Psychonomic Bulletin & Review
https://doi.org/10.3758/s13423-024-02539-8

BRIEF REPORT

Dynamics of sensory and decisional biases in perceptual decision making: Insights from the face distortion illusion

Yi Gao[1] · Sixing Chen[2] · Dobromir Rahnev[1]

Accepted: 19 June 2024

Abstract
Bias in perceptual decision making can have both sensory and decisional origins. These distinct sources of bias are typically seen as static and stable over time. However, human behavior is dynamic and constantly adapting. Yet it remains unclear how sensory and decisional biases progress in distinct ways over time. We addressed this question by tracking the dynamics of sensory and decisional biases during a task that involves a visual illusion. Observers saw multiple pairs of peripherally presented faces that induce a strong illusion making the faces appear distorted and grotesque. The task was to judge whether one of the last two faces had true physical distortion (experimentally introduced in half of the trials). Initially, participants classified most faces as distorted as exemplified by a liberal response bias. However, over the course of the experiment, this response bias gradually disappeared even though the distortion illusion remained equally strong, as demonstrated by a separate subjective rating task without artificially distorted faces. The results suggest that the sensory bias was progressively countered by an opposite decisional bias. This transition was accompanied by an increase in reaction times and a decrease in confidence relative to a condition that does not induce the visual illusion. All results were replicated in a second experiment with inverted faces. These findings demonstrate that participants dynamically adjust their decisional bias to compensate for sensory biases, and that these two biases together determine how humans make perceptual decisions.

Keywords Bias · Perceptual decision making · Confidence · Flashed face distortion Effect · Face inversion

Introduction

Perceptual decision making can be influenced by both sensory and decisional biases. Sensory biases stem from the processing of sensory evidence and are exemplified by visual illusions like the Muller-Lyer illusion (Witt et al., 2015). In contrast, decisional biases arise within the decision-making process and are exemplified by adjustments to response strategies in the absence of differences in sensory evidence (Gold & Shadlen, 2007). Most previous research has attempted to determine whether a particular bias occurs due to sensory (Jazayeri & Movshon, 2007; Rahnev, 2021b; Shams & Kim, 2010), decisional (Bang & Rahnev, 2017; Bowen et al., 2020; Odgaard et al., 2003), or both sensory and decisional processes (Linares et al., 2019; Rahnev & Denison, 2018). The underlying assumption in this line of work is that, unless the task structure changes (e.g., by introducing different payoff schemes), sensory and decisional biases are stable and do not change over time.

However, research has shown that human behavior is inherently dynamic and that our decisions often undergo changes over time. These dynamics are well documented in fields such as reinforcement learning (Shteingart & Loewenstein, 2014) but occur for virtually all human behavior, including low-level perceptual effects such as visual adaptation (Gao et al., 2022; Webster, 2015) and perceptual learning (Goldstone, 1998). Human decisions are also known to vary on a short time scale based on immediate previous stimuli and decisions (Frund et al., 2014; Yu & Cohen, 2008), as well as drift unpredictably over longer time scales (Norton et al., 2017; Wei et al., 2021). Critically, even perceptual changes, such as visual adaptation, have been argued to have both sensory and decisional components (Witthoft et al., 2018). Therefore, it is likely that sensory and decisional biases are dynamically adjusted over time.

✉ Yi Gao
yi.gao0525@outlook.com

1 School of Psychology, Georgia Institute of Technology, Atlanta, GA, USA

2 Department of Psychology, New York University, New York, NY, USA

Published online: 09 July 2024

Yet no study to date has reported the existence of dynamic changes of either sensory or decisional biases that occur predictably and consistently across participants in the absence of changes to the task structure. Most studies simply average the data across time under the implicit assumption that both types of biases remain largely stable. However, if the two biases do change over time, such averaging can lead to misleading results (Gao et al., 2019, 2021). Therefore, it is critical to track the evolution of sensory and decisional biases to determine their temporal dynamics.

Here we sought to determine if we can identify changes in sensory and decisional biases over the course of a single session with a perceptual task. Participants completed a task that induces the flashed face distortion effect (FFDE), a visual illusion where the presentation of a series of faces leads to distorted or grotesque face perception (Balas & Pearson, 2019; Bowden et al., 2019; Tangen et al., 2011; Wen & Kung, 2014). The face distortion illusion typically requires the presentation of multiple faces in spatial alignment. Although the underlying mechanisms are largely unidentified, this phenomenon offers a valuable opportunity to investigate how sensory and decision biases dynamically evolve throughout the development of the illusory percept. The task was to judge whether one of the last two faces had true physical distortion, which was experimentally introduced in half of the trials. Critically, in addition to the illusion condition where we present a series of faces, we also included a control condition that did not produce face distortion. We found that the illusion and control condition initially featured vastly different response bias (participants were much more likely to indicate that faces are distorted in the illusion compared with the control condition), but that this difference disappeared over time. These results were driven by a stable sensory bias that was progressively countered by an opposite decisional bias, an affect that could also be observed in reaction time and confidence data. Our findings thus demonstrate that decisional biases are dynamically adjusted to compensate for sensory bias.

Methods

Participants

We recruited 27 participants for Experiment 1, tested with upright faces (12 females and 15 males, mean age = 19.74 years, *SD* = 1.68, range: 18–25). We additionally recruited another 27 participants for a control experiment with inverted faces (nine females and 18 males, mean age = 19.96 years, *SD* = 3.07, range: 18–31). These sample sizes were chosen to achieve 80% statistical power given a medium effect size (0.5). We recruited another 20 participants for Experiment 2 (five females and 15 males, mean age = 20.35 years, *SD* = 1.93, range: 18–26). All participants had normal or corrected-to-normal vision and provided a signed consent form approved by the Institutional Review Board of the Georgia Institute of Technology. Participants were recruited from the Georgia Institute of Technology and were compensated for 1 SONA credit or $10/hour.

