivid to the imager. Also, low-pass filtered stimuli, which can approximate the blurry phenomenological experience of imagery relative to normal vision, would be more readily imagined than non-filtered or high-pass filtered stimuli. These predictions might make intuitive sense, but how can they be investigated empirically?

Binocular rivalry

Despite its important role in both everyday functioning, the private nature of mental images has made it a difficult phenomenon to probe and understand in the past. However, recent advances in technology, particularly in fMRI as well as objective psychophysical methods like the binocular rivalry paradigm, have allowed us to understand the workings of the imagining mind and brain with much greater clarity.

Binocular rivalry has been used extensively to study how visual imagery affects perception (Pearson et al., 2008). With the underlying model of visual perception and

SPATIAL FREQUENCY, FILTER CONDITION, IMAGERY VIVIDNESS

imagery sharing common neural mechanisms, Pearson and others have found the top-down bias effect of imagery on subsequent perception during binocular rivalry. The classic binocular rivalry method consists of presenting conflicting visual patterns one to each eye, such as a green vertical grating shown to the left eye and a red horizontal grating shown to the right eye. Because of competitive interactions in the visual cortex, only one pattern is perceptually dominant at any one time, and the other image is suppressed from awareness. This competition results in perception spontaneously flipping from one pattern to the other and back. Pearson et al. designed a paradigm where participants were first presented with a fixation dot on a dark screen for 10 seconds and were told to view the screen passively, to imagine a red horizontal grating, or to imagine a green vertical grating. Participants were then presented with the binocularly rivalrous stimuli for 0.75 seconds, after which they would report what they perceived using key presses. The researchers found that imagining a specific pattern would strongly bias participants' subsequent perception of the pattern during binocular rivalry, meaning if the participant imagined a green vertical grating prior to being shown the binocularly rivalrous stimuli, they were more likely to report seeing the green vertical grating than if they had not imagined anything or than if they had imagined the red horizontal grating. Other objective methods of investigating mental imagery include the pupillary light response task (Kay et al., 2022) and fMRI decoding approaches (Albers et al., 2013; Cui et al., 2007; Harrison & Tong, 2009; Koenig-Robert & Pearson, 2019; Naselaris et al., 2015; Stokes et al., 2009).

Predictive coding provides a possible account for why binocular rivalry occurs. In Hohwy et al. (2008), the authors break binocular rivalry down to two key problems: selection and alternation. The problem of selection asks which stimulus dominates and why a fusion of the two rivaling stimuli does not occur, while the problem of alternation asks why the stimuli appear to switch from one to the other and back. For the selection problem, say a face and a

SPATIAL FREQUENCY, FILTER CONDITION, IMAGERY VIVIDNESS

house are rivaling. The possible hypotheses in this scenario are 1) it is a face + house, 2) it is a face, or 3) it is a house. The first hypothesis has a high likelihood because it can predict the sensory evidence to a large extent, but it has a very low prior because it is unlikely that causes in the environment will cause a percept that is the fusion of a face and house. On the other hand, the second and third hypotheses have lower likelihoods because neither one can fully explain the sensory evidence, but both have relatively high priors because of our past experience with faces and houses. Therefore, no hypothesis enjoys both high prior and high likelihood probability, thus the hypothesis with the highest prior wins and results in the highest posterior. If the face has a higher prior, then given equal likelihood, it will dominate. For the alternation problem, when a hypothesis explains the sensory data enough for a temporarily stable percept, there is high-level inhibition of other hypotheses. However, there remains strong error signals from the suppressed percept that are unexplained, while the sensory signals from the dominant percept are explained away by the dominant hypothesis. The error signals are transmitted to higher-level areas through feedforward connections, eventually inducing a switch in the dominant hypothesis since the brain works to minimize prediction error. This results in instability in perceptual dynamics and is why alternation occurs, according to the predictive coding account.

Theoretically, how does mental imagery influence binocular rivalry under predictive coding? Considering that feedback connections are how predictions are encoded (Rao & Ballard, 1999), and that imagery and perception have been shown to share neural mechanisms (Ishai et al., 2000; Dijkstra et al. 2019; Dijkstra et al., 2020), mental imagery of a stimulus should work to increase the prior of that stimulus, leading to a higher posterior and thereby increasing the perceptual dominance of a stimulus during rivalry. According to this account, in the binocular rivalry paradigm of Pearson et al. (2008), the top-down bias effect of mental imagery may be due to the hypothesis of the stimulus being imagined having an

SPATIAL FREQUENCY, FILTER CONDITION, IMAGERY VIVIDNESS

even higher prior than during normal perception. Thus, that hypothesis enjoys both a relatively high prior and a high likelihood, which results in the dominance of the imagined stimulus.

Current study

In order to test the aforementioned predictions about the nature of mental images, the current study used a binocular rivalry paradigm similar to Pearson et al. (2008) to investigate the effects of spatial frequency and stimulus complexity on the vividness of visual imagery. Previous studies have examined the nature of mental images and the influence of stimulus properties on imagery vividness (Finke & Kosslyn, 1980; Kunen & May, 1980; Finke & Kurtzman, 1981a; Finke & Kurtzman, 1981b; Dijkstra et al., 2017), but they generally relied on subjective reports of vividness and are susceptible to personal biases. Subjectivity is greatly reduced (though not completely eliminated) with the binocular rivalry technique, hence it was the investigative method of choice in this study. Additionally, previous studies have not examined the nature of mental images under a predictive coding framework, which is a powerful account of visual perception, mental imagery, and brain function overall, as illustrated above. We formalized the predictions about the vividness of mental images into two hypotheses and tested them in two experiments. We expected that imagery strength would decrease with higher spatial frequencies, since imagery representations have a lower level of detail as a result of the larger receptive fields of neurons in higher-level visual areas, which are responsible for generating these representations. We also expected that imagery strength would be higher for low-pass filtered stimuli than non-filtered or high-pass filtered stimuli, since low-pass filtered stimuli can better simulate the blurry phenomenological experience of visual imagery. We chose Gabors and radial checkerboards as simple stimuli and faces as complex stimuli because we wanted to capture the dynamics of visual imagery across stimulus complexity.

