Motivated unseeing: expected loss biases perceptual decisions by suppressing visual representations

Matan Mazor1,2, Itay Goetz3, & Clare Press1,2

1 Birkbeck, University of London

2 Wellcome Centre for Human Neuroimaging, UCL

3 University of Bamberg

Author note

Correspondence concerning this article should be addressed to Matan Mazor, Malet Street, London WC1E 7HX. E-mail: [mtnmzor@gmail.com](mailto:mtnmzor@gmail.com)

Abstract

Our perceptual decisions are affected not only but what we perceive, expect and believe, but also by what we want to perceive. Effects of motivated perception have been identified in previous studies, with monetary reward and losses biasing visual discrimination, and affecting neural activity in visual cortex. However, it remains unclear whether these behavioural and neuronal effects are driven by a selective enhancement of the perception of motivation-congruent stimuli, or alternatively, by a selective suppression of motivation-incongruent ones. In 6 pre-registered online studies (N=100x6), we leverage reward/loss and presence/absence asymmetries and find evidence for a selective suppression of undesired visual representations relative to a neutral control condition. Our results shed light on a putative mechanism by which motivation affects perception: the suppression of undesired percepts.

*Keywords:* keywords

*Word count:* X

Motivated unseeing: expected loss biases perceptual decisions by suppressing visual representations

# Introduction

The year is 2021 and the world is experiencing a global pandemic. People routinely get tested for the virus: this is done by placing a sample on a test kit and waiting for the result, with the presence of the virus indicated by a test line emerging beneath the control line. You have a flight later today, so you take a test, really hoping for it to come out negative. You place the sample on the kit and patiently stare at it, waiting for the test line to emerge. Is your perception of the test line altered by your motivation not to see it? Do you selectively unsee the undesired outcome? Or maybe your motivation not to see the line is paradoxically making you more likely to perceive it?

Previous research indicates that what we want can determine what we see, or report to see. In addition to reports of increased sensitivity to reward-associated stimuli (Libera & Chelazzi, 2006; Pessoa & Engelmann, 2010), recent findings further suggest that motivation has a biasing effect on perception (Balcetis, Dunning, & Granot, 2012; Leong, Dziembaj, & D’Esposito, 2021; Leong, Hughes, Wang, & Zaki, 2019). According to these studies, people are more likely to see what they want to see. For example, Leong et al. (2019) presented participants with ambiguous overlays of face and scene images, with the face being more dominant in some trials, and the scene in others. Motivation was manipulated by associating face-dominant and scene-dominant images with a loss or gain of points on different trials. Even when informed that they should respond as accurately as possible, participants were more likely to report perceiving the rewarded stimulus – an effect that has been shown to reflect a combination of increased gain for motivated sensory representations in visual cortex and a non-perceptual decision bias, driven by subcortical mechanisms (Leong et al., 2021, 2019).

These and similar findings reveal considerable influences of motivation on perception, but they leave open the question of whether these effects are driven by the enhancement of desired representations, the suppression of undesired representations, or both. It is difficult to arbitrate between these two mechanisms based on findings in the literature, because in most experimental paradigms the two stimulus categories are mutually exclusive and exhaustive. This makes the suppression of one category hard to dissociate from the enhancement of the alternative category; subjects may appear to enhance motivation-congruent representations, while in fact they suppress the representation of motivation-incongruent ones.

To address these open questions, we conducted six online experiments (N=100x6). In Experiments 1-3, subjects performed two alternative forced choice (2AFC) discrimination tasks. We compared the effects of reward and loss manipulations against a neutral condition, and found marked effects of loss on response bias, with no parallel effects for reward. In Experiments 4-6 we asked whether the effects of loss on decision bias reflected a suppression of the stimulus category associated with loss, or alternatively, an enhancement of the alternative stimulus category. In a detection task, associating stimulus absence or presence with loss resulted in weak and unstable biasing of the decision criterion. Drift diffusion modelling revealed suppression effects of motivation on drift rate for the undesired stimulus category, again in line with a motivated unseeing account.

# Experiment 1

## Methods

We report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in the study.

### Participants.

The study was approved by the Research Ethics Committee of Birkbeck, University of London (study ID number 2122050). Participants were recruited via Prolific, and gave informed consent prior to their participation. The criteria for recruitment were a Prolific approval rate of 95%, English as their first language, and an age between 18 and 60. The study took about 15 minutes, and participants were paid 7.5 pounds per hour. The top-performing 30% of participants received an additional bonus payment of £1.

### Procedure.

Figure 1 illustrates the experimental design. In a near-threshold categorisation task, participants reported whether they saw an image of a bird or a fish. At the beginning of the experiment, participants were given instructions and their understanding was then checked by a multiple-choice question. Then, participants were presented with a practice block, which was repeated until accuracy on the task reaches 75%. The practice round was followed by the main part of the experiment, comprising six blocks of 28 trials. At the beginning of each trial a fixation cross appeared on the screen for 500-1000 milliseconds, followed by three grayscale images: a forward mask (presented for 50 ms), the target image (presentation time calibrated to achieve 70% accuracy), and a backward mask (presented until a decision is made). All images were from the noisy animals stimulus set curated, processed, and kindly shared by, Meuwese, Loon, Lamme, and Fahrenfort (2014). Participants’ task was to judge whether the second image was of a bird or a fish. They were asked to indicate their decision using the J and F keys on the keyboard, counterbalanced across participants. The next trial did not begin until participants pressed one of these keys to indicate their decision.