Procedure

All experiments consisted of two conditions designed to either induce or not induce a visual illusion (Fig. 1A). In the illusion condition, we presented 15 flashes of faces that induced a known visual illusion called the flashed face distortion effect (Balas & Pearson, 2019; Bowden et al., 2019; Tangen et al., 2011; Wen & Kung, 2014). In this illusion condition, faces presented after the first few flashes begin to appear increasingly distorted and even grotesque. In contrast, in the control condition, we presented a single flash of faces that did not induce a perceptual illusion. In both cases, a single flash included two different faces presented on the left and right of fixation, following a standard design in the literature (Balas & Pearson, 2019; Wen & Kung, 2014). While it is possible to present a single face on the left or right of fixation, such a design would risk participants inadvertently diverting their gaze from the fixation dot.

In Experiment 1, participants' task was to indicate whether any of the last pair of faces was physically distorted. To make the task meaningful, we manually created artificially distorted faces (Fig. 1B). On half the trials from each condition, either the left or the right face was artificially distorted, whereas on the other half of the trials, both faces were undistorted. In the illusion condition, the first 14 flashes always had only undistorted faces. Prior to the experiment, we informed participants that they might experience the face distortion illusion in certain instances.

In Experiment 1, each trial began with the presentation of a red dot at fixation for a random duration between 800 to 1,300 ms. Depending on the condition, we then presented either 15 or a single flash of faces (height = 6 degrees, width = 6 degrees). We included two stimulus manipulations—flash duration and stimulus eccentricity. Specifically, on half the trials, each flash lasted for 250 ms, whereas on the other half of the trials, each flash lasted for 1,000 ms. Independently, on half the trials, the faces were presented at 2° eccentricity, whereas on the other half of the trials, the faces were presented at 8° eccentricity. These manipulations were included as part of an independent effort to replicate previous findings regarding how flash duration and eccentricity affect FFDE (Balas & Pearson, 2019), which was published separately (Gao et al., 2024). Nevertheless, we confirmed that our main results of an interaction between the response bias for the illusion and control conditions were

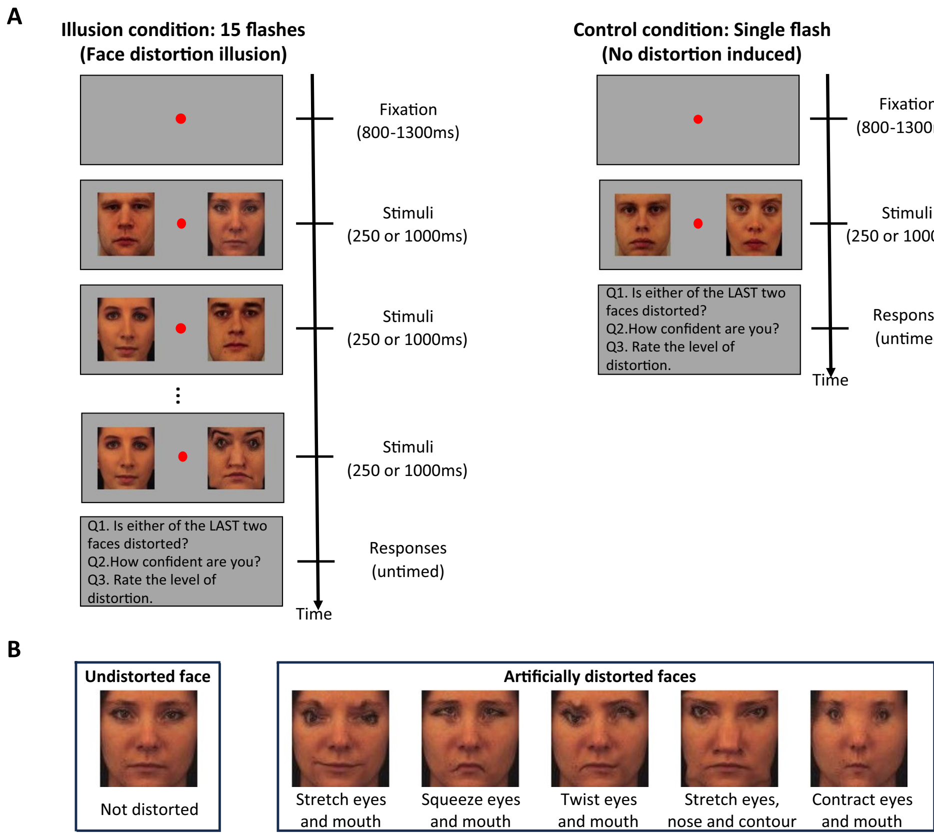

Fig. 1 Task and stimuli. **A** Example trials from the illusion and control conditions. The illusion condition featured 15 flashes of faces that led to a strong visual illusion where faces appeared distorted and grotesque. The control condition featured a single flash of faces and did not induce any illusionary distortion. All faces on a given trial were presented for either 250 or 1,000 ms (same duration for all faces). Participants answered three questions after each trial: Q1. Is either of the last two faces distorted? Q2. How confident are you? Q3. Rate the level of distortion. **B** An example of an undistorted face and artificially distorted faces created by different methods. (Color figure online)

present separately in each experimental condition (Supplementary Fig. 2).