Methods

Participants

A total of 16 volunteers (9 female, 7 male; aged 20–30) participated in this experiment. Participants included students or employees at Vanderbilt University recruited through word of mouth. They all had normal or corrected-to-normal vision and gave written informed consent before participating in the experiment. They were monetarily compensated for their time (\$15/hour). The study was carried out with approval from the Vanderbilt University Institutional Review Board. The data was collected over five days.

Design

Two experiments were completed by all participants, and both employed a 3×2 factorial design. The first experiment varied spatial frequency along three levels (low, medium, and high) and imagined stimulus along two levels (green vertically oriented Gabor and red horizontally oriented Gabor). The second experiment varied filtering along three levels (no filter, low-pass filter, and high-pass filter) and imagined stimulus along two levels (green face and red radial checkerboard). All factors were manipulated within-subject. The dependent variable of both experiments was imagery strength, which was defined as the percentage of trials where the imagined stimulus matched the dominant stimulus during binocular rivalry. Imagery strength (% primed) was calculated using the equation $n_p / (n - n_x) \times 100$, with n_p being the number of trials where the imagined and dominant stimuli match, n being the total number of trials, and n_x being the number of trials where a mixed percept was reported (Bergmann et al., 2016). For each experiment, participants completed 120 trials, which was split into 6 blocks of 20 trials each, with short breaks in between each block to minimize fatigue. The order of conditions was randomized, and each participant completed one block of every condition (spatial frequency \times imagined stimulus, filtering \times imagined stimulus).

Materials

After participants gave informed consent but before they started the experimental trials, they were given the Vividness of Visual Imagery Questionnaire (Marks, 1973), or VVIQ. It consists of 16 items grouped into four settings, and participants were asked to think about specific scenarios that conjure up a mental image. They were then asked to rate the vividness of the image on a five-point scale, with one being the most vivid and five being the least vivid. Though the VVIQ is a subjective measure of imagery vividness, and thus prone to participant self-bias, studies have concluded that the VVIQ has high reliability (Campos & Pérez-Fabello, 2009; Jankowska & Karwowski, 2022) and can predict individual variations in imagery strength as measured using the binocular rivalry method (Pearson et al., 2011).

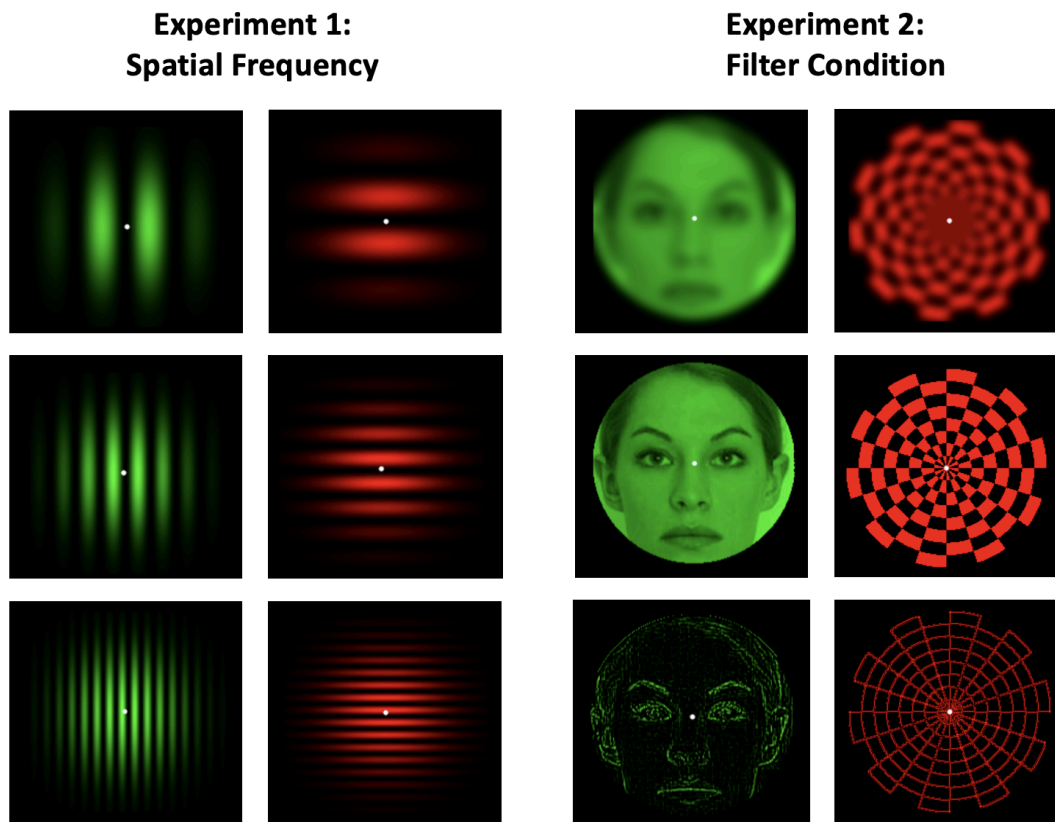
Participants completed all trials in a darkened room, using an individually calibrated mirror stereoscope to facilitate binocular rivalry. Their head position was stabilized at 50 cm from the screen using a chin rest. The stimuli were presented on a CRT monitor (Dell E773C; resolution: 1280 x 1024 pixels; refresh rate: 60 Hz; screen dimensions: 40 cm x 30 cm) and created using MATLAB R2019a (The MathWorks) using the Psychophysics Toolbox extension (Brainard, 1996; Pelli, 1997; Kleiner et al., 2007) running on MacOS X, version 10.11.6. Fixation disparity was eliminated by physically adjusting the tilt of the mirrors to line up a display consisting of nonius lines, namely a cross at the center of each visual field, a medial vertical line below the left cross, and a medial vertical line above the right cross. The fusion of the crosses and lines into a unified percept indicated successful alignment.