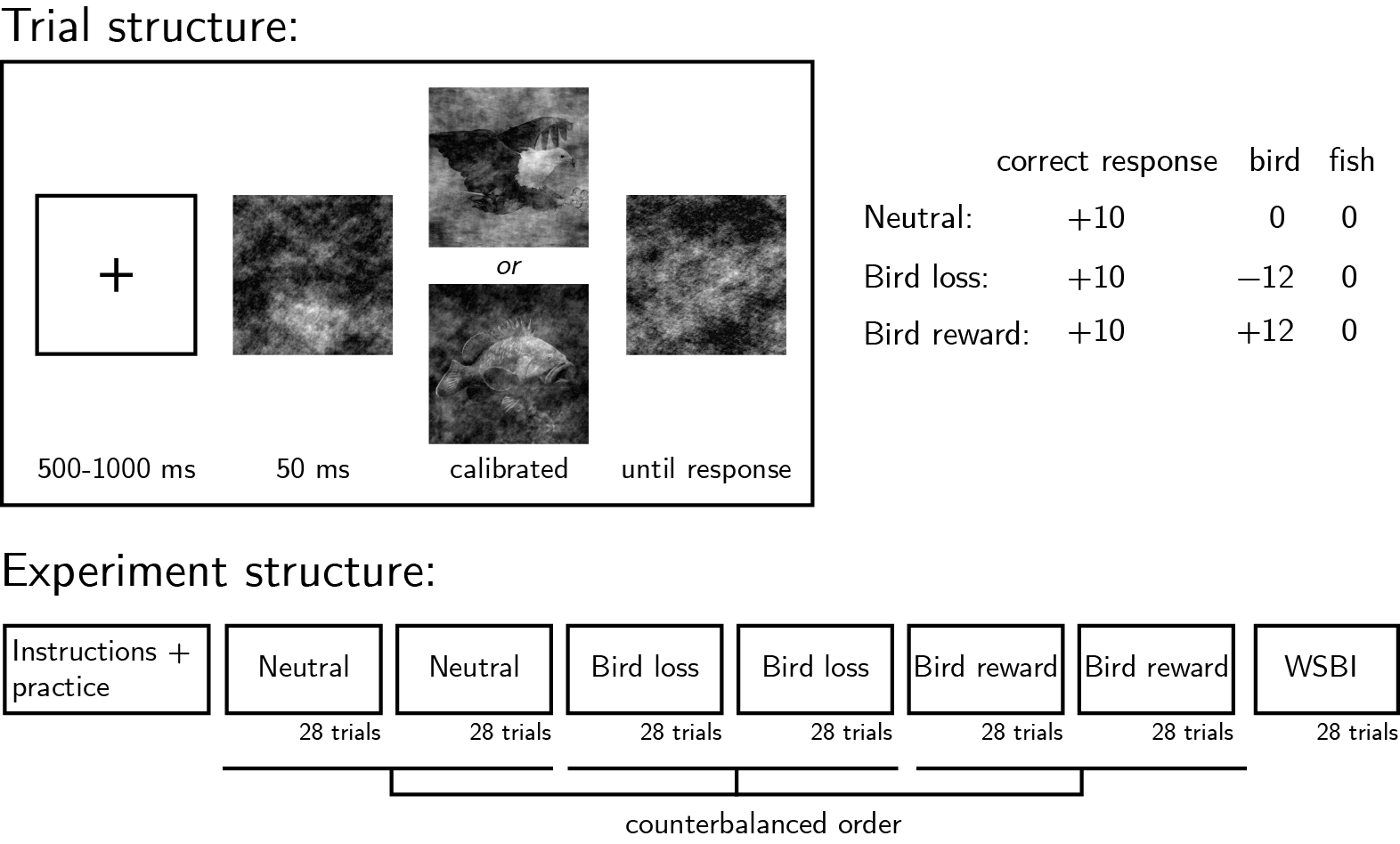
In addition to a base payment, participants were told that they could accumulate points in the task, and that these points will determine whether or not they get a bonus payment. Specifically, that a bonus payment of £1 will be given to the top 30% of highest performing participants.

In all blocks, participants were told that correct responses award them with 10 points. In addition, on two bird-reward blocks, bird images were rewarded with an extra 12 points, and on two bird-loss blocks, bird images were associated with a loss of 12 points. Two neutral blocks included gains or losses for birds. The order of the six blocks were randomized, with the constraint that the two blocks of each condition were always presented consequently.

Importantly, in all blocks participants were motivated to make accurate judgments regarding the content of the presented image, but on some blocks the presence of a bird (regardless of their decision) was associated with a gain of points, while on others the presence of a bird was associated with a loss of points. Two multiple-choice questions ensured participants understood these instructions (for example, “To get many points I need to…” correct answer: “press F when the second image contained a bird and J when it contained a fish”, and not “press F as many times as possible”; “In addition to getting 10 points for accurate responses, in this block whenever a bird image is presented…” correct answer: “I automatically get 12 additional points”, and not “I get 12 additional points only if I pressed F”). Neutral blocks were preceded by the response mapping comprehension question (“To get many points I need to…” correct answer: “press F when the second image contained a bird and J when it contained a fish”). An additional question (“In addition to getting 10 points for accurate responses, in this block whenever a bird image is presented…”, correct response: “I don’t get or lose additional points”) was presented only to those participants for which ‘neutral’ was not the first experimental condition.

Task difficulty was calibrated by adjusting the target presentation time, starting at 60 ms and following a one-up-two-down procedure with a multiplicative step size of 0.9, which moved closer to 1 following each change in the direction of the calibration process. The calibration ran in the background throughout the entire task.

Upon completion of the main part of the experiment, participants completed the white bear suppression inventory (Wegner & Zanakos, 1994).



*Figure* *1.*  Experimental design.Top left: Trial structure. The target image’s presentation time was calibrated to achieve 70% accuracy. Top right: incentive structure on different experimental blocks. Bottom: Overall experiment structure.

