



The push–pull of serial dependence effects: Attraction to the prior response and repulsion from the prior stimulus

Patrick Sadil¹ · Rosemary A. Cowell^{2,3} · David E. Huber²

Accepted: 30 May 2023 / Published online: 11 August 2023
© The Psychonomic Society, Inc. 2023

Abstract

In the “serial dependence” effect, responses to visual stimuli appear biased toward the last trial’s stimulus. However, several kinds of serial dependence exist, with some reflecting prior stimuli and others reflecting prior responses. One-factor analyses consider the prior stimulus alone or the prior response alone and can consider both variables only via separate analyses. We demonstrate that one-factor analyses are potentially misleading and can reach conclusions that are opposite from the truth if both dependencies exist. To address this limitation, we developed two-factor analyses (model comparison with hierarchical Bayesian modeling and an empirical “quadrant analysis”), which consider trial-by-trial combinations of prior response and prior stimulus. Two-factor analyses can tease apart the two dependencies if applied to a sufficiently large dataset. We applied these analyses to a new study and to four previously published studies. When applying a model that included the possibility of both dependencies, there was no evidence of attraction to the prior stimulus in any dataset, but there was evidence of attraction to the prior response in all datasets. Two of the datasets contained sufficient constraint to determine that both dependencies were needed to explain the results. For these datasets, the dependency on the prior stimulus was repulsive rather than attractive. Our results are consistent with the claim that both dependencies exist in most serial dependence studies (the two-dependence model was not ruled out for any dataset) and, furthermore, that the two dependencies work against each other.

Keywords Serial dependence · Tilt aftereffect · Priming · Sequence effects

Recently encountered visual information alters perceptual decisions. Often, the alteration manifests as a repulsion from the prior stimulus, as in perceptual phenomena like the tilt aftereffect (Gibson & Radner, 1937; Webster, 2015). More recently, the opposite pattern has been reported (Cicchini et al., 2014; Fischer & Whitney, 2014; Kiyonaga et al., 2017), in which the influence of the prior stimulus on current perceptual judgments appears to be attractive. That is, when participants are presented with a series of stimuli of the same type (e.g., gratings that differ in orientation across trials, Fig. 1, Left), responses to the current stimulus appear biased *toward* the orientation of the stimulus on

the immediately preceding trial (Fig. 1, Right). This attractive effect has been labeled “serial dependence” (Fischer & Whitney, 2014). Like repulsive aftereffects, attractive serial dependencies have been observed for a wide range of stimulus classes, including orientation, spatial location, motion direction, numerosity, timing, identity, gaze direction, ensemble statistics, attractiveness, and gender (Alais et al., 2018; Fischer et al., 2020; Fornaciai & Park, 2018b; Jepma et al., 2014; Liberman et al., 2014; Suárez-Pinilla et al., 2018; Taubert et al., 2016a, b; Xia et al., 2016), suggesting that the effect reflects fundamental mechanisms of perception.

The attractive serial dependence effect has been explained as reflecting a “continuity field” (Fischer & Whitney, 2014), which promotes visual stability by biasing current perceptual representations toward the recent past through a (potentially weighted) averaging mechanism (Alais et al., 2018; Cicchini et al., 2018; Fischer & Whitney, 2014). However, it is unclear whether this attraction is towards the low-level visual properties of the prior stimulus or towards the higher-level percept of the prior stimulus (which manifests as attraction

✉ Patrick Sadil
psadil@umass.edu

¹ Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD USA

² Institute of Cognitive Science, University of Colorado Boulder, Boulder, CO, USA

³ Department of Psychology and Neuroscience, University of Colorado Boulder, Boulder, CO, USA

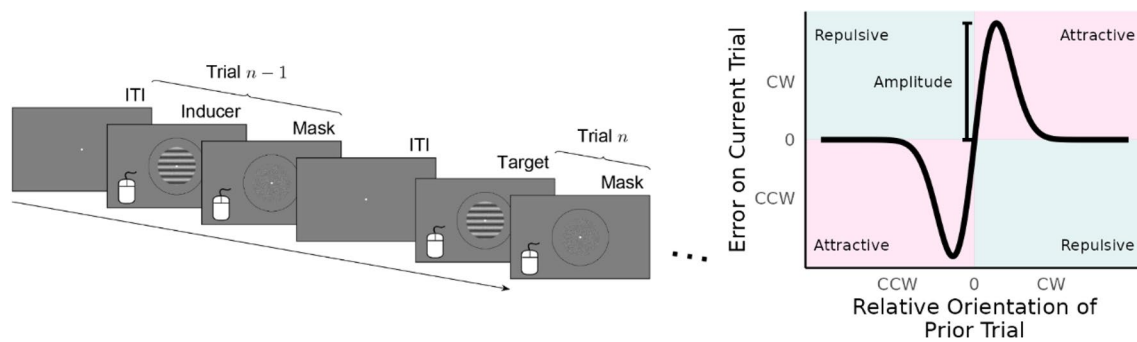


Fig. 1 Left: Experiment 1 Schematic. On each trial, participants were presented with a grating stimulus and asked to report its orientation. Responses could begin when the stimulus appeared, but the stimulus was replaced by a mask after 200 ms. Stimuli are not to scale. ITI = intertrial interval. Right: Graphical Depiction of Serial Dependence. Errors on the current trial are plotted as a function of the orientation difference between the current trial and the prior trial. If the

error is in the same rotational direction as the orientation difference (e.g., both Clockwise/CW or both Counterclockwise/CCW), this suggests attraction, whereas if they are of opposite rotations, this suggests repulsion, as indicated by the colored quadrants. The bias effect typically operates within a restricted range of differences between the current trial and the prior trial

towards the prior response, in the case of a perceptual error on the prior trial). Adjudicating between these two possibilities is difficult because the response on the current trial can be influenced by many factors (see also Jepma et al., 2014; Jones et al., 2013) that sum together in a complicated mixture to influence behavior. In addition, outside of the serial dependence literature, there is substantial evidence that a response on one trial can make participants more likely to make a similar response on the next trial (Annis et al., 2018; Jesteadt et al., 1977), an effect that we refer to as “response hysteresis” (Schwiedrzik et al., 2014). We use this term for any observed attraction to the prior response, while noting that it could arise from a variety of psychological mechanisms, including changes in beliefs about the base rate of stimuli (Zhang et al., 2014; cf. Urai et al., 2019), systematic changes in how much evidence participants require for initiating responses (Wagenmakers et al., 2004), a tendency to repeat decisions without regard to the stimulus (Akaishi et al., 2014; Braun et al., 2018), or one interpretation of continuity field theory (in which the response simply provides an index of the participant’s higher-level percept). But whatever the source of response hysteresis, its existence poses a problem when drawing conclusions about serial dependence effects: If the response to the prior trial is accurate, then attraction to the prior stimulus and response hysteresis are confounded.