In both experiments, the screen background was set to the lowest luminance level throughout to make sure visual imagery had minimal interference from incoming visual signals. The stimuli used in the first experiment were orthogonally oriented green and red Gabor patches. The green vertically oriented Gabor was presented to the left eye and the red horizontally oriented Gabor was presented to the right eye. The spatial frequency conditions

SPATIAL FREQUENCY, FILTER CONDITION, IMAGERY VIVIDNESS

were quantified as follows: low – 0.75 cycles per degree (cpd); medium – 1.5 cpd; high – 3 cpd. Both Gabors spanned 4° in visual angle and were each centered on a white fixation dot covering 0.8°. The spatial phase of the Gabors were randomized for each block. Mean luminance starting values were 7.8 cd/m² for both the green and red Gabors, which were then individually adjusted for each participant to compensate for eye dominance and green/red equiluminance (see Procedure section).

The stimuli used in the second experiment were faces and radial checkerboards. An image of a face with neutral expression was found on Google Image Search and cropped and processed into a greyscale image, which was then shown through the green channel as a green-tinted face. The checkerboard stimulus was coded in MATLAB and consisted of alternating red and black patches radiating out from a center point. The face was presented to the left eye and the checkerboard to the right eye. Both the face and the checkerboard spanned 4° in visual angle and were centered on a white fixation dot covering 0.8°. Similar to the first experiment, green and red luminance values were individually adjusted for each participant.

Figure 1*Stimuli used in Experiments 1 & 2*

Note: Low, medium, and high spatial frequency Gabors were shown in Experiment 1.

Low-pass filtered, non-filtered, and high-pass filtered faces and radial checkerboards were shown in Experiment 2. The green stimuli were always presented to the left eye, while the red stimuli were always presented to the right eye.

Procedure*Adjustment and calibration*

Participants gave written informed consent, filled in the VVIQ, and their total score was recorded. They were then shown the experimental setup and completed the mirror alignment of the mirror stereoscope using the nonius lines display. After that, they completed an adaptive procedure previously described by Pearson et al. (2008) to adjust the relative strength of the contrast and luminance of the stimuli. This was done to control for individual

SPATIAL FREQUENCY, FILTER CONDITION, IMAGERY VIVIDNESS

differences in perceptual bias that could result from eye dominance and/or green/red equiluminance (perceiving either color as brighter than the other even when both have identical color intensity values). Peak luminance was adjusted by first showing the participant the rivalry display briefly (0.75 s) at a given contrast, then having them indicate the dominant stimulus using key presses. The participant was then shown the dominant stimulus at full contrast for 4 seconds, which leads to adaptation and a greater likelihood of perception switching to the other stimulus during the next rivalry presentation (Pearson and Clifford, 2005). If the same stimulus was dominant twice in a row, the contrast level of the dominant stimulus was reduced and the contrast level of the suppressed stimulus was increased. This procedure was iterated until the intervening full contrast stimulus could induce a perceptual switch in most or all rivalry presentations. After completing the adaptive procedure for peak luminance, the perceptual strength of the two stimuli would be equal, and we would have greater confidence that perceptual bias was due mainly to the top-down effect of imagery instead of the bottom-up effect of stimulus properties.

Experiment 1

For the first experiment, participants were instructed to imagine either a green vertical Gabor patch by the cue “G” or a red horizontal Gabor patch by the cue “R”. The letter cue was shown for 0.75 seconds. Then, a 7-second imagery period was presented, consisting of only a white central fixation dot. The participant was instructed to fixate on the dot while imagining the indicated Gabor around the dot. After that, the stimuli were shown for 0.75 seconds, one to each eye. Participants were then shown a question mark for 1.5 seconds, cueing them to indicate which percept they saw using key presses. Using their right hand, they pressed the “1” key on the numeric keypad if they saw mainly the green vertical Gabor, the “3” key if they mainly saw the red horizontal Gabor, and the “2” key if they saw an

SPATIAL FREQUENCY, FILTER CONDITION, IMAGERY VIVIDNESS

approximately equal mixture of the two stimuli. If mixed stimuli were indicated on more than 20% of the trials in a block, the block was excluded from analysis.

Experiment 2

For the second experiment, participants were instructed to imagine either a face by the cue “F” or a radial checkerboard by the cue “C”. The timeline of each trial was identical to the first experiment. Participants pressed “1” if they mainly saw the green face, “3” if they mainly saw the red checkerboard, and “2” if they saw an approximately equal mixture of both stimuli. At the end of the experiment, participants were debriefed and given monetary compensation for their time.