#### Randomization.

The order and timing of experimental events was determined pseudo-randomly by the Mersenne Twister pseudorandom number generator, initialized in a way that ensures registration time-locking (Mazor, Mazor, & Mukamel, 2019).

### Rejection criteria.

Participants were excluded if their accuracy fell below 50% in one or more experimental conditions (neutral, bird-loss, bird-reward). We also excluded participants for having extremely fast or slow reaction times in one or more of the tasks (below 100 milliseconds or above 5 seconds in more than 25% of the trials). Trials with response time below 100 milliseconds or above 5 seconds were excluded from the response-time analysis.

## Results

110 participants were recruited from Prolific for Experiment 1.

Median completion time was 17.44 minutes. Mean accuracy was 0.69. Participants reported seeing a bird on 0.53 of the trials. Overall we excluded 10 participants based on our exclusion criteria, leaving 100 participants for the main analysis.

First, we asked whether our reward manipulation had an effect on response bias. The proportion of reports of seeing a bird was similar on bird-reward (0.53) and neutral (0.53) blocks (, 95% CI , , ). In contrast, associating bird stimuli with point loss had a robust effect on response bias: the proportion of reports of seeing a bird on bird-loss blocks (0.50) was significantly lower compared to in neutral blocks (, 95% CI , , ).

Response time analysis revealed no main effects of condition or response, and no interaction between the two (all s <1). Specifically, response times to bird stimuli were unaffected by the reward (, 95% CI , , and loss (, 95% CI , , ) manipulations. Similarly, response times to fish stimuli were unaffected by the reward (, 95% CI , , ) and loss (, 95% CI , , ) manipulations.

# Experiment 2

In Exp. 1, we found that participants had an overall bias to report seeing birds, and that this response bias disappeared when bird images were associated with losing points. However, and in contrast with previous reports of motivated perception (Balcetis et al., 2012; Leong et al., 2021, 2019), we found no effect of associating bird images with a reward on response bias. Since fish images were never associated with the loss of gain of points, we could not tell whether this loss/reward asymmetry reflects an asymmetry between our reward and loss manipulations, such that only the loss manipulation is effective, or alternatively, whether it reflects an asymmetry in the perception of birds and fish (such that fish images are more amenable to enhancement, or bird images more amenable to suppression). To this end, Exp. 2 replicated the results of Exp. 1 with one critical change: the loss or gain of points was associated with fish, rather than bird images.

## Methods

Exp. 2 was identical to Exp. 1 with the sole difference that fish images, instead of bird images, were associated with loss or gain of points.

## Results

101 participants were recruited from Prolific for Experiment 1.

Median completion time was 22.01 minutes. Mean accuracy was 0.70. Participants reported seeing a bird on 0.56 of the trials. Overall, we excluded 5 participants based on our exclusion criteria, leaving 96 participants for the main analysis.

Similar to Exp. 1, here too only the loss manipulation was effective in biasing participants responses relative to a neutral baseline. The proportion of reports of seeing a fish was similar on fish-reward (0.45) and neutral (0.44) blocks (, 95% CI , , ). In contrast, the proportion of reports of seeing a fish on fish-loss blocks (0.40) was significantly lower compared to in neutral blocks (, 95% CI , , ).

In Exp. 2, a repeated measured analysis of variance revealed a significant effect of response (bird or fish; ). The motivation manipulations had no significant effect on response time (), and there was no interaction between response and condition (<1).

Despite the absence of a main effect for condition, we followed our pre-registered reaction time analysis for response times (see [osf.io/nj9es](https://osf.io/nj9es)). This analysis revealed that response times for fish (target) images were significantly slower in fish-reward (median RT: 852 ms) compared to neutral blocks (median RT: 839 ms; , 95% CI , , ; Wilcoxon sign rank test: , ). A smaller difference of 28.02 ms between bird responses in reward and neutral blocks was not significant (, ). The loss manipulation had no effect on fish (, 95% CI , , or bird (, 95% CI , , ) response times.

# Experiment 3

Exp. 2 confirmed that the absence of a reward effect in Exp. 1 was not due to differences between the representation of bids and fish. Instead, in the first two experiments we found no effect of associating bird or fish images with a reward on response bias on decision time.

We considered two candidate reasons for our failure to observe an effect of motivation on perception. First, it may be that a bonus payment of £1 was insufficient to reliably affect participants’ behaviour. For comparison, Leong et al. (2019) offered up to $20 of bonus (67% of the base payment), Leong et al. (2021) offered up to $10 (50% of the base payment) and Balcetis et al. (2012) offered a raffle ticket with a top prize of $100 (1000% of the base payment). Second, we considered the possibility that our stimuli were perceived as emotionally aversive, rendering them more difficult to associate with a positive outcome. Previous studies used face/scene images (Leong et al., 2021, 2019) or numbers and letters (Balcetis et al., 2012) which are perhaps more neutral.

In Exp. 3 we have made several changes to our paradigm, to increase our sensitivity to effects of motivation on perception. First, we increased our bonus payment, and offered to double the payment of the highest scoring 30% of our subjects, and triple the payment of the highest scoring 5%. We also included a leaderboard screen to increase motivation. Second, instead of fish and bird images, we used emotionally neutral noisy grating images as stimuli. Finally, to make sure that subjects fully understand that only the accuracy of their responses affects their bonus payment, we included two additional lures to our comprehension questions.