The conflation of attraction to the prior stimulus and response hysteresis is exacerbated by standard analysis practices, one-factor analyses that consider the prior stimulus alone or the prior response alone, or consider both factors but in separate analyses. In a standard one-factor analysis, the magnitude of dependence is determined by non-parametrically smoothing the errors or by fitting the derivative of a Gaussian function (Fig. 1, Right). The derivative is a

convenient model since it has an amplitude parameter that corresponds to the maximum average error elicited by prior stimuli, which is taken as the magnitude of serial dependence. We use this derivative function as a tool to highlight the kinds of biases in the current trial that would arise from the summation of two dependencies with opposite signs, one whose amplitude is positive (Fig. 2A, first column), and one whose amplitude is negative (Fig. 2A, second column). A typical one-factor analysis—plotting a dependence as a function of either the prior orientation or prior response separately—shows only a summation of these two latent influences, with the summation influenced by participants’ response variability (Fig. 2A, third column). For example, if the two forces mirrored each other perfectly and participants were relatively accurate, the latent influences would combine to produce a flat line—an observed lack of dependence despite robust latent dependencies (Fig. 2A, bottom row).

The example in Fig. 2A, bottom row, demonstrates just how misleading one-factor analyses can be. In this example, a one factor-analysis based on the prior stimulus produces a null result. In addition, a one factor-analysis based on the prior response produces a null result. However, both conclusions are incorrect. In reality, there is both an effect of the prior stimulus and the prior response. Thus, it is potentially misleading to use a one-factor analysis, and yet all prior studies in the serial dependence literature are based one-factor analyses. Furthermore, the situation in Fig. 2A is not uncommon. To preview our results, of the five datasets that we analyzed, two of them produced seemingly null results for one-factor analyses based on only the prior stimulus or prior response and yet a two factor-analysis revealed this to be caused by offsetting dependencies. Furthermore, the results of all 5 datasets were consistent with offsetting dependencies.

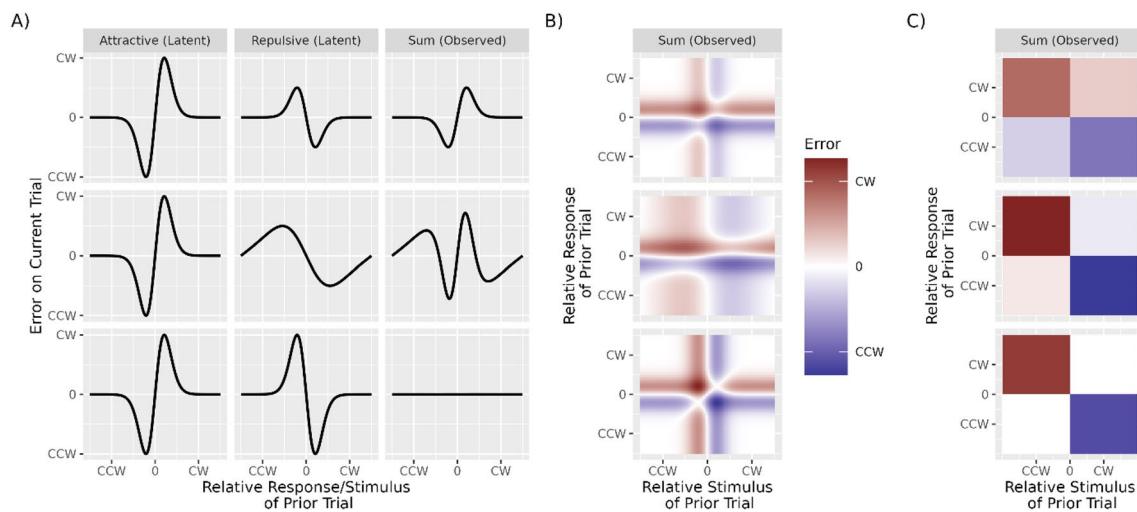


Fig. 2 Analysis of serial dependence effects, showing a traditional one-factor analysis (**A**) and a two-factor analysis (**B**) that examines combinations of prior response and prior stimulus. **A**) Hypothetical results of a serial dependence experiment, in which the prior stimulus induces two effects simultaneously. The first column shows an attractive dependence, the second a repulsive dependence, and these are summed to give the third column, the observed dependence. The repulsive effect differs across the rows, while the attractive effect is the same in all three rows. In each column, the x -axis of the analyzed dependence might reflect the relative stimulus of the prior trial or the relative response of the prior trial. However, the observed results would only be a simple sum of the two effects (third column) if both dependencies are of the same type (both prior stimulus, or both prior response) or in the special case that they are different (one prior stimulus, one prior response) but accuracy is perfect, making them equiv-

alent. **B**) Predicted error for all combinations of prior stimulus and prior response. This two-factor analysis allows the identification of two dependencies when one is caused by prior stimulus and the other by prior response. The plots in (**B**) show all combinations of the prior response, drawn from the first column of (**A**), and prior stimulus, drawn from the second column of (**A**), with these two effects summing to produce a predicted error. When the prior trial is accurate, the combination corresponds to the predictions along the diagonal of the plots in (**B**), which is equivalent to the summation plots in the third column of (**A**). The colors indicate the sign and magnitude of the predicted error for the current trial. **C**) Predicted error for combinations of prior response and stimulus, averaged to aid visualization. A typical experiment will not contain every possible combination of prior stimulus and response, but key information is retained by averaging errors within quadrants. (Color figure online)