Each combination of number of flashes, duration, and eccentricity was repeated twice per block with each block having 16 trials. After the face presentation, participants saw a single response screen with three questions. The first question asked participants to indicate whether either of the last two faces was distorted. The second question asked participants to rate their confidence using a 4-point scale (*not confident at all*, *somewhat confident*, *very confident*, *extremely confident*). Finally, the third question asked participants to provide a subjective rating of the level of distortion (*not distorted*, *minor distortion*, *major distortion*, *extreme distortion*). This question was included as part of our independent effort to replicate previous findings in the literature (Balas & Pearson, 2019) and was not analyzed here. Participants were allowed unlimited time to respond. Each question appeared on the same response screen, but each subsequent question was shown only after the participant responded to the previous question. Participants did not receive any feedback following their responses. Participants completed 10 blocks corresponding to a total of 160 trials.

In Experiment 1, we presented all faces upright. This standard approach to induce FFDE may lead to increased sensitivity in the illusion but not in the control condition over time, complicating response bias interpretation. To address this, we conducted a replication experiment with all faces presented in an inverted orientation. This replication aimed to equalize learning effects between illusion and control conditions by minimizing the disparity in sensitivity change over time.

Experiment 1 contained artificially distorted faces, which allowed us to measure response bias. We conducted Experiment 2 to test whether the strength of the distortion illusion changed over time. The experiment had the same design as Experiment 1, except for the following two changes. First, we only used normal faces (i.e., no faces were artificially distorted). Second, because we were only interested in participants' subjective ratings, participants only indicated the subjective level of distortion (third question from Experiment 1) but did not need to judge whether faces were objectively distorted or provide confidence (since there were no artificially distorted faces). These changes were designed to directly test the strength of participants' subjective distortion, independent of their judgments regarding objective distortion (which made Question 3 difficult to interpret in Experiment 1).

To confirm that the face distortion illusion did not weaken over time, we sent out a brief, one-question survey a few weeks following Experiment 1. We asked participants, "Did this illusion change over the course of the experiment?" We note that all participants were aware that they were experiencing a visual illusion. First, we had informed them about the illusion in the preexperiment instructions. Second, participants knew that only the last flash could contain artificially distorted faces, but people begin to experience distortion starting after only a few face flashes. Therefore, it is easy to infer that the distortion one experiences before the last flash is due to an illusion. Participants chose among five response options, indicating whether the illusion became much weaker, a little weaker, a little stronger, much stronger, or that it stayed the same. In Experiment 2, we collected the same survey response as in Experiment 1 immediately upon completion of the experiment.

Stimuli

To generate the artificially distorted faces, we employed five distinct distortion methods. The first four methods used the nudge function of PicMonkey Photo Editor and Graphic Design Maker (picmonkey.com). All original faces were positioned on an 8 by 8 grid, and individual points on each face were shifted manually to produce different distortions. The four artificial face distortions involved (1) stretching the eyes and mouth, (2) squeezing the eyes and mouth, (3) twisting the eyes and mouth, and (4) stretching the eyes, nose, and contour of the face (see Fig. 1B for examples of distorted faces). For the fifth distortion method, we utilized custom MATLAB codes (The MathWorks, Natick, MA) to contract the eyes and mouth of each face using algorithms described in Webster and Maclin (1999) and Yamashita et al. (2005) (see Fig. 1B for an example distorted face). In addition, we adjusted the luminance of the face images to match the average luminance of the face set. The original face images were selected from the Karolinska directed emotional faces database (Lundqvist et al., 1998). We chose 15 male and 15 female faces and created five distorted versions of each face according to all five methods described above. The faces presented were randomly drawn so that in the illusion condition each of the 30 identities was shown exactly once, whereas in the control condition two identities were selected randomly.

Analyses

For each block, we computed sensitivity (d') and response criterion (c) based on the signal detection theory (Green & Swets, 1966) formulas:

$$d\prime = \varphi^{-1}(hitrate) - \varphi^{-1}(falsealarmrate)$$

and

$$c = -\frac{\varphi^{-1}(hitrate) + \varphi^{-1}(falsealarmrate)}{2},$$

where φ^{-1} represents the inverse of the cumulative standard normal distribution transforming the hit rate and false alarm rate to Z scores. We coded objectively distorted faces as the nontarget and nondistorted faces as the target. As such, a larger c value indicates a bias towards saying "distorted" more frequently, which indexes the presence of the flashed face distortion illusion.

We performed two-way analysis of variance (ANOVA) with repeated measures to compare d', the response criterion, confidence rating, and reaction time (RT) between the illusion and control conditions, with block number and condition as the independent factors. The RT was computed as the time it took participants to respond to the initial question regarding whether either of the two faces was distorted. To maximize power, we merged the data from different flash durations and eccentricities. Separating eccentricity and flash duration into distinct factors would necessitate estimating sensitivity (d') and response bias (c) from merely four trials per participant, rendering the estimation noisy. Consequently, we opted to aggregate the data over both eccentricity and flash duration.

In addition, we fitted the data of the 10 blocks with linear regression models and performed paired t tests on the slopes of the models to determine how d', c, confidence, and RT changed over blocks. We report effect size for all t tests and ANOVAs and Bayes factors (Krekelberg, 2022) for all t tests and ANOVAs where $p > 0.05$.

Results

Experiment 1: Sensory bias progressively counteracted by an opposite decisional bias.