## Methods

In a near-threshold categorisation task, participants reported whether a noisy grating was tilted to the right (clockwise) or to the left (anticlockwise). At the beginning of each trial a fixation cross appeared on the screen for 500-1000 milliseconds, followed by three grayscale images: a forward mask (presented for 50 ms), the target image (presentation time calibrated to achieve 70% accuracy), and a backward mask (presented until a decision is made). The forward and backward masks comprised black and white concentric circles. The target image was a noisy visual grating: 200x200 pixels, where 90% of the pixels show random grayscale values, and the remaining 10% are sampled from a grating (frequency: 62 pixels, phase: random), tilted 45 degrees clockwise or counterclockwise on different trials. Participants’ task was to judge whether the grating was tilted clockwise or counterclockwise. They were asked to indicate their decision using the J and F keys on the keyboard, counterbalanced across participants. The next trial did not begin until participants pressed one of these keys to indicate their decision.

In addition to a base payment, participants were told that they can accumulate points in the task, and that these points will determine whether or not they get a bonus payment. Specifically, that we will double the payment of the top 30%, and triple the payment of the top 5% participants.

Similar to Exp. 1 and 2, in all blocks participants were be told that correct responses award them with 10 points. In addition, on two reward blocks, a target orientation (randomized between subjects) was rewarded with an extra 12 points, and on two loss blocks, the same target orientation was associated with a loss of 12 points. Two neutral blocks included no gains or losses for the target orientation. The order of the six blocks was randomized, with the constraint that the two blocks of each condition are always be presented consequently.

Task difficulty was calibrated by adjusting the target presentation time, starting at 150 ms and following a one-up-two-down procedure with a multiplicative step size of 0.9, which moved closer to 1 following each change in the direction of the calibration process. The calibration ran in the background throughout the entire task.

Upon completion of the main part of the experiment, participants completed the white bear suppression inventory (Wegner & Zanakos, 1994).

## Results

132 participants were recruited from Prolific for Experiment 3.

Median completion time was 14.48 minutes. Mean accuracy was 0.76. Participants reported seeing the target orientation on 0.49 of the trials. Overall we excluded 32 participants based on our exclusion criteria, leaving 100 participants for the main analysis.

The proportion of reports of seeing the target orientation was similar on reward (0.48) and neutral (0.49) blocks (, 95% CI , , ). In contrast, the proportion of reports of seeing the target orientation on loss blocks (0.46) was significantly lower compared to in neutral blocks (, 95% CI , , ). This is consistent with the results from Exp. 1 and Exp. 2: only loss, but not reward, had an effect on response bias.

A two-way repeated measures analysis of variance revealed a significant effect of motivation condition on response time (), and a marginally significant interaction between condition and response (). This pattern was driven by a general slowing in the loss condition, which was more pronounced for the stimulus associated with a loss. Specifically, response times in loss blocks were slower by 33.48 ms for the stimulus associated with a loss (, ; Wilcoxon sign rank test: , ), and by 25.21 ms for the alternative stimulus (, ; Wilcoxon sign rank test: , ). No significant effects on RT were observed for the reward manipulation.

# Experiment 4

In Experiments 1-3, we replicated previously observed effects of motivation on perceptual decisions (Balcetis et al., 2012; Leong et al., 2021, 2019). Specifically, we found that subjects were less likely to report seeing stimuli that were associated with point loss in a perceptual discrimination task, even when incentivized to report accurately. In contrast, we observed no effects of point gain on perceptual decision making.

In Exp. 4 we focused on the effects of expected loss on perception, and asked whether they operated by means of suppressing the representation of stimuli associated with stimulus loss, or alternatively, by enhancing the representation of the alternative stimulus category. To this end, we utilized the asymmetry between inference about visual presences and absences. While stimulus presence is inferred from perceptual evidence, no evidence can be available to support inferences about stimulus absence (Mazor, 2021; Mazor & Fleming, 2020, 2022; Mazor, Moran, & Fleming, 2021). As such, we reasoned that in perceptual detection decisions between stimulus presence and absence, only representations of presence can be enhanced or suppressed. This means that a shift in response bias following an association between stimulus presence and point loss cannot be driven by the enhancement of a representation of ‘nothing’.

## Methods

Exp. 4 was similar to Exp. 3 with the several modifications.

First, instead of making discrimination judgments based on the orientation of a grating, participants reported whether a vertical grating appeared or not in a noisy stimulus. On half of the trials, the stimulus consisted of random patterns of grayscale dots, and on the other half 20% of the pixels were sampled from an image of a vertical grating. Like in Exp. 1-3, participants indicated their decision using the J and F keys on the keyboard, counterbalanced across participants.

Importantly, throughout the experiment target-absent images were referred to as “the Noise image” and target-present images were referred to as “the Stripes image”. We chose this balanced framing, instead of one where gratings stimuli are described as with or without stripes, in order to make sure any differences between the two motivation manipulations cannot be explained by priming effects for the word “stripes”.

Similar to Exp. 1-3, in all blocks participants were told that correct responses award them with 10 points. In addition, on two presence-loss blocks, the presence of a grating was associated with an immediate loss of 12 points, and on two absence-loss blocks, the absence of a grating was associated with an immediate loss of 12 points. The order of the six blocks was randomized, with the constraint that the two blocks of each condition are always be presented consequently.

In Exp. 4, we changed the structure of comprehension checks to accommodate the fact that all motivation manipulations now involved point loss. Specifically, before presence-loss and absence-loss blocks, subjects answered the question “In addition to getting 10 points for accurate responses, in this block I lose points whenever…” (correct response in presence-loss blocks: “the target image is the Stripes image, and regardless of whether I press J of F.”, and not “the target image is the Stripes image and I press J.”, correct response in absence-loss blocks: “the target image is the Noise image, and regardless of whether I press J of F.”, and not “the target image is the Noise image and I press F.”).