This problem is widely acknowledged, and several studies have sought to isolate separate dependencies experimentally (Bae & Luck, 2020; Cicchini et al., 2017; Fischer & Whitney, 2014; Fornaciai & Park, 2019; Fritsche et al., 2017; Pascucci et al., 2019b; Pascucci & Plomp, 2021; Suárez-Pinilla et al., 2018), but these studies often produce contradictory results. This might reflect a common challenge: the psychological operations that contribute to different sequential effects, like attraction to the prior stimulus or response hysteresis, are not easily manipulated independently. For example, one approach instructs participants to withhold their responses on a subset of trials, asking whether the attraction persists on trials after which there was no response (Fischer & Whitney, 2014; Fornaciai & Park, 2019; Pascucci et al., 2019b; Pascucci & Plomp, 2021). However, a lack of a motor response does not imply a lack of a perceptual decision (Pascucci et al., 2019b), and the prior decision may be sufficient to cause dependencies (Akaishi et al., 2014). To reduce the role of unrealized decisions, some studies instructed participants to ignore the stimulus (Pascucci et al., 2019b) or informed them that prior stimuli were irrelevant to the task (Fornaciai & Park,

2018a), but this introduces a confound with attention, and the continuity field is believed to be sensitive to how well participants attend to the prior stimulus (Fischer et al., 2020; Fischer & Whitney, 2014). To assess attention, some researchers have asked participants to report on a subset of the otherwise ignorable stimuli (Pascucci et al., 2019b; Pascucci & Plomp, 2021), but then it is difficult to guarantee that attentional checks do not reintroduce unrealized decisions. In summary, although it is straightforward to manipulate participants' responses, it is more challenging to manipulate their decisions while also equating every aspect of perceptual processing.

Given the difficulties inherent to isolating dependencies experimentally, we explored whether dependencies could be isolated statistically. The standard serial dependence paradigm, in which participants simply report the stimulus on each trial, may in some cases provide sufficient constraint to untangle the effect of the prior stimulus from the effect of the prior response, provided that (1) the analysis considers combinations of prior response and prior stimulus (i.e., a two-factor analysis) and (2) there is sufficient trial-by-trial variability in these combinations (e.g., participants'

accuracy varies across trials). In the hypothetical scenario of Fig. 2A, if one dependency (e.g., first or second column) is driven by the prior response and the other is driven by the prior stimulus (e.g., the other column), a one-factor analysis of either the prior stimulus or prior response is one that collapses over the other factor. Rather than collapsing over one factor, errors can be assessed across all possible combinations of prior stimulus and prior response (Fig. 2B) with an explicit model, or the model's results can be compared with a model-free approach (e.g., averaging across combinations of prior stimulus and response, as in Fig. 2C). These two-factor analyses can provide insight into cases where the one-factor approach fails. For instance, in the scenario where the two effects perfectly mirror each other (third row of Fig. 2), the trials for which the prior response is not perfectly accurate exhibit systematic errors in the current trial that can be used to identify the two sequential dependencies (Fig. 2B–C).

We developed a two-factor, statistical modeling approach that can identify two dependencies, where the combinations of prior stimulus and response arise from natural variations in trial-by-trial accuracy. The approach first uses model comparison to determine whether a dataset is constraining enough to identify multiple dependencies. If there is sufficient constraint, the parameters of the model that is best supported by the data can be inspected to determine the direction of those dependencies. We applied this framework to a new experiment that collected considerably more data than a typical serial dependency study. We then applied the technique to previously published datasets (Fischer & Whitney, 2014; Pascucci et al., 2019b; Samaha et al., 2019).

Methods

Experiment 1

As discussed above, disentangling stimulus and response effects requires variability between prior responses and prior stimuli. To gain the data needed to tease apart a dependency on the prior stimulus from one arising from the prior response, we used a challenging orientation judgment task (i.e., performance was not perfectly accurate) with a high trial count (i.e., good statistical power in terms of combinations of prior stimulus and prior response). The stimulus parameters and experimental procedure followed the design of Samaha et al. (2019). This task involved briefly presented stimuli corrupted by noise. As such, the results of this study may not generalize to all serial dependence tasks, in which the stimuli are uncorrupted by noise and participants are allowed to view the stimuli for extended durations. Hence, we additionally analyzed four published datasets that used a range of protocols (described below).

Participants

Power analyses indicated that approximately thirteen participants would be sufficient both to detect a single dependence and to reliably distinguish between one versus two dependencies (Appendix B). To allow for the possibility that data from some participants would be unusable, sixteen participants were run, including author PS, with all except PS being awarded course credit. All participants had normal or corrected-to-normal vision, and all provided usable data. The procedure was approved by the University of Massachusetts Amherst Institutional Review Board.

Stimulus parameters

Stimuli were presented on an LCD monitor (ASUS VG248QE, 1,920 × 1,080 cm, 100 Hz refresh rate, 1,920 × 1,080 resolution), viewed from approximately 60.96 cm. Stimuli were displayed using the Psychophysics Toolbox (Version 3.0.14; Brainard, 1997; Pelli, 1997) and custom MATLAB code (MathWorks, 2015).

Throughout the experiment, participants fixated on a light gray dot (0.08°). Grating stimuli (sine wave with 1.5 cycles per degree and phase 0 subtending a circular region of 2°) were presented on a medium gray background. Mask stimuli consisted of white noise rendered at 100% contrast. Participants were cued to make orientation reproduction responses with a circle (6° radius) and made the response by clicking a mouse near the circle (within 50 pixels of the circle).

The signal-to-noise ratio of the grating was reduced by averaging the grating with white noise (Samaha et al., 2016, 2019). The contrast of the grating was determined by a pilot study (three participants, data not shown), in which white noise (100% contrast) was averaged with a grating stimulus presented at a range of contrast levels. The main experiment used a contrast (10%) that elicited responses that were within $\pm 25^\circ$ of the true orientation approximately 80% of the time.