Response *bias* diminishes over time for the illusion condition only

In Experiment 1, participants completed a face distortion detection task over 10 blocks. Critically, there were two conditions: an illusion condition that made even normal faces appear distorted and a control condition that did not induce illusory percepts. We first examined how response bias changed over time in the illusion and control conditions. We found an interaction between condition and block number, $F(9, 234) = 6.19$, $p = 7.8 \times 10^{-8}$, $\eta_p^2 = 0.19$. Post hoc analysis shows that, initially, participants were much more biased towards saying "distorted" in the illusion compared with the control condition (all Bonferroni-corrected $ps < 0.002$, all $BF_{10} > 262.68$), but that this difference disappeared by the last three blocks (all uncorrected $ps > 0.64$, all $BF_{01} > 4.44$; Fig. 2A). To capture the change over time, we computed the slope of the response criterion over the 10 blocks and found that the response criterion became progressively less biased over blocks in the illusion condition, $t(26) = 4.09$, $p = 0.0004$, $BF_{10} = 81.55$, Cohen's $d = 0.79$, but remained relatively stable over blocks in the control condition, $t(26) = 11.54$, $p = 0.14$, $BF_{01} = 1.72$, Cohen's $d = 0.30$. Critically, the response criterion decreased at a faster rate in the illusion than in the control condition, $t(26) = 5.23$, $p = 1.8 \times 10^5$, $BF_{10} = 1.2 \times 10^3$, Cohen's $d = 1.01$. This pattern of results demonstrates that participants experienced strong distortion percepts in the illusion condition. Like many other illusions, the face distortion illusion does not become weaker over time, which suggests that participants adjusted their decisional bias over time to counteract the perceptual bias caused by the illusion.

One possibility is that the decrease in response bias in the illusion condition is due to a corresponding change in sensitivity. To check whether this is the case, we examined how sensitivity (d') changed over time in each condition (Fig. 2B). We found an interaction between condition and block number, $F(9, 234) = 3.02$, $p = 0.002$, $\eta_p^2 = 0.10$. Specifically, sensitivity did not change over time in the control condition, $t(26) = 0.45$, $p = 0.65$, $BF_{01} = 4.46$, Cohen's $d = 0.09$, but increased slightly across blocks in the illusion condition, $t(26) = 3.68$, $p = 0.001$, $BF_{10} = 31.93$, Cohen's $d = 0.71$. These results are consistent with the notion that participants exhibited faster learning in the more unusual illusion condition, as also suggested by a significant difference in the rate of d' increase, $t(26) = 2.94$, $p = 0.007$, $BF_{10} = 6.52$, Cohen's $d = 0.57$. These results do not exclude the possibility that the response bias is partially explained

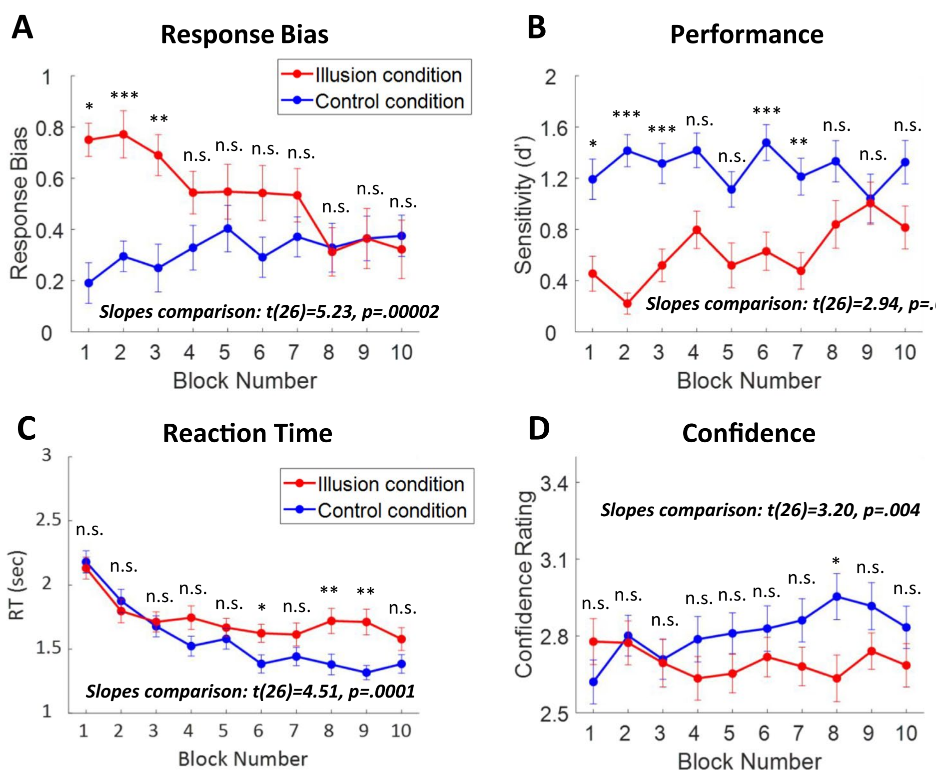

Fig. 2 Results for Experiment 1 with upright faces. **A** Response bias sharply decreased over time in the illusion but remained stable in the control condition. **B** Sensitivity (*d'*) improved slightly in the illusion but remained largely stable in the control condition. **C** RT decreased over time in both conditions, but the decrease was smaller for the illusion compared with the control condition. **D** Confidence was initially similar across the two conditions but became gradually lower for the illusion compared with the control condition. Overall, these results are consistent with the notion that the perceptual bias induced by the face distortion illusion was gradually counteracted by an opposite decisional bias. Error bars indicate *SEM*. Asterisks indicate the results of paired *t* tests after Bonferroni correction between the control and illusion conditions. ***$p < .001$. **$p < .01$. *$p < .05$. n.s., $p > .05$. (Color figure online)

by d' but make it unlikely that the shift in response bias is solely explained by the change in sensitivity.