Finally, we included an open-ended question at the end of the experiment, probing subjects about their strategies (“We are interested to know if you had any strategy that you used to maximize your point count. Don’t worry - this will not affect your bonus payment!”).

## Results

123 participants were recruited from Prolific for Experiment 4.

Median completion time was 14.37 minutes. Mean accuracy was 0.79. Participants reported target presence on 0.54 of the trials. Overall we excluded 20 participants based on our exclusion criteria, leaving 103 participants for the main analysis.

To our surprise, neither presence- nor absence-loss manipulations had significant effects on decision bias. The proportion of target-presence reports in neutral blocks (0.55) was similar to both presence-loss (0.53; , 95% CI , , ) and absence-loss blocks (0.55; , 95% CI , , ).

A two-way repeated measured analysis of variance revealed the expected effect of response (target presence or absence) on response time (), driven by target absence responses being slower by 74.83 ms on average. Motivation condition had no significant effect on response time (), and so did the interaction between condition and response ().

## Exploratory analysis: White Bear Suppression Inventory (WBSI)

At the end of the experiment, subjects completed the White Bear Suppression Inventory (WBSI), designed to quantify individual tendencies to suppress unwanted thoughts. Higher WBSI scores correspond to a higher tendency to suppress, or distract oneself away from, one’s thoughts. The mean WBSI score was 51.00 and consistent with the mean scores reported in Wegner and Zanakos (1994) (-) and (Blumberg (2000) (). Relibability was high, with a Cronbach alpha of 0.93.

Subjects’ WBSI scores significantly correlated with the effect of presence-loss on response bias (, 95% CI , , ), such that subjects that scored higher on the WBSI inventory tended to report seeing more gratings when the presence of gratings was associated with point loss. In contrast, WBSI scores did not correlate at all with the effect of absence loss on response bias (, 95% CI , , ).

Following Blumberg (2000), we decomposed the WBSI inventory into three constituent factors: unwanted intrusive thoughts (an example item: ‘I have thoughts that I cannot stop’), thought suppression (‘There are things that I try not to think about’), and self-distraction (‘I often do things to distract myself from my thoughts’). While intrusive thoughts did not significantly correlate with the effects of presence-loss and absence-loss on response bias (, 95% CI ; , 95% CI ), the effect of presence-loss on response bias was significantly correlated with both thought suppression (, 95% CI , , ) and self-distraction (, 95% CI , , ). Participants who reported higher levels of thought suppression and self-distraction were more likely to report seeing a grating when the presence of a grating was associated with point loss, relative to the neutral condition. Thought suppression and self-distraction were not associated with the effects of absence loss on response bias (, 95% CI ; , 95% CI ).

# Experiment 5

In Exp. 4, decision bias was not significantly affected by associating target presence or absence with point loss. This was surprising, given the findings from Exp. 1-3, where subjects were less likely to report seeing stimuli that were associated with losing points. We hypothesized that the reason for these null results may have been the overall higher accuracy levels in Exp. 4 (mean accuracy in the included sample: 0.81, d’= 2.05 compared to 0.78, d’= 1.64 in Exp. 3). Based on relatively low accuracy levels in a small pilot sample, and given the overall lower d’ in detection compared to discrimination tasks, we increased the visibility levels in Exp. 4 from 10% to 20%. This increase may have made the task too easy, as a result masking potential effects of motivation on decision bias. In Exp. 5, we brought visibility back to 10%. Apart from this one change, Exp. 5 was perfectly identical to Exp. 4.

## Methods

Exp. 5 was identical to Exp. 4 except for the stimulus visibility level in target-present (‘stripes’) trials, which was now set to 10%.

## Results

121 participants were recruited from Prolific for Experiment 5.

Median completion time was 17.73 minutes. Mean accuracy was 0.75. Participants reported target presence on 0.54 of the trials. Overall we excluded 21 participants based on our exclusion criteria, leaving 100 participants for the main analysis.

In Exp. 5, associating stimulus absence with point loss had a significant effect on decision bias. The proportion of target-presence reports was higher absence-loss (0.59) compared to neutral blocks (0.55; , 95% CI , , ).

In contrast, associating stimulus presence with point loss had no significant effect on decision bias (0.53 in presence loss compared to 0.55 in neutral blocks ; , 95% CI , , ). Still, the effects of presence-loss and absence-loss were not significantly different in magnitude (, 95% CI , , ), leaving open the possibility that this difference in significance between the two conditions is driven by sampling noise.

We next turned to response times. Like in Exp. 4, a two-way repeated measured analysis of variance revealed an effect of response (target presence or absence) on response time (), driven by target absence responses being slower by 64.20 ms on average. Motivation condition also had a significant effect on response time (), driven by a general slowing down in the presence-loss condition relative to the neutral (, 95% CI , , ) and absence-loss conditions (, 95% CI , , ).

Furthermore, the effects of response and motivation conditions presented a significant interaction with response time (). Responses were slower in the presence loss condition relative to neutral for both ‘target present’ (, 95% CI , , ) and ‘target absent’ responses (, 95% CI , , ). In contrast, the absence loss condition had a significant slowing down effect on response time only for ‘target absent’ responses (, 95% CI , , ), but not for ‘target present’ responses (, 95% CI , , ), with a significant condition by response interaction (, 95% CI , , ).

### White Bear Suppression Inventory.

In Exp. 5, we pre-registered our hypothesis for a significant correlation between WBSI scores and their constituent factors and presence loss effects on decision bias.

The mean WBSI score was 50.57 and consistent with the mean scores reported in Wegner and Zanakos (1994) (-) and (Blumberg (2000) (). Relibability was high, with a Cronbach alpha of 0.95. WBSI scores did not correlate with presence-loss effects on decision bias (, 95% CI , , ).