Procedure

The trial structure is illustrated in Fig. 1 (Left). Each trial began with the presentation of a grating, surrounded by a circle. To encourage fixation, a centrally presented dot was visible continuously. The grating was replaced by a mask after 200 ms, but participants could initiate responses immediately after target onset.¹ The mask remained on the

¹ Although participants could initiate responses immediately, responses were rarely finalized before the mask appeared. Across participants, over 99% of response times were longer than 200 ms, meaning that the stimulus was no longer visible at the time when the response was made for over 99% of trials.

screen until participants finished their response. A circle surrounded the grating and mask, and participants responded by using a computer mouse to click on the circle. They were instructed to report the orientation of the grating stimulus by clicking a point on the circle such that an imagined line connecting the center of the grating (i.e., the fixation point) to the clicked point would be parallel to the stripes of the grating. They first practiced the task for 10 trials with an experimenter present and available to answer questions about the task. Participants could respond on either side of the circle. Trials were separated with a variable fixation period (randomly determined on each trial with a draw from a discrete uniform distribution ranging from 300–500 ms in steps of 20 ms).

The orientation on each trial was drawn at random from a discrete, uniform distribution, sampling integers between 0–179°. Participants completed 15 blocks, each with 101 trials, yielding 1,515 trials per participant, and 24,240 trials in total. The median duration of the experiment was 48 minutes (range: 34–86).

Published datasets

Four published datasets were reanalyzed (Fischer & Whitney, 2014, Experiment 1b; Pascucci et al., 2019a, Experiments 1 and 2; Samaha et al., 2019). See Table 1 for an overview of the methods. Note that, in the original publications, the data were preprocessed to a different extent than in our analyses. Preprocessing in the original studies included (A) excluding outlier responses (Pascucci et al., 2019b; Samaha et al., 2019), (B) centering each participant's errors to remove any overall clockwise or counterclockwise bias in that individual (Pascucci et al., 2019b; Samaha et al., 2019), and/or (C) regressing out biases caused by the orientation on the current trial (Pascucci et al., 2019b), whereby participants respond less accurately for certain orientations, regardless of the stimuli or responses preceding those orientations (e.g., Jastrow, 1892; Wei & Stocker, 2015). Preprocessing steps (B) and (C) were built into our main analyses (model comparison, described below) by including regressors for these effects. Of note, any bias for particular orientations can cause a spurious dependence on the prior response, and so they must be considered when modeling response dependencies (Appendix A; Fritsche, 2016).

Analyses

All datasets were analyzed in three ways: (1) Different hierarchical Bayesian models were fit to the data and compared with each other, assessing whether the data were sufficiently constraining to distinguish between effects of the prior orientation and prior response (a two-factor analysis); (2)

standard one-factor serial dependence analyses (e.g., plotting smoothed errors) based on the raw data (these plots will not necessarily match the previously published plots, which were based on pre-processed data); and (3) a nonparametric quadrant analysis (a two-factor analysis) that sorts trials into one of four quadrants depending on the relative direction of the prior stimulus and prior response. When reported, confidence intervals are across participants (Morey, 2008).

Bayesian analyses

Three hierarchical Bayesian models were fit to each dataset. The models allowed for one of the three combinations of effects of the prior orientation and prior response (i.e., two models with only one dependency and one model with both dependencies). All three models also included terms to capture rotational bias (i.e., a participant's tendency to err in a clockwise or counterclockwise direction by a small amount on all trials) and biases towards particular salient angles. The latter was parameterized as either a bias towards the nearest cardinal angle (e.g., 0° or 90°), and thus a bias away from oblique angles, or a bias towards oblique angles (e.g., 45°, 135°), and thus away from cardinal angles, with the latter possibility termed “the oblique effect” in the literature (Jastrow, 1892; Wei & Stocker, 2015). Here, we provide a brief description of the Bayesian models that were used to analyze the data. A more detailed description of the models, including how they handle preferences for oblique/cardinal angles, is given in Appendix A.

The models were additive, assuming that error on the current trial reflected the summation of any dependencies from the prior trial along with biases that did not depend on the prior trial (i.e., rotational and periodic biases). One model included a dependence on the prior stimulus, a second included a dependence on the prior response, and a third included both dependencies. Added together, the serial dependence biases and the cardinal/oblique/rotational biases determined the average error as a function of the current orientation and the stimulus/response from the prior trial. Trial-by-trial variability in this average was modeled with a normal distribution.

Serial dependencies in orientation were modeled using the derivative of the density function for a circular normal distribution². All models were estimated hierarchically, allowing for parameters to vary by participant while simultaneously allowing the participant-specific parameters to constrain each other via a “population-level” distribution (Kruschke, 2015, Chapter 8). The population-level distribution can be informally understood as an average across participants—an estimate produced by analyzing participants

² A circular distribution was used to account for the circular nature of the stimuli; when the errors peak far enough away from 0, a typical Normal distribution would have a substantial discontinuity at $\pm 90^\circ$.

Table 1 Protocols for studies of serial dependence in orientation judgments

Study	Experiment	Inducer duration	Inducer eccentricity	Inducer-target position	Inducer contrast	Inducer SF	Mask duration	Mask-to-response delay	ITI average ^a	Trials
Current Study	1	200	0	Same	10 ^b	1.5	0	0	400	24240
Fischer Whitney (2014)	1b ^c	500	6.5	Same	25	0.33	1000	250	2000	3296
Pascucci et al. (2019a)	1	400	8.5	Varied	50	0.5	400	500	500	5600
Pascucci et al. (2019b)	2	400	0	Same	50	1.2	400	500	500	4400
Samaha Switzky Postle (2019)	1	33	0	Same	8.5 ^{b,d}	1.5	0	6300 ^e	700	6000

Eccentricity distance in degrees visual angle from fixation to the center of the stimuli; *Contrast* Percent Michelson; *SF* Spatial Frequency, cycles per degree; *ITI* Inter-trial Interval. All timing is reported in milliseconds

^aIn some experiments ‘Average ITI’ includes both the ITI and a period of cueing at the start of the target trial. In all experiments, the reported ITI does not include participants’ response time

^bContrast refers to Michelson contrast of gratings, which were convolved with white noise

^cIn the original experiment, this was listed as a control experiment for experiment 1. In experiment 1, the orientations on each trial were explicitly counterbalanced, whereas in 1b they were randomized

^dThe reported amplitude averages across two conditions. The conditions kept the signal-to-noise ratio of the stimuli equal, but halved the grating and noise contrasts

^eThis is an average across five conditions (600, 3450, 6300, 9150 and 12000). Authors observed that delay did not have a substantial effect

separately and then summarizing those individual analyses. Analyzing them hierarchically rather than separately accounts for individual differences while simultaneously pooling information across participants to gain greater reliability.