To address this issue, we conducted a control experiment ($N = 27$) that used an identical design except that all faces were inverted instead of upright. This design led to an equivalent change in d' over time for the illusion and control conditions, such that there was not interaction between condition and block number, $F(9, 234) = 0.49$, $p = 0.88$, $BF_{01} = 169.86$, $\eta_p^2 = 0.02$ (Supplementary Fig. 1). Nevertheless, we still found a significant interaction between condition and block number for response bias, $F(9, 234) = 6.04$, $p = 1.3 \times 10^{-7}$, $\eta_p^2 = 0.19$. These results replicate our findings with upright faces and show that the response bias effects occur even in the absence of corresponding changes in d'.

Shift in response *bias* in the illusion condition is accompanied by corresponding shifts in RT and confidence

The results so far suggest that participants continuously adjusted their decisional bias to counteract the perceptual bias in the illusion condition. We therefore examined whether such decisional adjustment would be reflected in a change in RT and confidence over time. This was indeed the case: Both RT and confidence were similar for the illusion and control conditions in the early blocks for diverged in the later blocks (Fig. 2C–D). Indeed, we found an interaction between condition and block number for both RT, $F(9, 234) = 4.63$, $p = 1.2 \times 10^{-5}$, $\eta_p^2 = 0.15$, and confidence, $F(9, 234) = 2.8$, $p = 0.004$, $\eta_p^2 = 0.10$. Specifically, the rate of decrease in RT over time was slower for the illusion compared with the control condition, $t(26) = 4.51$, $p = 1.2 \times 10^{4}$, $BF_{10} = 222.08$, Cohen's $d = 0.87$. Similarly, the rate of increase in confidence was higher for the control compared with the illusion condition, $t(26) = 3.20$, $p = 0.004$, $BF_{10} = 11.14$, Cohen's $d = 0.62$. These results are consistent with a gradual increase in decisional bias over time in the illusion condition, resulting in slower decisions and decreased subjective confidence in the illusion compared with the control condition.

Experiment 2: The face distortion illusion does not weaken over time.

The results of Experiment 1 are consistent with the interpretation that the perceptual bias induced by the face distortion illusion is counteracted over time by an opposite decisional bias. However, this interpretation relies on the assumption that the face distortion illusion did not weaken over time. To track how the strength of the face distortion illusion changed over time, we conducted Experiment 2 where no artificially distorted faces were included. We asked participants after each trial to rate the subjective strength of face distortion they perceived. Note that we collected similar ratings in Experiment 1, but the question there was ambiguous because the presence of artificially distorted faces likely led some participants to report their belief that a face is artificially distorted rather than simply indicate their subjective experience. By removing artificially distorted faces altogether in Experiment 2, we were able to obtain purer ratings of participants' subjective experience.

We found a very large main effect of condition, $F(1, 171) = 127.99$, $p = 7.0 \times 10^{-10}$, $\eta_p^2 = 0.93$, confirming that participants' distortion ratings were significantly higher in the illusion compared with the control condition (Fig. 3). Critically, there was no main effect of block number, $F(9, 171) = 0.92$, $p = 0.51$, $BF_{01} = 691.69$, $\eta_p^2 = 0.09$, and no interaction between block and condition, $F(1, 171) = 1.89$, $p = 0.06$, $BF_{01} = 135.37$, $\eta_p^2 = 0.09$, demonstrating that the strength of the face distortion illusion did not change over time. This conclusion was further supported by a direct test for the slope in the illusion, $t(19) = 0.06$, $p = 0.95$, $BF_{01} = 4.30$, Cohen's $d = 0.01$, and the control condition, $t(19) = 2.45$, $p = 0.02$, $BF_{10} = 2.50$, Cohen's $d = 0.55$.

In addition, after the completion of each experiment (Experiment 1, Experiment 2, and the control experiment), we asked participants whether the illusion changed over the course of the experiment. Participants consistently reported experiencing the face distortion illusion as maintaining its intensity or, in some instances, becoming slightly stronger over time (see Supplementary Results and Supplementary Fig. 3). Overall, these results strongly support the notion that the strength of the face distortion illusion either remained consistent over time or perhaps even slightly increased. Therefore, the decrease in response bias in Experiment 1 is most consistent with decisional bias counteracting the

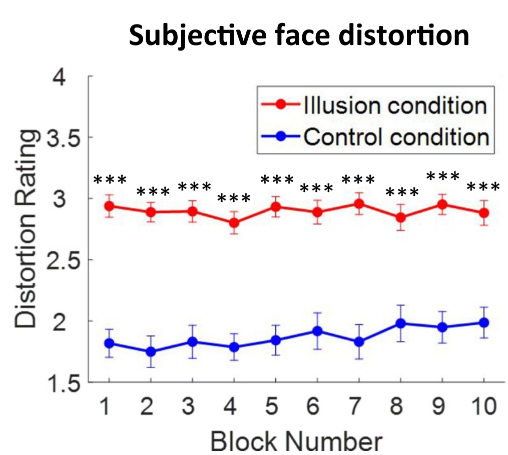

Fig. 3 Experiment 2: The face distortion illusion does not weaken over time. Subjective face distortion ratings in Experiment 2. The distortion strength did not change over time in the illusion condition, while it slightly increased for the control condition. Error bars indicate *SEM*. Asterisks indicate the results of Bonferroni-corrected paired *t* tests between the control and illusion conditions. ***$p < .001$. (Color figure online)

sensory bias, rather than the sensory bias becoming weaker over time.