To our surprise, this time WBSI scored did not correlate with absence-loss effects on decision bias (, 95% CI , , ), and there was no significant difference between the correlation of WBSI scores with the presence-loss and absence-loss manipulations ().

# Experiment 6

In Exp. 4, decision bias was not significantly affected by associating target presence or absence with point loss. We hypothesized that the reason for these null results may have been the overall higher accuracy levels in Exp. 4, and in Exp. 5 we decreased visibility from 20% to 10%. Accuracy in Exp. 5 was still relatively high (0.76, d’=1.71). In both experiments, a clear biasing effect of the presence-loss manipulation emerged when restricting the analysis to subjects with d’ below 2.0 only. In Exp. 6, we reduced the visibility level further to 5%, to decrease d’ and allow an effect to appear.

## Methods

Exp. 6 was identical to Exp. 4 and 5 except for the stimulus visibility level in target-present (‘stripes’) trials, which was now set to 5%.

## Results

134 participants were recruited from Prolific for Experiment 5. Due to a bug on the Prolific system, a subset of 18 subjects that took part in Exp. 6 already participated in previous experiments.

Median completion time was 18.68 minutes. Mean accuracy was 0.70. Participants reported target presence on 0.48 of the trials. Overall we excluded 34 participants based on our exclusion criteria, leaving 100 participants for the main analysis.

In Exp. 6, associating stimulus absence with point loss had no significant effect on decision bias. The proportion of target-presence reports in absence-loss (0.49) was numerically higher but not significantly different then in neutral blocks (0.47; , 95% CI , , ).

Similarly, response bias on blocks in which stimulus presence was associated with point loss was not significantly different from response bias in neutral blocks (0.47; , 95% CI , , ). The effects of presence-loss and absence-loss were not significantly different in magnitude (, 95% CI , , ).

We next turned to response times. Unlike in Exp. 4 and 5, here a two-way repeated measured analysis of variance revealed no effect of response (target presence or absence) on response time (). Motivation condition too had no significant effect on response time (), as did the interaction between condition and response ().

### White Bear Suppression Inventory.

The mean WBSI score was 51.45 and consistent with the mean scores reported in Wegner and Zanakos (1994) (-) and (Blumberg (2000) (). Relibability was high, with a Cronbach alpha of 0.94.

WBSI scores were negatively correlated with presence-loss effects on decision bias (, 95% CI , , ). This is opposite to what we found in Exp. 4, where the correlation between WBSI and presence-loss effects was negative. The difference between the correlations observed in Exp. 4 and 6 was statistically significant, suggesting that it is unlikely to reflect sampling noise only ( 3.63, ).

Similar to Exp. 4 and 5, WBSI scored did not correlate with absence-loss effects on decision bias (, 95% CI , , ). In contrast to Exp. 5, the correlation between WBSI scores and presence-loss effects was significantly more negative than the correlation of WBSI scores with absence-loss effects ().

# Experiment 7

In Exp. 4-6, stimulus presentation time was dynamically staircased throughout test blocks. This staircasing may have masked bias effects, since the loss mainpulation is expected to have an effect on accuracy via changes to the decision criterion. Experiment 7 will be similar to Exp. 6, without the staircasing procedure.

## Methods

Exp. 7 was identical to Exp. 5 except for the stimulus presentation time which was now set to 250 for all trials. In addition, in light of high rejection rates due to miscomprehension of task instructions, we now changed the way we treated comprehension check failures: participants were given feedback about their errors, instead of being presented with the task instructions again.

## Results

171 participants were recruited from Prolific for Experiment 7.

Median completion time was 15.77 minutes. Mean accuracy was 0.71. Participants reported target presence on 0.49 of the trials. Overall we excluded 37 participants based on our exclusion criteria, leaving 134 participants for the main analysis.

In Exp. 7, associating stimulus absence with point loss had a significant effect on participants’ decision bias. The proportion of target-presence reports in absence-loss (0.53) was significantly higher than in neutral blocks (0.49; , 95% CI , , ).

Similarly, the proportion of target-presence responses on blocks in which stimulus presence was associated with point loss was 0.45 and significantly reduced compared to neutral blocks (, 95% CI , , ). The effects of presence-loss and absence-loss were not significantly different in magnitude (, 95% CI , , ).

We next turned to response times. Similar to Exp. 6, a two-way repeated measured analysis of variance revealed no effect of response (target presence or absence) on response time (). Motivation condition too had no significant effect on response time (). However, the interaction between the two did have a significant effect on response time (): subjects were significantly faster to report target absence when target presence was associated with point loss (, 95% CI , , ), and significantly faster to report target presence when target absence was associated with point loss (, 95% CI , , ). The respective effects of these manipulations on target presence and target absence responses were not significant ().

### White Bear Suppression Inventory.

The mean WBSI score was 51.87 and slightly higher than the mean scores reported in Wegner and Zanakos (1994) (-) and (Blumberg (2000) (). Relibability was high, with a Cronbach alpha of 0.92.