Models were constructed and fit using the Stan language (Carpenter et al., 2016), using its interface (RStan, 2.18.2) for the R computing language (R Core Team). Stan draws samples from an approximation of the posterior distribution using a modified version of the Hamiltonian Monte Carlo algorithm (Duane et al., 1987; Hoffman & Gelman, 2014). The validity of this approximation was assessed in two ways. First, chains were monitored for divergences, an indication that the numerical simulation methods in the algorithm are compromised (Betancourt, 2017). Second, the split- \hat{R} (“split r-hat”) for each parameter was calculated (Gelman & Rubin, 1992). In all results, there were no divergences, and for each parameter, the split- \hat{R} was below 1.1 (Gelman et al., 2013, Chapter 14).

Model comparison was based on the models’ predictive abilities with an approximation to a leave-one-out (LOO) cross-validation score (Vehtari et al., 2017, 2015). Specifically, models were compared based on their Pareto-smoothed, importance sampling leave-one-out plus (PSIS-LOO+) score, as calculated with the loo software package (Vehtari et al., 2020).

Median window

In addition to the two-factor Bayesian modeling, each dataset was also analyzed with a one-factor analysis (e.g.,

Fig. 2A), a median window based on either the prior stimulus or response. The median error in the current trial for each participant was calculated using a sliding, median window, centered on 200 equally spaced orientation differences between -90° and 90° (width of $\pm 12^\circ$, following Samaha et al., 2019). The smoothed error for orientation differences was calculated between both (1) the prior and current orientations and (2) the response to the prior stimulus and the current orientation. To eliminate boundary artifacts, windows near $\pm 90^\circ$ included circularly wrapped copies of the data (e.g., data from stimulus orientations of $78-89^\circ$ were considered to be adjacent to an orientation of -90°). The across-participant distribution of median errors was then used to calculate confidence intervals.

Quadrant plots

The hierarchical Bayesian models make parametric assumptions about the functional form of the underlying serial dependencies (i.e., circular normal distributions) to increase statistical power. To check that the data show features consistent with the conclusions from the model comparison, we also developed a nonparametric two-factor quadrant plot analysis that can be used to visualize the data in terms of separate dependencies. This was achieved by sorting each trial into one of the four categories for the combinations of whether the prior stimulus was clockwise or counterclockwise as compared with the current stimulus, and whether the prior response was clockwise or counterclockwise as compared with the current stimulus, as shown in Fig. 2C. For each of these four possible prior stimulus/response

categories, current trial errors were averaged within participants, and confidence intervals were calculated across the averages.

For the quadrant plots, if the prior stimulus is the only dependence, there should be a main effect of prior stimulus in the 2×2 grid of quadrants (as seen in Fig. 3A, repulsion from the prior stimulus results in a clockwise bias for the left column and a counterclockwise bias for the right column). If the prior response is the only dependence, there should be a main effect of prior response (as seen in Fig. 3B, attraction to the prior response results in a clockwise bias for the top row and a counterclockwise bias for the bottom row). If both dependencies exist, there should be an interaction. For instance, if the two dependencies are in opposite directions, the two “accurate” quadrants (upper-right and lower-left) will not display much of a bias because the dependencies counteract each other, whereas the two “inaccurate” quadrants (upper-left and lower-right) will show biases, with one clockwise and the other counterclockwise (see Fig. 3C). In contrast, if the two dependencies are in the same direction, the two accurate quadrants will show a bias, with one clockwise and the other counterclockwise, while the two inaccurate quadrants will not display much of a bias because the

two dependencies counteract each other (see Fig. 3D). Thus, the quadrant analysis can display whether the data are consistent with the presence of two serial dependencies that are either in opposition or in agreement (this is also true of the hierarchical Bayesian model, as seen in Appendix Fig. B2, which breaks down the power analysis according to whether the direction of the two dependencies is the same or opposite, revealing equivalent statistical power in both cases).

We use the quadrant plot as a visual aid, rather than a statistical test, owing to a sampling bias that makes it unsuitable for inferential tests. This sampling bias results in the average magnitude of difference from the prior stimulus/response being larger in the accurate quadrants compared with the inaccurate quadrants. For instance, the average extent to which the prior stimulus is clockwise from the current stimulus (i.e., the right column) is likely to be greater for the upper-right accurate quadrant than for the lower-right inaccurate quadrant. This is because whenever the prior stimulus is very clockwise from the current stimulus (a point falling to the far right of the x -axis), the response to the prior stimulus is very likely to also be clockwise from the current stimulus (thus falling into the upper-right quadrant), whether the participant made a small or a large error: only extremely