Data processing and analysis

Participant response data was cleaned and processed within MATLAB. Blocks with more than 20% mixed responses were excluded. Referring to the equation used to calculate imagery strength as percent primed, $n_p / (n - n_m - n_x) \times 100$, n_p was calculated as the number of trials where the letter cue matched the key press. For example, if on one trial the letter cue was an “R” and the participant indicated “3” (they saw mostly the red horizontal Gabor), then that trial would be counted towards n_p . On the other hand, if on another trial the letter cue was a “R” and the participant indicated “1” (they saw mostly the green vertical Gabor), then that trial would not be counted towards n_p .

After data processing, overall imagery strength of each participant was tested for correlation with their VVIQ scores for both experiments. A one-way ANOVA was done to compare the mean imagery strength of the low, medium, and high spatial frequency Gabors, and a one-way ANOVA was also done to compare the mean imagery strength of the low-pass, high-pass, and unfiltered stimuli. Post hoc comparisons were performed using the Tukey HSD test for the ANOVA results. In both experiments, one-sample t-tests were conducted to compare the mean imagery strength of each spatial frequency and filter

condition against chance-level imagery bias. Also, one-sample t-tests were performed to compare the mean imagery strength of each stimulus type against chance.

Results

We hypothesized that imagery strength would decrease as spatial frequency increased due to the larger receptive fields of the higher-order areas that are presumably responsible for generating mental images. We also hypothesized that imagery strength of low-pass filtered stimuli would be greater than non-filtered stimuli, which would have greater imagery strength compared to high-pass filtered stimuli.

We first investigated the mean imagery strength across participants for spatial frequencies (Experiment 1) and filter conditions (Experiment 2). After filtering out trial blocks with greater than 20% mixed trials reported, four participants' data were removed from the final analysis as more than half of their trial blocks met the exclusion criteria. 12 participants' data were analyzed, where a one-way ANOVA was done to compare the effect of spatial frequency on mean imagery strength for low, medium, and high spatial frequency Gabors (Figure 2A). No significant differences were found between mean imagery strength at low ($M = 46.38$, $SD = 11.31$), medium ($M = 50.36$, $SD = 21.34$), or high spatial frequency ($M = 51.37$, $SD = 19.08$); $F(2, 33) = 0.26$, $p = .77$. Post-hoc Tukey HSD tests were not run due to these non-significant results. Therefore, the hypothesis that imagery strength decreases as spatial frequency increases was not supported, nor was there any consistent effect of imagery at all, which will be discussed below.

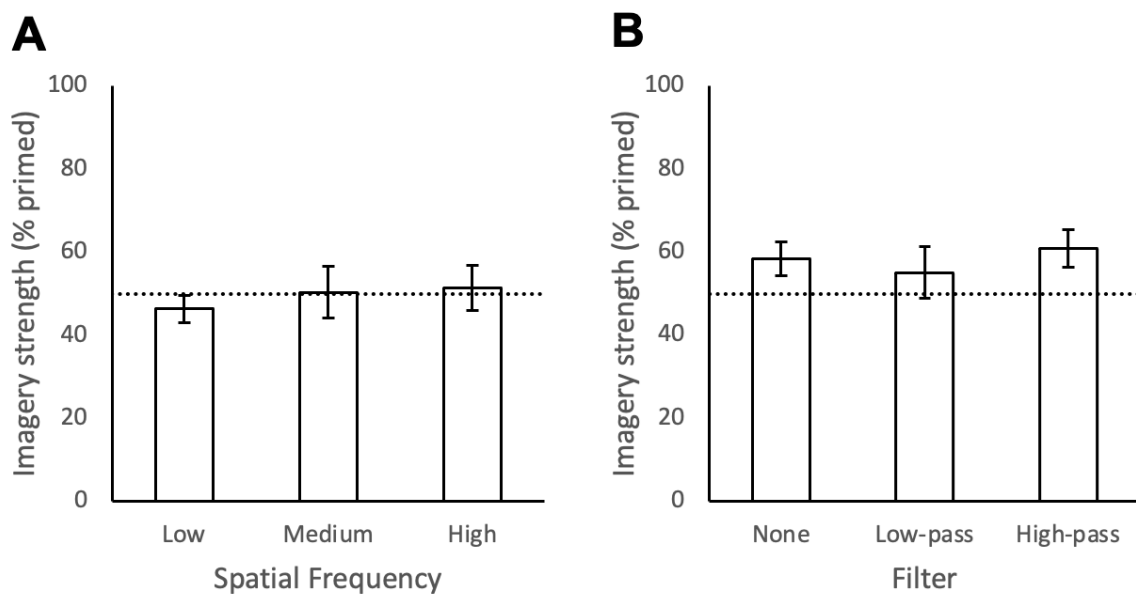
A one-way ANOVA was performed to compare the effect of filter condition on imagery strength in non-filtered, low-pass filtered, and high-pass filtered face and checkerboard stimuli (Figure 2B). No significant differences were found in imagery strength between non-filtered ($M = 58.39$, $SD = 14.06$), low-pass filtered ($M = 55.05$, $SD = 21.63$),

SPATIAL FREQUENCY, FILTER CONDITION, IMAGERY VIVIDNESS

and high-pass filtered stimuli ($M = 60.92$, $SD = 15.51$); $F(2, 33) = 0.34$, $p = .71$. Post-hoc Tukey HSD tests were not run due to these non-significant results. Therefore, the hypothesis that imagery strength increases from high-pass filtering to no filtering to low-pass filtering was not supported.