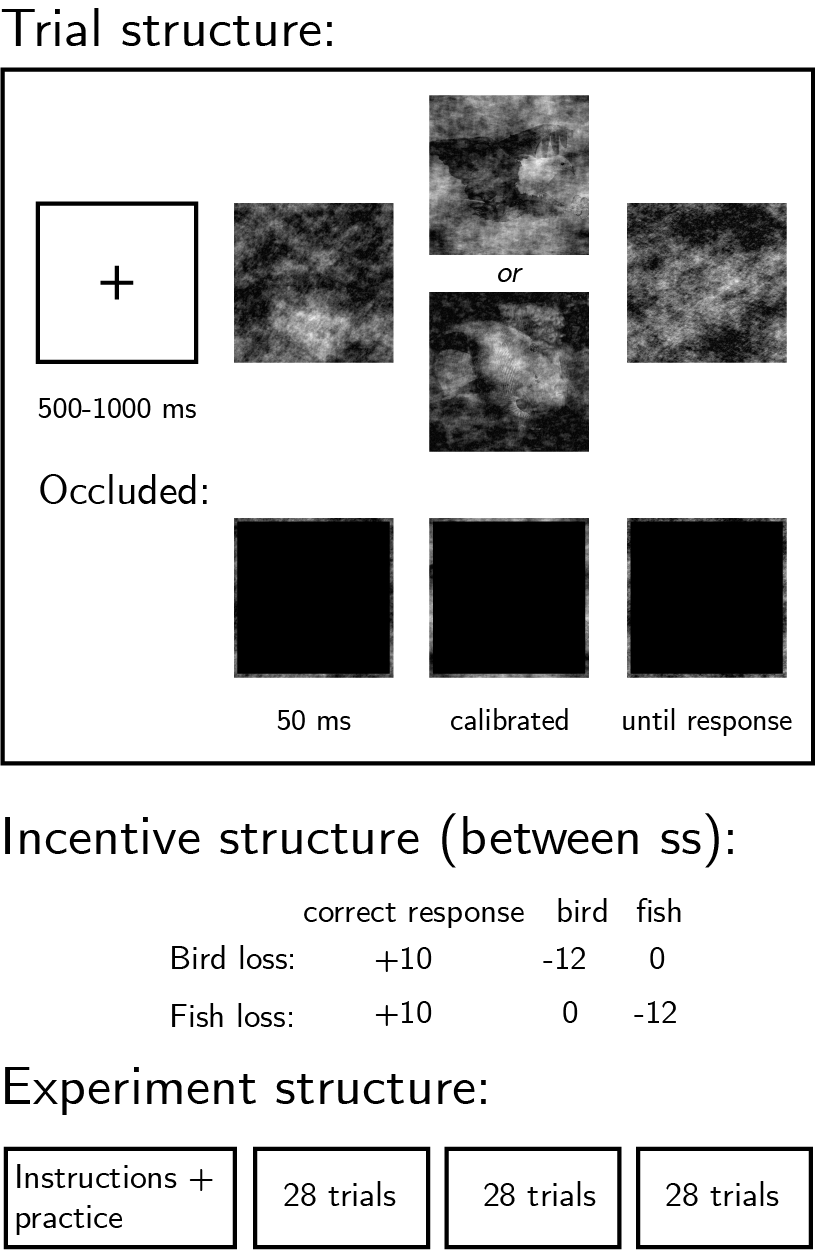
WBSI scores were not reliably correlated with presence-loss effects on decision bias (, 95% CI , , ), not with the effects of absence-loss on decision bias (, 95% CI , , ).

# Experiment 8

Assuming that representations of the target-absent stimulus are not amenable to suppression at the perceptual level, symmetric decision bias effects of associating stimulus presence or absence with point loss seem inconsistent with a sensory suppression account. Here we examined an alternative account, namely that motivated seeing effects emerge not at the perceptual, but at the decision level.

## Methods

Exp. 8 followed the general structure of Exp. 1 and 2, where subjects decided whether noisy images included a bird or a fish. However, we introduced several important changes. First, the motivation manipulation was tested in a between-subject manner: for different participants, either fish or bird images were associated with point loss. And second, on a random subset of trials, stimuli were masked by a black square. On these trials, subjects were asked to guess the content of the image underneath the square, in the absence of any perceptual input. These trials served as a non-perceptual control condition, where decisional biases cannot be driven by perception.



*Figure* *2.*  Experimental design, Exp. 8. Top left: Trial structure. Presentation time will be calibrated to achieve 70% accuracy. On occluded trials, a black square will occlude the images. Middle: incentive structure for different subjects. Bottom: Overall experiment structure.