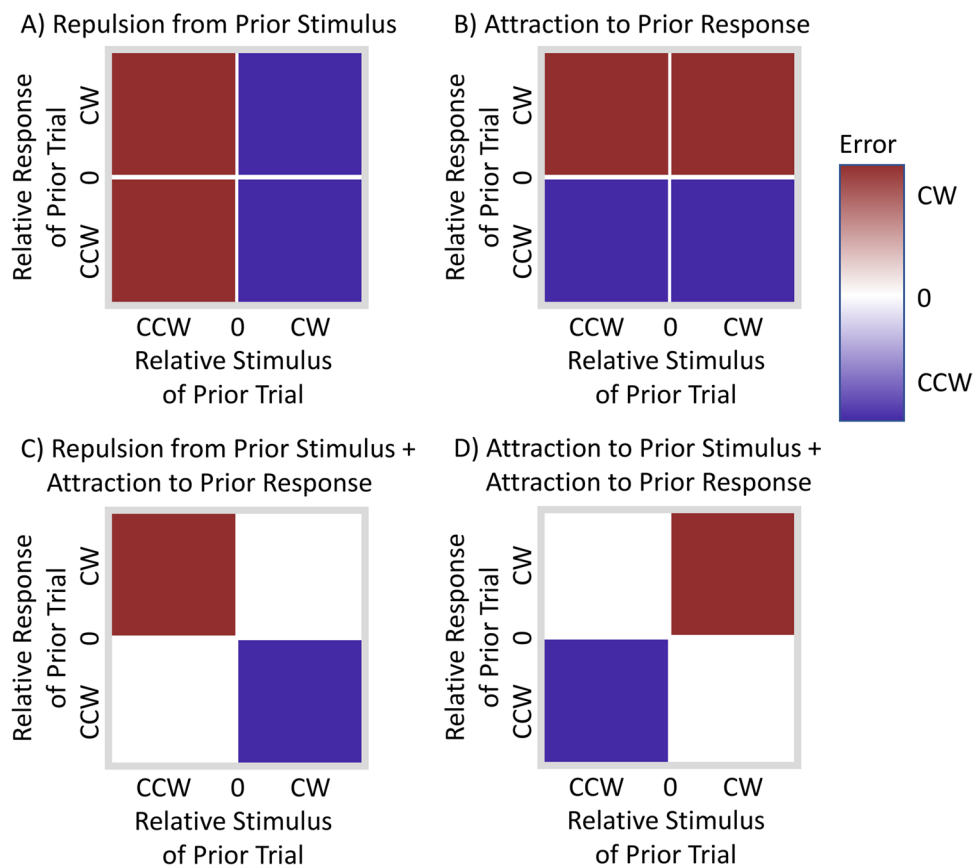


Fig. 3 A guide to interpreting quadrant plots, with examples of a dependence only on the prior stimulus **A**), a dependence only on the prior response **B**), or two dependencies that are in opposite directions **C**) or the same direction **D**). (Color figure online)

large errors on the prior trial would cause the prior response to cross into the inaccurate (lower-right) quadrant. In contrast, whenever the prior stimulus is only slightly clockwise from the current stimulus (falling just to the right of zero on the x -axis), the response to the prior stimulus is much more likely to fall counterclockwise from the current stimulus (thus falling into the lower-right quadrant): even a small error on the prior trial might cause the response to cross into the bottom-right quadrant. Thus, on average, data falling in the upper-right quadrant will be biased towards the right side of the quadrant and data falling in the lower-right quadrant will be biased towards the left (more central) side of the quadrant.). The critical point is that, if there is only one dependency, the main effect outcomes seen in Fig. 3A and B are unlikely, because of this sampling bias. To reach statistical conclusions about whether two dependencies exist versus one, or whether the dataset contains sufficient constraint to answer this question, we recommend the use of model comparison, followed by the use of quadrant plots as a check to see if the data generally support the conclusions reached with model comparison.

Results

Bayesian analysis

Bayesian models were applied to the trial-level data. Three models were fit that allowed for either (1) an effect of only the prior orientation, (2) an effect of only the prior response, or (3) an effect of both the prior orientation and the prior response. Models were compared by their PSIS-LOO+ score (Vehtari et al., 2017, 2015), which estimates the model's ability to accurately predict held-out data. This should adjust for the extra model flexibility inherent in the model with both dependencies. In Experiment 1 (Fig. 4), model comparison favored the model with both effects as compared with either of the single effect models (difference \pm standard error: full model vs. stimulus only: -103.72 ± 16.04 ; full model vs. response only: -149.99 ± 19.61). Thus, Experiment 1 provided enough constraint to estimate two separate dependencies.

With five times more data than is typically collected in serial dependence studies (see Table 1), Experiment 1 had higher statistical power for dissociating the effect of the prior response from the effect of the prior stimulus (see also Supplementary Fig. C3, which shows that the magnitude of errors in Experiment 1 was comparable to the other datasets, suggesting that model comparison was reliable for Experiment 1 primarily because of trial count). We next applied model comparison to four previously published datasets (Fig. 4). Across the datasets, only one other—Pascucci et al., Experiment 1—provided enough constraint to reliably separate stimulus and response effects. That is, although

one-factor analyses of the remaining three experiments suggested dependencies (Appendix A, Appendix C), the model comparison did not allow a definitive conclusion as to whether there were dependencies on the previous stimulus, previous response, or both (i.e., at least two models tied for being the most predictive, Fig. 4).

To assess why the models could not be reliably distinguished in some cases, we performed a best-case scenario simulation study that assumed independence between prior responses and prior stimuli and eliminated individual differences and bias effects other than prior responses and prior stimuli. The models were fit to artificial datasets generated in this best-case scenario manner, with different datasets generated assuming either one or two dependencies (Appendix B). These simulations had two goals: first, to confirm that a comparison of Bayesian models could in principle determine whether a dataset contained one versus two dependencies, and, second, to estimate how many trials would be required for that determination to be reliable in this best-case situation. The simulations revealed that approximately 10,000 trials are required to reliably distinguish between two effects (this trial count is not broken down by subject because this situation assumed no individual differences). At 5,000 trials, two components may be identified in only 75% of experiments. Considering that none of the 4 previously published datasets contained more than 10,000 total trials (Table 1), it is unsurprising that model comparison failed to reach clear conclusions for three of them. In light of these results, we report here additional analyses of the two experiments that were able to separate stimulus and response dependencies, namely the present experiment, which contained more than 10,000 trials, and Experiment 1 of Pascucci et al. (for completeness, analyses for all experiments are presented in Appendices A and C).