Discussion

Perceptual decision making contains at least two components: sensory processing, primarily occurring within an early time window in occipital areas, and decision making, primarily occurring later in parietal and frontal areas (Mostert et al., 2015). The sensory and decisional components of perceptual decision making can separately bias the final perceptual decision, but distinguishing between these two types of biases has been challenging, and the dynamics of these biases are virtually unknown. Here, we show that sensory and decisional biases can be dissociated during a single session in which participants completed a task that induced a face distortion illusion. We showed that response criterion, which encompasses both sensory and decisional biases, decreases substantially over the course of the session. Using a separate distortion rating task, we demonstrated that the sensory signal remained unchanged over the course of the session, signifying that it is the decision bias that was dynamically adjusted by participants. In accordance with this interpretation, we also found corresponding changes in both RT and confidence. These findings offer compelling evidence for the differentiation of sensory and decisional biases and reveal that people dynamically adjust their decisional bias to compensate for sensory information judged to be illusory.

Studies to date have struggled separating sensory and decisional biases, making it difficult to track individual changes separately. For example, while signal detection theory allows separating sensitivity and response bias (Green & Swets, 1966), the response bias encompasses both sensory and decisional biases (Bang & Rahnev, 2017; Rungratsameetaweemana & Serences, 2019; Rungratsameetaweemana et al., 2018; Witt et al., 2015). Similarly, some studies have attempted to map the parameters of the drift diffusion model to sensory versus decisional effects (Germar et al., 2014; Voss et al., 2008), but this mapping has been shown not to hold in all cases (Chen & Rahnev, 2023; Sánchez-Fuenzalida et al., 2023; Starns et al., 2012; White & Poldrack, 2014). Some researchers have proposed that examining the distributions of confidence and RT as a function of feature strength can distinguish sensory from decisional biases (Gallagher et al., 2019, 2021; Maldonado Moscoso et al., 2020), but this conclusion has also been challenged (Chen & Rahnev, 2023). Thus, while many signatures of sensory versus decisional effects have been examined, clearly distinguishing between the two biases remains a challenge.

Here, we take a different approach to dissociating sensory from decisional biases. First, we use a strong visual illusion (the flashed face distortion effect) that is known to lead to large sensory biases. Because they induce consistent and large sensory bias, visual illusions are an ideal test case for tracking compensation by decisional biases. Second, we track the sensory biases via a separate task. This approach allows us to sidestep the difficulties of using the same set of data to separately infer sensory and decisional biases. Our approach has similarities with the approach by Sánchez-Fuenzalida et al. (2023), who also used a visual illusion to induce a large sensory bias. However, unlike us, Sánchez-Fuenzalida et al. used a reproduction task to distinguish sensory from decisional biases and did not examine the dynamics of these biases over time.

The distinct dynamics of sensory and decisional biases provide critical insights into the progression of perceptual judgments. Our study indicates that repeated perceptual judgments in general may involve a decisional bias that progressively counteracts the sensory bias. While the perceptual judgments change over time, it is plausible that there is no change over time in the sensory cortex, but with changes in decisional areas in the parietal/frontal cortex. In fact, this possibility may explain previous findings where repeated visual adaptation reduced the aftereffects measured by behavioral judgments but without observing changes in the sensory cortex (Dong et al., 2016, 2020). More generally, our approach allows follow-up studies to map the neural correlates of sensory and decisional biases.

We interpret our results as showing a dynamic decisional bias adjustment. Another way of thinking about our results is that participants learned to dissociate the experienced illusion (illusory distortions) from the physical face distortions. We believe that these two ways of describing our results are equivalent to each other: By changing their decisional bias in the illusion condition, participants are in effect downweighing the influence of the illusory percepts. Critically, it is unlikely that participants changed the task they were performing. Indeed, they were informed about the existence of the illusion during the instructions, so they were aware of their task from the very beginning. Therefore, the effects we observe cannot be due to an initial confusion about the nature of the task and instead are in line with decisional adjustments that participants were able to make over the course of the experiment.

Our findings suggest that participants gradually changed their decisional bias for the illusion but not for the control condition. Therefore, in the second half of Experiment 1, participants were maintaining different sets of decision criteria for the illusion and control conditions. This may seem surprising as several previous studies have suggested that people may struggle with maintaining two independent sets of criteria for interleaved conditions (Gorea & Sagi, 2000, 2002; Gorea et al., 2005). Nevertheless, follow-up research has demonstrated that while separately maintained sets of criteria undergo a process of "criterion attraction," they can nonetheless remain strongly differentiated (Rahnev, 2021a; Zak et al., 2012). Our experiment made it especially easy to maintain separate sets of decision criteria because of the immediately obvious difference between illusion trials (consisting of 15 flashes) and control trials (consisting of a single flash). Furthermore, the long stimulus presentation in the illusion condition afforded participants additional time to adjust their criteria. Researchers planning to conduct studies on the dynamics of decisional bias should be aware of the problem of criterion attraction and ensure that their design makes it easy for participants to maintain separate sets of decision criteria for different conditions.

Supplementary information The online version contains supplementary material available at https://doi.org/10.3758/s13423-024-02539-8.

Acknowledgements We thank Minzhi Wang for his help with data collection and Corey J. Breeland for his help with creating distorted faces.

Author contributions Yi Gao and Dobromir Rahnev conceived the experiment; Yi Gao programmed and conducted the experiment; Yi Gao, Sixing Chen, and Dobromir Rahnev wrote the manuscript.

Funding This work was supported by the National Institute of Health (award: R01MH119189) and the Office of Naval Research (award: N00014-20–1-2622).

Data availability All data and codes are available at: https://osf.io/h3jkn/.

Declarations

Conflicts of interest The authors have no competing interests to declare that are relevant to the content of this article.

Ethics approval The study was approved by the Institutional Review Board of the Georgia Institute of Technology.