Figure 2

Mean imagery strength of spatial frequencies and filter conditions



Note: A: Mean imagery strength (% primed) of low, medium, and high spatial frequency Gabor stimuli. B: Mean imagery strength (% primed) of non-filtered, low-pass filtered, and high-pass filtered face and radial checkerboard stimuli. Error bars represent standard error of the mean. Dotted line represents chance-level bias at 50%.

Additionally, the mean imagery strength of all conditions were compared against chance-level bias (50%) using one-sample t-tests. In Experiment 1, no significant differences were found between mean imagery strength for all spatial frequencies and chance (Table 1A). In Experiment 2, no significant difference was found between mean imagery strength for low-pass filtering and chance (Table 1B). The non-filtered condition was marginally

SPATIAL FREQUENCY, FILTER CONDITION, IMAGERY VIVIDNESS

significantly different from chance, and the high-pass filter condition was significantly different from chance (Table 1B). This shows that imagery significantly biased subsequent perception during rivalry only for the high-pass filter condition, and slightly biased perception for the non-filtered condition, but not for any of the other conditions.

Table 1A

Mean imagery strength of all spatial frequency conditions vs chance (50%)

Condition	Mean (%)	SD	df	<i>t</i> statistic	<i>p</i> -value
Low spatial frequency	46.38	11.31	11	1.11	.29
Medium spatial frequency	50.36	21.34	11	0.06	.95
High spatial frequency	51.37	19.08	11	0.25	.81

Note: One-sample *t*-test results of mean imagery strength of all spatial frequency conditions vs chance (50%).

Table 1B

Mean imagery strength of all filtering conditions vs chance (50%).

Condition	Mean (%)	SD	df	<i>t</i> statistic	<i>p</i> -value
Non-filtered	58.39	14.06	11	2.07	.06
Low-pass filtered	55.05	21.63	11	0.81	.44
High-pass filtered	60.92	15.51	11	2.44	.03

Note: One-sample *t*-test results of mean imagery strength of all filtering conditions vs chance (50%). High-pass filtered condition is significantly different from chance, and non-filtered condition is marginally significant from chance (*p*-values in bold).

We also explored whether imagery of each type of stimuli would have biased subsequent rivalry perception, similarly using one-sample *t*-tests. Chance level was 25% because during any given trial, the probability of seeing a particular stimulus was 50%, and the probability of the perceived stimulus matching the imagined stimulus was 50%. Hence, $50\% \times 50\% = 25\%$. An extremely significant difference was found in mean imagery strength between the checkerboard stimulus and chance (Table 2). No other stimulus type showed significant differences with chance (Table 2). This demonstrates that imagery only

SPATIAL FREQUENCY, FILTER CONDITION, IMAGERY VIVIDNESS

significantly biased subsequent rivalry perception for the checkerboard stimuli and not for any other stimulus type.

Table 2

Mean imagery strength of all stimulus types vs chance (25%)

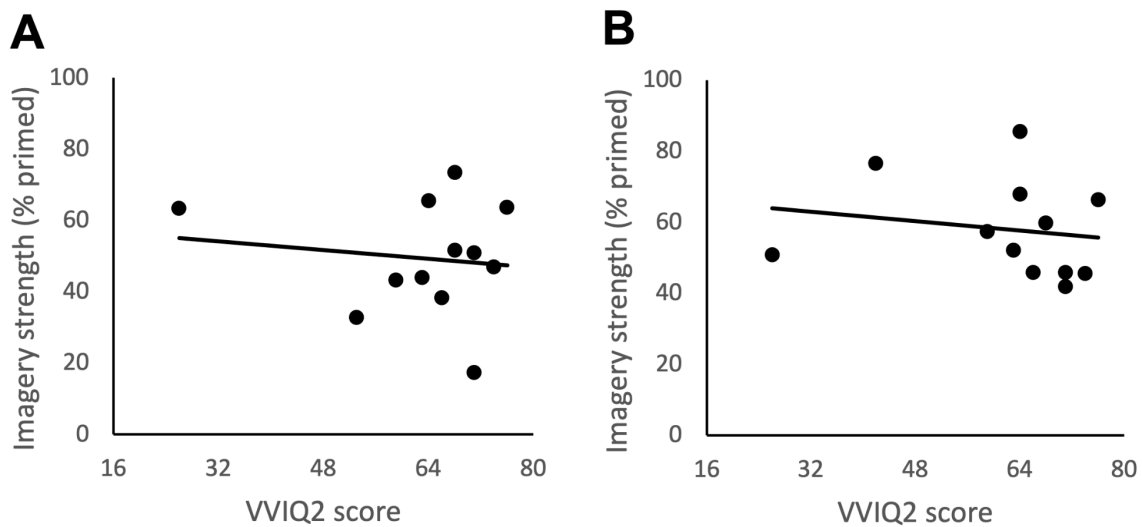
Stimulus type	Mean (%)	SD	df	<i>t</i> statistic	<i>p</i> -value
Green Gabors	28.24	14.80	35	1.31	.20
Red Gabors	21.13	16.36	35	1.42	.17
Face	23.31	13.11	35	0.78	.44
Checkerboard	34.81	12.87	35	4.57	<i>p</i> < 0.001

Note: One-sample *t*-test results of mean imagery strength of all stimulus types vs chance (25%). Imagery strength of checkerboards was significantly greater than chance (*p*-value in bold).