Before interpreting the results, we assessed whether the winning model adequately captured the data by comparing observed data to predicted data in terms of errors smoothed with a median window and plotted separately as a function of either the prior orientation or prior response (Fig. 5, left column for observed data and middle column for data generated from the model with both dependencies). Note that the model predictions exhibited “peripheral bumps” (Fritsche et al., 2017), whereby the dependence on the prior response appears to swap from attractive to repulsive at approximately $\pm 45^\circ$. This occurred even though each derivative of a circular normal distribution function is either exclusively repulsive or exclusively attractive. In the simulated datasets, the peripheral bumps reflect the sum of two overlapping dependencies; the same effect is illustrated in the middle row of Fig. 2A.

Critically, the directions of the two dependencies in the models were unconstrained: both effects could be attractive, both could be repulsive, or one could be attractive and the other repulsive. Nevertheless, for both experiments that

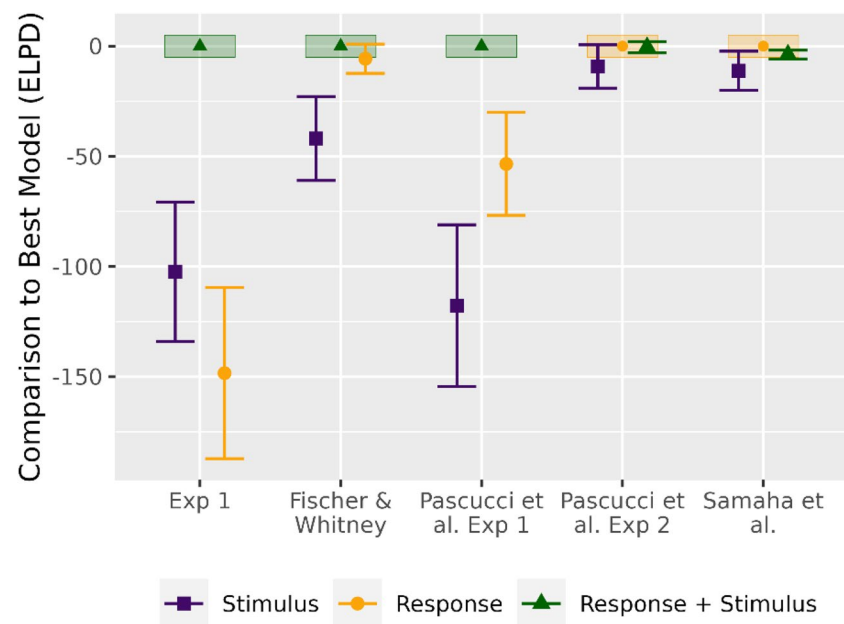


Fig. 4 Model comparison (predictions for held out data) between a model with only the prior stimulus, only the prior response, or both dependencies. In all experiments, the model with both was either the most predictive model or tied for most predictive. The y-axis shows the relative expected log predictive density, estimated with PSIS-

LOO+ for each model, as compared with the most predictive model (i.e., the most predictive model is always 0, with error shown by the shaded rectangle). Error bars extend two standard errors of the mean of the difference. (Color figure online)

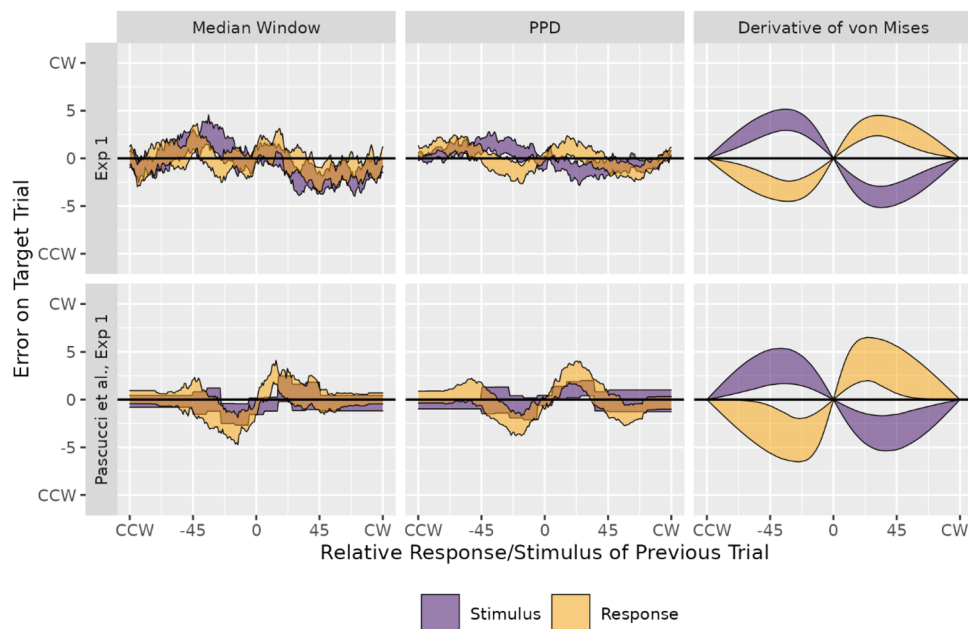


Fig. 5 Model results from the model with both dependencies. Model predictions (middle) can be compared with observed data (left), and each estimated dependency is shown in isolation (right), revealing that errors are attracted to the prior response and repelled from the prior stimulus. The two experiments for which model comparison could distinguish stimulus and response effects are in the different rows. Left: Preprocessed data were smoothed with a moving median window. Ribbons span the 95% confidence intervals, across participants. Middle: To check whether the model captured the data, the model with both dependencies was used to generate artificial

datasets using parameters estimated from the real data. The generated datasets were preprocessed as in the left column. Ribbons span the 95% highest density interval across the preprocessed data. Right: The 95% highest density interval for the posteriors of the estimated dependencies. Unlike the median smoothing (left/middle column) that summates the dependencies, these plots show each dependency in isolation after removing the contribution of the other dependency and after removing any preferences for oblique or cardinal angles and any bias to rotate responses. PPD = Posterior Predictive Distribution. (Color figure online)

provided sufficient constraint for disentangling stimulus and response effects, the results indicated an attraction to the prior response and repulsion from the prior stimulus (Fig. 5, right column). Furthermore, as seen in the third column of Appendix Fig. C4, the results from the other three datasets were consistent with the claim that repulsion from the prior stimulus and attraction to the prior response is a general property of serial dependence experiments. More specifically, when applying the model that contained both dependencies, none of the datasets produced results that were reliably attractive to the prior stimulus (i.e., the 95% highest density interval for the prior stimulus effect was either indeterminate between attraction versus repulsion or decidedly favoring repulsion) and all datasets produced results that were reliably attractive to the prior response (i.e., the 95% highest density interval for the prior response effect was decidedly favoring attraction).