Consent to participate All participants provided written consent to participate in the study.

Consent for publication All participants signed informed consent to publish their data.

References

Balas, B., & Pearson, H. (2019). The flashed face distortion effect does not depend on face-specific mechanisms. *Scientific Reports*, *9*(1). https://doi.org/10.1038/s41598-018-37991-9

Bang, J. W., & Rahnev, D. (2017). Stimulus expectation alters decision criterion but not sensory signal in perceptual decision making. *Scientific Reports, 7*(1), 17072. https://doi.org/10.1038/s41598-017-16885-2

Bowden, J., Whitaker, D., & Dunn, M. J. (2019). The role of peripheral vision in the flashed face distortion effect. *Perception, 48*(1), 93–101. https://doi.org/10.1177/0301006618817419

Bowen, H. J., Marchesi, M. L., & Kensinger, E. A. (2020). Reward motivation influences response bias on a recognition memory task. *Cognition*, *203*, Article 104337. https://doi.org/10.1016/j.cognition.2020.104337

Chen, S., & Rahnev, D. (2023). Signatures proposed to index perceptual effects emerge in a purely cognitive task. *PsyArXiv*.

Dong, X., Du, X., & Bao, M. (2020). Repeated contrast adaptation does not cause habituation of the adapter. *Frontiers in Human Neuroscience*,. https://doi.org/10.3389/fnhum.2020.589634

Dong, X., Gao, Y., Lv, L., & Bao, M. (2016). Habituation of visual adaptation. *Scientific Reports*, *6*(1), Article 19152. https://doi.org/10.1038/srep19152

Frund, I., Wichmann, F. A., & Macke, J. H. (2014). Quantifying the effect of intertrial dependence on perceptual decisions. *Journal of Vision, 14*(7), 9–9. https://doi.org/10.1167/14.7.9

Gallagher, R. M., Suddendorf, T., & Arnold, D. H. (2019). Confidence as a diagnostic tool for perceptual aftereffects. *Scientific Reports*, *9*(1), Article 7124. https://doi.org/10.1038/s41598-019-43170-1

Gallagher, R. M., Suddendorf, T., & Arnold, D. H. (2021). The implied motion aftereffect changes decisions, but not confidence. *Attention, Perception, & Psychophysics, 83*(8), 3047–3055. https://doi.org/10.3758/s13414-021-02331-z

Gao, Y., Pieller, J., Webster, M. A., & Jiang, F. (2022). Temporal dynamics of face adaptation. *Journal of Vision*, *22*(11), Article 14. https://doi.org/10.1167/jov.22.11.14

Gao, Y., Wang, M., & Rahnev, D. (2024). Objectively quantifying subjective phenomena: Measuring the flashed face distortion effect. *Cognition*, *250*, Article 105861. https://doi.org/10.1016/j.cognition.2024.105861

Gao, Y., Webster, M. A., & Jiang, F. (2019). Dynamics of contrast adaptation in central and peripheral vision. *Journal of Vision*, *19*(6), Article 23. https://doi.org/10.1167/19.6.23

Gao, Y., Webster, M. A., & Jiang, F. (2021). Changes of tuning but not dynamics of contrast adaptation with age. *Vision Research, 187*, 129–136. https://doi.org/10.1016/j.visres.2021.03.015

Germar, M., Schlemmer, A., Krug, K., Voss, A., & Mojzisch, A. (2014). Social influence and perceptual decision making. *Personality and Social Psychology Bulletin, 40*(2), 217–231. https://doi.org/10.1177/0146167213508985

Gold, J. I., & Shadlen, M. N. (2007). The neural basis of decision making. *Annual Review of Neuroscience, 30*(1), 535–574. https://doi.org/10.1146/annurev.neuro.29.051605.113038

Goldstone, R. L. (1998). Perceptual learning. *Annual Review of Psychology, 49*(1), 585–612. https://doi.org/10.1146/annurev.psych.49.1.585

Gorea, A., Caetta, F., & Sagi, D. (2005). Criteria interactions across visual attributes. *Vision Research, 45*(19), 2523–2532. https://doi.org/10.1016/j.visres.2005.03.018

Gorea, A., & Sagi, D. (2000). Failure to handle more than one internal representation in visual detection tasks. *Proceedings of the National Academy of Sciences, 97*(22), 12380–12384. https://doi.org/10.1073/pnas.97.22.12380

Gorea, A., & Sagi, D. (2002). Natural extinction: A criterion shift phenomenon. *Visual Cognition, 9*(8), 913–936. https://doi.org/10.1080/13506280143000638

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. John Wiley.

Jazayeri, M., & Movshon, J. A. (2007). A new perceptual illusion reveals mechanisms of sensory decoding. *Nature, 446*(7138), 912–915. https://doi.org/10.1038/nature05739

Krekelberg, B. (2022). *BayesFactor* (2.3.0). Zenodo. 10.5281/zenodo.7006300

Linares, D., Aguilar-Lleyda, D., & López-Moliner, J. (2019). Decoupling sensory from decisional choice biases in perceptual decision making. *ELife*, *8*. https://doi.org/10.7554/eLife.43994

Lundqvist, D., Flykt, A., & Ohman, A. (1998). *The Karolinska directed emotional faces* (KDEF) [CD-ROM]. Department of Clinical Neuroscience, Psychology section, Karolinska Institutet.