Lastly, we investigated whether participants' VVIQ2 scores would predict the degree to which imagery biased subsequent perception, that is, whether the subjective report of imagery vividness would be positively correlated with the relatively objective measure of imagery strength. The imagery strength of each participant was plotted against their VVIQ2 score (Figure 3). In both experiments, there was no significant correlation between imagery strength and VVIQ2 score at the significance level of $p < .05$; Experiment 1: $r(11) = -0.13$, $p = .69$; Experiment 2: $r(11) = -0.17$, $p = .60$.

Figure 3

Individual participant imagery strength vs VVIQ2 score in Experiment 1 and 2



Note: A: Imagery strength (% primed) vs VVIQ2 score in Experiment 1. B: Imagery strength (% primed) vs VVIQ2 score in Experiment 2. Solid line indicates line of best linear fit.

Discussion

In this study, we investigated several hypotheses regarding the imagery strength of various stimuli across two experiments. We investigated whether mean imagery strength would decrease with increasing spatial frequency, the reason being that the larger receptive fields of the higher-order areas responsible for generating mental images would not afford the necessary resolution for higher spatial frequency images. We found no evidence to support this hypothesis, as there were no significant differences between mean imagery strength for the low, medium, and high spatial frequency categories of Gabor stimuli. Therefore, we cannot make any definite conclusions about the validity of the top-down model of mental image generation. Another hypothesis was that mean imagery strength would increase from high-pass filtering to no filtering to low-pass filtering, since the phenomenological experience of visual imagery is often described as blurry and less crisp than normal vision, which is what

SPATIAL FREQUENCY, FILTER CONDITION, IMAGERY VIVIDNESS

the low-pass filtering was intended to imitate. We found no significant differences between the filtering conditions; therefore, the hypothesis is not supported.

The mean imagery strength of all spatial frequency and filtering conditions were compared against chance level to see if imagery successfully biased subsequent perception during rivalry. The non-filtered and high-pass filtered conditions showed marginally significant differences in comparison to chance, implying that the low-pass filtered condition (supposedly the most “realistic”) was surprisingly the least vividly imagined, thus biasing subsequent rivalry perception the least out of the three filter conditions. This could be due to the consistent perceptual dominance of one stimulus over the other. Some participants reported seeing the red radial checkerboard in several consecutive trials, despite the letter cue and the imagined percept being the green face. This phenomenon can be attributed to calibration only being done on the medium spatial frequency and the non-filtered conditions, which was done because we assumed that the balance of perceptual strength of the two stimuli would remain the same across conditions. The lack of calibration for the low-pass and high-pass filter conditions, as well as for the low and high spatial frequency conditions, may have contributed to improper stimulus strength balancing, and thus the lack of significant findings. In future investigations, all stimulus conditions should be calibrated to ensure proper balancing; only then can the biasing effect of imagery be isolated.

Additionally, the mean imagery strength for each type of stimulus was compared against chance level, and the radial checkerboard surprisingly showed an extremely significant difference. As mentioned above, this could be due to the consistent perceptual dominance of the checkerboard stimuli as caused by lack of calibration for two of the three filter conditions. Also, the regular repeating pattern of the checkerboard rivaling against the non-regular-patterned face could have caused piecemeal rivalry, i.e., small patches of the checkerboard breaking through the percept of the face, leading participants to report seeing

SPATIAL FREQUENCY, FILTER CONDITION, IMAGERY VIVIDNESS

the checkerboard much more often than the face. In future investigations, perhaps a different control stimulus could be used, for example a house or a car, to decrease the chances of piecemeal rivalry occurring.

Finally, we examined the correlation between participants' VVIQ2 scores and their imagery strength (percent of trials primed), and found that for both experiments, there was no significant correlation between the former, subjective measure and the latter, more objective measure of imagery vividness. This finding is in line with Dijkstra et al. (2019), who also found no significant correlation between VVIQ2 scores and imagery priming. In contrast, Pearson et al. (2011) found a significant positive correlation between these two measures. Thus, more evidence is needed to establish whether this positive correlation is robust. In our sample, we had individual participants that violated the expected correlation in two ways: 1) they had a low VVIQ2 score but relatively high imagery strength, or 2) they had a high VVIQ2 score but low imagery strength (Figure 1). These violations could point to either, or both measures being unreliable in accurately measuring one's vividness or strength of visual imagery. However, given that Pearson et al. (2011) found a significant positive correlation between VVIQ2 score and imagery strength, the violations in the present study could simply be due to experimental errors and flaws, which will be discussed below.

There are several limitations to the current study in addition to the aforementioned issues with calibration and piecemeal rivalry. The sample size was small at 16 participants, especially after elimination due to mixed percept trials exceeding 20%, reducing the sample to 12 participants. Though there was a wide range in self-reported and measured imagery strength, the current study would have benefited from a greater number of participants to confirm whether the non-significant findings were due to wrong hypotheses or methodological errors. The calibration procedure could have been more precise, since a few participants reported vertical misalignment of the stimuli during calibration; however, the

SPATIAL FREQUENCY, FILTER CONDITION, IMAGERY VIVIDNESS

misalignments were mostly resolved once the actual experiment began as participants became more familiar with binocular rivalry and adapted to this novel perceptual experience. In future experiments, calibration should involve both vertical and horizontal micro-adjustments using key presses to achieve perfect alignment. Another limitation of the current study was the lack of manually mixed stimuli (simulating piecemeal rivalry) to check for demand characteristics. This manipulation has been done in many of Pearson's studies (e.g., Pearson et al., 2008; Pearson et al., 2011) and involves 10% of trials showing faux-piecemeal rivalry stimuli, to which participants should press "2" to indicate a mixed percept in response. Failure to do so would indicate demand characteristics or misunderstanding of the instructions, and the participant's data would be excluded from analysis. Inclusion of this manipulation in future experiments would improve the quality of the data collected and ensure proper isolation of the biasing effect of imagery on subsequent perception in rivalry.