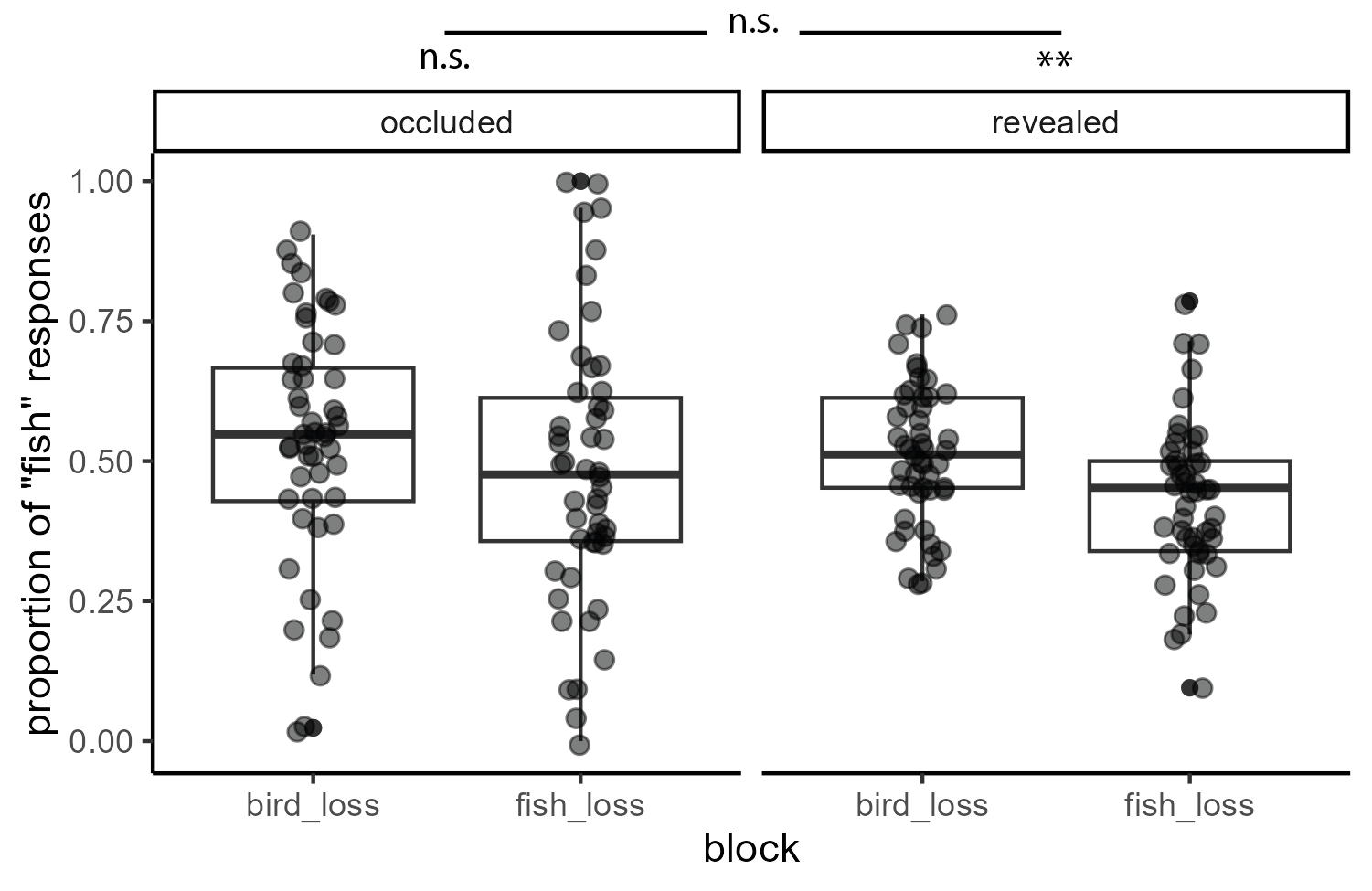
## Results

138 participants were recruited from Prolific for Experiment 7.

Median completion time was 9.30 minutes. Mean accuracy was 0.65 in non-occluded trials and 0.50 in occluded trials. Participants reported seeing a fish on 0.49 of the trials. Overall we excluded 38 participants based on our exclusion criteria, leaving 100 participants for the main analysis.

Consistent with the results from Exp. 1 and 2, the loss manipulation affected response bias in non-occluded trials. Participants for which bird stimuli were associated with point loss reported seeing a fish on 0.51 of the trials, significantly more than the proportion of fish reports in the group of participants for whom fish images were associated with point loss (0.44; , ).

Importantly, a similar effect was observed in occluded trials (mean proportion of fish reports among bird-loss participants: 0.54, and among fish-loss participants: 0.48), but this effect was not significant (, ). Importantly, the difference in response bias between occluded and non-occluded trials was itself not different between the two groups (, ), providing no support for a perceptual effect beyond the biasing effect in occluded trials.



*Figure* *3.*  Propotion of fish responses as a function of group (bird-loss or fish-loss) and condition (occluded or not) in Exp. 8.

Response times were similar for stimuli that were or were not associated with point loss, both when stimuli were visible (, 95% CI , , ), and when they were occluded (, 95% CI , , ) with no significant interaction between stimulus value and occlusion on response time (, 95% CI , , ).

# References

Balcetis, E., Dunning, D., & Granot, Y. (2012). Subjective value determines initial dominance in binocular rivalry. *Journal of Experimental Social Psychology*, *48*(1), 122–129.

Blumberg, S. J. (2000). The white bear suppression inventory: Revisiting its factor structure. *Personality and Individual Differences*, *29*(5), 943–950.

Leong, Y. C., Dziembaj, R., & D’Esposito, M. (2021). Pupil-linked arousal biases evidence accumulation toward desirable percepts during perceptual decision-making. *Psychological Science*, *32*(9), 1494–1509.

Leong, Y. C., Hughes, B. L., Wang, Y., & Zaki, J. (2019). Neurocomputational mechanisms underlying motivated seeing. *Nature Human Behaviour*, *3*(9), 962–973.

Libera, C. D., & Chelazzi, L. (2006). Visual selective attention and the effects of monetary rewards. *Psychological Science*, *17*(3), 222–227.

Mazor, M. (2021). *Inference about absence as a window into the mental self-model*.

Mazor, M., & Fleming, S. M. (2020). Distinguishing absence of awareness from awareness of absence. *Philosophy and the Mind Sciences*, *1*(II).

Mazor, M., & Fleming, S. M. (2022). Efficient search termination without task experience. *Journal of Experimental Psychology: General*.

Mazor, M., Mazor, N., & Mukamel, R. (2019). A novel tool for time-locking study plans to results. *European Journal of Neuroscience*, *49*(9), 1149–1156.

Mazor, M., Moran, R., & Fleming, S. M. (2021). Metacognitive asymmetries in visual perception. *Neuroscience of Consciousness*, *2021*(1), niab025.

Meuwese, J. D., Loon, A. M. van, Lamme, V. A., & Fahrenfort, J. J. (2014). The subjective experience of object recognition: Comparing metacognition for object detection and object categorization. *Attention, Perception, & Psychophysics*, *76*(4), 1057–1068.

Pessoa, L., & Engelmann, J. B. (2010). Embedding reward signals into perception and cognition. *Frontiers in Neuroscience*, *4*, 17.

Wegner, D. M., & Zanakos, S. (1994). Chronic thought suppression. *Journal of Personality*, *62*(4), 615–640.