Maldonado Moscoso, P. A., Cicchini, G. M., Arrighi, R., & Burr, D. C. (2020). Adaptation to hand-tapping affects sensory processing of numerosity directly: Evidence from reaction times and confidence. *Proceedings of the Royal Society B: Biological Sciences*, *287*(1927), Article 20200801. https://doi.org/10.1098/rspb.2020.0801

Mostert, P., Kok, P., & de Lange, F. P. (2015). Dissociating sensory from decision processes in human perceptual decision making. *Scientific Reports*, *5*(1), Article 18253. https://doi.org/10.1038/srep18253

Norton, E. H., Fleming, S. M., Daw, N. D., & Landy, M. S. (2017). Suboptimal criterion learning in static and dynamic environments. *PLOS Computational Biology*, *13*(1), Article e1005304. https://doi.org/10.1371/journal.pcbi.1005304

Odgaard, E. C., Arieh, Y., & Marks, L. E. (2003). Cross-modal enhancement of perceived brightness: Sensory interaction versus response bias. *Perception & Psychophysics, 65*(1), 123–132. https://doi.org/10.3758/BF03194789

Rahnev, D. (2021a). A robust confidence–accuracy dissociation via criterion attraction. *Neuroscience of Consciousness*, *2021*(1). https://doi.org/10.1093/nc/niab039

Rahnev, D. (2021b). Response bias reflects individual differences in sensory encoding. *Psychological Science, 32*(7), 1157–1168. https://doi.org/10.1177/0956797621994214

Rahnev, D., & Denison, R. N. (2018). Suboptimality in perceptual decision making. *Behavioral and Brain Sciences*, *41*. https://doi.org/10.1017/S0140525X18000936

Rungratsameetaweemana, N., Itthipuripat, S., Salazar, A., & Serences, J. T. (2018). Expectations do not alter early sensory processing during perceptual decision-making. *The Journal of Neuroscience, 38*(24), 5632–5648. https://doi.org/10.1523/JNEUROSCI.3638-17.2018

Rungratsameetaweemana, N., & Serences, J. T. (2019). Dissociating the impact of attention and expectation on early sensory processing. *Current Opinion in Psychology, 29*, 181–186. https://doi.org/10.1016/j.copsyc.2019.03.014

Sánchez-Fuenzalida, N., van Gaal, S., Fleming, S. M., Haaf, J. M., & Fahrenfort, J. J. (2023). Predictions and rewards affect decision-making but not subjective experience. *Proceedings of the National Academy of Sciences*, *120*(44). https://doi.org/10.1073/pnas.2220749120

Shams, L., & Kim, R. (2010). Crossmodal influences on visual perception. *Physics of Life Reviews, 7*(3), 269–284. https://doi.org/10.1016/j.plrev.2010.04.006

Shteingart, H., & Loewenstein, Y. (2014). Reinforcement learning and human behavior. *Current Opinion in Neurobiology, 25*, 93–98. https://doi.org/10.1016/j.conb.2013.12.004

Starns, J. J., Ratcliff, R., & White, C. N. (2012). Diffusion model drift rates can be influenced by decision processes: An analysis of the strength-based mirror effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*(5), 1137–1151. https://doi.org/10.1037/a0028151

Tangen, J. M., Murphy, S. C., & Thompson, M. B. (2011). Flashed face distortion effect: Grotesque faces from relative spaces. *Perception, 40*(5), 628–630. https://doi.org/10.1068/p6968

Voss, A., Rothermund, K., & Brandtstädter, J. (2008). Interpreting ambiguous stimuli: Separating perceptual and judgmental biases. *Journal of Experimental Social Psychology, 44*(4), 1048–1056. https://doi.org/10.1016/j.jesp.2007.10.009

Webster, M. A. (2015). Visual adaptation. *Annual Review of Vision Science, 1*(1), 547–567. https://doi.org/10.1146/annurev-vision-082114-035509

Webster, M., & Maclin, O. H. (1999). Figural aftereffects in the perception of faces. *Psychonomic Bulletin & Review, 6*(4), 647–653. https://doi.org/10.3758/bf03212974

Wei, S., Xie, Y., & Rahnev, D. (2021). Inferring serial correlation with dynamic backgrounds. *Proceedings of the 38th International Conference on Machine Learning*, 11047–11057.

Wen, T., & Kung, C. C. (2014). Using functional magnetic resonance imaging to explore the flashed face distortion effect. *Journal of Vision*, *14*(12). https://doi.org/10.1167/14.12.29

White, C. N., & Poldrack, R. A. (2014). Decomposing bias in different types of simple decisions. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*(2), 385–398. https://doi.org/10.1037/a0034851

Witt, J. K., Taylor, J. E. T., Sugovic, M., & Wixted, J. T. (2015). Signal detection measures cannot distinguish perceptual biases from response biases. *Perception, 44*(3), 289–300. https://doi.org/10.1068/p7908

Witthoft, N., Sha, L., Winawer, J., & Kiani, R. (2018). Sensory and decision-making processes underlying perceptual adaptation. *Journal of Vision, 18*(8), 10. https://doi.org/10.1167/18.8.10

Yamashita, J. A., Hardy, J. L., De Valois, K. K., & Webster, M. A. (2005). Stimulus selectivity of figural aftereffects for faces. *Journal of Experimental Psychology: Human Perception and Performance, 31*(3), 420–437. https://doi.org/10.1037/0096-1523.31.3.420

Yu, A. J., & Cohen, J. D. (2008). Sequential effects: Superstition or rational behavior? *Advances in Neural Information Processing Systems, 21*, 1873–1880.

Zak, I., Katkov, M., Gorea, A., & Sagi, D. (2012). Decision criteria in dual discrimination tasks estimated using external-noise methods. *Attention, Perception, & Psychophysics, 74*(5), 1042–1055. https://doi.org/10.3758/s13414-012-0269-0