In conclusion, the current study found no significant changes of mean imagery strength as functions of spatial frequency and filter condition. The mean imagery strength of the non-filtered and high-pass filtered stimuli were barely significantly different from chance, while the mean imagery strength of the checkerboard stimuli were extremely significantly different from chance. No significant correlation was found between imagery strength and VVIQ2 scores. These results do not support our original hypotheses and clearly reflect the flaws and limitations of the current study. Taken together, these results are not conclusive enough to provide evidence in support of the predictive coding model of visual imagery and perception. Nevertheless, the current study lays the foundation for future studies of the nature of mental imagery and the mechanisms that support it through careful manipulation of stimuli characteristics. In addition to spatial frequency and stimulus complexity, future investigations could examine the interaction between factors like contrast level, stimulus complexity, and even participant factors like aphantasia or hyperphantasia. This will allow us to empirically

SPATIAL FREQUENCY, FILTER CONDITION, IMAGERY VIVIDNESS

test hypotheses that a predictive coding model of perception would generate with regards to mental imagery, from which we can gain a deeper understanding of the brain mechanisms that support imagery and perception.

References

- Albers, A. M., Kok, P., Toni, I., Dijkerman, H. C., & de Lange, F. P. (2013). Shared representations for working memory and mental imagery in early visual cortex. *Curr. Biol.* 23, 1427–1431. <https://doi.org/10.1016/j.cub.2013.05.065>
- Bergmann, J., Genç, E., Kohler, A., Singer, W., & Pearson, J. (2015). Smaller Primary Visual Cortex Is Associated with Stronger, but Less Precise Mental Imagery. *Cerebral Cortex*, 26(9), 3838–3850. <https://doi.org/10.1093/cercor/bhv186>
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10(4), 433–436. <https://doi.org/10.1163/156856897x00357>
- Campos, A., & Pérez-Fabello, M. J. (2009). Psychometric quality of a revised version vividness of Visual Imagery Questionnaire. *Perceptual and Motor Skills*, 108(3), 798–802. <https://doi.org/10.2466/PMS.108.3.798-802>
- Cui, X., Jeter, C. B., Yang, D., Montague, P. R., & Eagleman, D. M. (2007). Vividness of mental imagery: Individual variability can be measured objectively. *Vision Research*, 47(4), 474–478. <https://doi.org/10.1016/j.visres.2006.11.013>
- Dijkstra, N., Ambrogioni, L., Vidaurre, D., & van Gerven, M. (2020). Neural dynamics of perceptual inference and its reversal during imagery. *ELife*, 9, e53588. <https://doi.org/10.7554/eLife.53588>
- Dijkstra, N., Bosch, S. E., & van Gerven, M. A. (2017). Vividness of Visual Imagery Depends on the Neural Overlap with Perception in Visual Areas. *Journal of Neuroscience*, 37(5), 1367–1373. <https://doi.org/10.1523/JNEUROSCI.3022-16.2016>
- Dijkstra, N., Bosch, S. E., & Gerven, M. A. J. van. (2019). Shared Neural Mechanisms of Visual Perception and Imagery. *Trends in Cognitive Sciences*, 23(5), 423–434. <https://doi.org/10.1016/j.tics.2019.02.004>

SPATIAL FREQUENCY, FILTER CONDITION, IMAGERY VIVIDNESS

- Dijkstra, N., Hinne, M., Bosch, S. E., & van Gerven, M. A. J. (2019). Between-subject variability in the influence of mental imagery on conscious perception. *Scientific Reports*, 9(1). <https://doi.org/10.1038/s41598-019-52072-1>
- Finke, R. A., & Kosslyn, S. M. (1980). Mental imagery acuity in the peripheral visual field. *Journal of Experimental Psychology: Human Perception and Performance*, 6(1), 126–139. <https://doi.org/10.1037/0096-1523.6.1.126>
- Finke, R. A., & Kurtzman, H. S. (1981a). Area and contrast effects upon perceptual and imagery acuity. *Journal of experimental psychology. Human perception and performance*, 7(4), 825–832.
- Finke, R. A., & Kurtzman, H. S. (1981b). Mapping the visual field in mental imagery. *Journal of Experimental Psychology: General*, 110(4), 501–517. <https://doi.org/10.1037/0096-3445.110.4.501>
- Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360(1456), 815–836. <https://doi.org/10.1098/rstb.2005.1622>
- Hohwy, J., Roepstorff, A., & Friston, K. (2008). Predictive coding explains binocular rivalry: An epistemological review. *Cognition*, 108(3), 687–701. <https://doi.org/10.1016/j.cognition.2008.05.010>
- Harrison, S. A., & Tong, F. (2009). Decoding reveals the contents of visual working memory in early visual areas. *Nature*, 458(7238), 632–635. <https://doi.org/10.1038/nature07832>
- Ishai, A., & Sagi, D. (1995). Common mechanisms of visual imagery and perception. *Science*, 268(5218), 1772–1774. <https://doi.org/10.1126/science.7792605>

SPATIAL FREQUENCY, FILTER CONDITION, IMAGERY VIVIDNESS

Ishai, A., Ungerleider, L. G., & Haxby, J. V. (2000). Distributed neural systems for the generation of visual images. *Neuron*, 28(3), 979–990.

[https://doi.org/10.1016/s0896-6273\(00\)00168-9](https://doi.org/10.1016/s0896-6273(00)00168-9)

Jankowska, D. M., & Karwowski, M. (2022). How vivid is your mental imagery? Validation of the vividness of Visual Imagery Questionnaire in five Polish samples. *European Journal of Psychological Assessment*. Advance online publication.

<https://doi.org/10.1027/1015-5759/a000721>

Kay, L., Keogh, R., Andrillon, T., & Pearson, J. (2022). The pupillary light response as a physiological index of aphantasia, sensory and phenomenological imagery strength.

ELife, 11. <https://doi.org/10.7554/elife.72484>

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in psychtoolbox-3. *Perception*, 36(14), 1-16.

Koenig-Robert, R. & Pearson, J. (2019). Decoding the contents and strength of imagery before volitional engagement. *Sci. Rep.* 9, 3504.

<https://doi.org/10.1038/s41598-019-39813-y>

Koenig-Robert, R., & Pearson, J. (2020). Why do imagery and perception look and feel so different? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 376(1817), 20190703. <https://doi.org/10.1098/rstb.2019.0703>

Kosslyn, S. M., Thompson, W. L., & Alpert, N. M. (1997). Neural systems shared by visual imagery and visual perception: a positron emission tomography study. *NeuroImage*, 6(4), 320–334. <https://doi.org/10.1006/nimg.1997.0295>

Kunen, S., & May, J. G. (1980). Spatial frequency content of visual imagery. *Perception & Psychophysics*, 28(6), 555–559. <https://doi.org/10.3758/bf03198825>

SPATIAL FREQUENCY, FILTER CONDITION, IMAGERY VIVIDNESS

- Lee, T. S., & Mumford, D. (2003). Hierarchical Bayesian inference in the visual cortex. *Journal of the Optical Society of America. A, Optics, image science, and vision*, 20(7), 1434–1448. <https://doi.org/10.1364/josaa.20.001434>
- Marks, D. F. (1973). Visual imagery differences in the recall of pictures. *British Journal of Psychology*, 64(1), 17–24. <https://doi.org/10.1111/j.2044-8295.1973.tb01322.x>
- Naselaris, T., Olman, C. A., Stansbury, D. E., Ugurbil, K. & Gallant, J. L. (2015). A voxel-wise encoding model for early visual areas decodes mental images of remembered scenes. *Neuroimage*, 105, 215–228. <https://doi.org/10.1016/j.neuroimage.2014.10.018>
- O’Craven, K. M., & Kanwisher, N. (2000). Mental Imagery of Faces and Places Activates Corresponding Stimulus-Specific Brain Regions. *Journal of Cognitive Neuroscience*, 12(6), 1013–1023. <https://doi.org/10.1162/08989290051137549>
- Pearson, J., & Clifford, C. W. G. (2005). Mechanisms selectively engaged in rivalry: normal vision habituates, rivalrous vision primes. *Vision Research*, 45(6), 707–714. <https://doi.org/10.1016/j.visres.2004.09.040>
- Pearson, J., Clifford, C. W. G., & Tong, F. (2008). The Functional Impact of Mental Imagery on Conscious Perception. *Current Biology*, 18(13), 982–986. <https://doi.org/10.1016/j.cub.2008.05.048>
- Pearson, J., Naselaris, T., Holmes, E. A., & Kosslyn, S. M. (2015). Mental Imagery: Functional Mechanisms and Clinical Applications. *Trends in Cognitive Sciences*, 19(10), 590–602. <https://doi.org/10.1016/j.tics.2015.08.003>
- Pearson, J., Rademaker, R. L., & Tong, F. (2011). Evaluating the Mind’s Eye: The Metacognition of Visual Imagery. *Psychological Science*, 22(12), 1535-1542. <https://doi.org/10.1177/0956797611417134>

SPATIAL FREQUENCY, FILTER CONDITION, IMAGERY VIVIDNESS

- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision*, *10*(4), 437–442.
<https://doi.org/10.1163/156856897x00366>
- Rao, R., & Ballard, D. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature neuroscience*, *2*(1), 79–87. <https://doi.org/10.1038/4580>
- Slotnick, S. D., Thompson, W. L., & Kosslyn, S. M. (2005). Visual Mental Imagery Induces Retinotopically Organized Activation of Early Visual Areas. *Cerebral Cortex*, *15*(10), 1570–1583. <https://doi.org/10.1093/cercor/bhi035>
- Smith, A. T., Singh, K. D., Williams, A. L., & Greenlee, M. W. (2001). Estimating Receptive Field Size from fMRI Data in Human Striate and Extrastriate Visual Cortex. *Cerebral Cortex*, *11*(12), 1182–1190. <https://doi.org/10.1093/cercor/11.12.1182>
- Stokes, M., Thompson, R., Cusack, R. & Duncan, J. (2009). Top-down activation of shape-specific population codes in visual cortex during mental imagery. *J. Neurosci.*, *29*, 1565–1572. <https://doi.org/10.1523/JNEUROSCI.4657-08.2009>