Quadrant plots

The Bayesian analyses suggest that the prior stimulus and prior response exerted opposite influences on errors, but this conclusion appears to contradict the moving window analysis, which does not show a clear repulsion from the prior stimulus nor an unequivocal attraction to the prior response (Fig. 5, left column). As discussed in relation to Fig. 2, this may reflect the limitations of a one-factor analysis such as the median window (Fig. 2A), whereby two opposing dependencies can counteract each other. We therefore sought to visualize the effects suggested by the Bayesian model with an alternative, non-Bayesian analysis that we refer to as a “quadrant plot”. Quadrant plots group together trials according to whether the prior trial was clockwise or counterclockwise relative to the current trial, both in terms of the prior stimulus and prior response (e.g., Fig. 2C and 3). The winning Bayesian model predicted, for both datasets, that the response and stimulus dependencies should counteract each other, specifically in the two diagonal quadrants (bottom left and top right), resulting in nearly unbiased responses. The model also predicted that the two dependencies should work in the same direction for the off-diagonal quadrants, resulting in large errors in the upper-left and lower-right quadrants, but with these errors of opposite direction in the two quadrants. This pattern of results is clearly seen for both experiments that offered sufficient constraint to separate the effects (Fig. 6; quadrant plots for the remaining experiments are provided in Appendix Fig. C2).

Discussion

Recent stimuli and recent responses influence perceptual decisions. For instance, in visual aftereffects, such as the tilt aftereffect, perception is repelled away from recently viewed

stimuli. In contrast to visual aftereffects, the serial dependence effect appears to show that perception is attracted towards recently viewed stimuli. However, this appearance may be misleading, considering that stimuli elicit a cascade of processes that range from detecting visual primitives to making perceptual decisions, with effects at one level potentially counteracting effects at another level. We claim that analyses of serial dependence should assume dependencies on both the prior response and prior stimulus. To address this mixture of processes, we developed new analysis techniques, ran a new experiment that collected enough data to reliably apply these techniques (five times more than is typically collected), and reanalyzed four previously published serial dependence experiments. The results lead to two separate conclusions.

First, we draw the methodological conclusion that typical experimental procedures in the study of serial dependence can in theory tease apart these two sequential dependencies, although in practice the amount of data required is substantially larger than is typically collected. When using a one-factor analysis (i.e., an analysis that considers the prior stimulus and collapses over the prior response, or vice versa), perfectly opposing dependencies and relatively low response accuracy can lead to mistaken conclusions (e.g., the last row of Fig. 2A suggests the absence of any effect). However, a two-factor analysis (i.e., an analysis that considers combinations of prior stimulus and prior response) can identify both dependencies when there are a sufficient variety of combinations of the two variables (i.e., ample data on the “off-diagonal” in Fig. 2B; see also Figs. 3C and 6). For our analyses, we estimated that, at minimum, experiments should contain 10,000 total trials (Appendix B), substantially more than were collected in many serial dependence experiments (e.g., Table 1). This estimate represents a “best-case” scenario using simulated data, but the necessary total trial count, and the breakdown of that total into a certain number of subjects and trial count per subject, will depend on experiment-specific factors like the accuracy of responses, variability across participants, the magnitude of the individual effects, and the specific analysis method.

Our methodological conclusion regarding use of a two-factor analysis is related to, but different than use of a two-factor model (Pascucci et al., 2019b). Pascucci et al. compared the performance of a one-factor model (attraction to the prior stimulus) versus a two-factor model that contained both a repulsion from the prior stimulus and an attraction to the prior response, with these factors summing up to produce a single response function. They applied both models to data, finding that the two-factor model provided a better explanation. For instance, the two-factor model explained an experiment in which repulsion from the prior stimulus was found when the prior stimulus was unreported, but attraction to the prior stimulus was found when the prior stimulus

was reported. However, this was a one-factor analysis that considered only the orientation of the prior stimulus rather than considering trial-by-trial combinations of prior stimulus and prior response. As shown in Fig. 2, a one-factor analysis can be misleading, and this is still true even if a two-factor model is fit to the one-factor analysis. For instance, the current experiment produced a completely null result when using a one-factor analysis of prior stimulus. Application of the two-factor model to data collapsed over prior response would lead to a parameter identifiability problem because the results would be equally explained by setting both factors to zero, or by having two factors that perfectly counteracted each other. In contrast, as shown in the model results in Fig. 5 and the quadrant analysis in Fig. 6, when using a two-factor analysis, the apparently null result is revealed as two highly reliable factors that counteracted each other. In support of Pascucci et al.'s two-factor model, when we used a two-factor analysis, the two-factor model did significantly better than either one-factor model at explaining the data for two of the five datasets, and for the other three datasets, the two-factor model was either the best model or not reliably different than the best model (the latter cases occurred when the one-factor model with only prior response was the best model). In no case was there evidence against the two-factor model. Thus, we claim that typical serial dependence experiments contain both a dependence on the prior stimulus and a dependence on the prior response (see also Moon & Kwon, 2022). More importantly, we find that conclusions about the direction of any dependencies should be based on trial-by-trial two-factor analyses.

Second, we conclude that when applied to data from the present experiment and from four previously published studies, our analysis failed to support the existence of an

attractive effect towards the distal properties of the prior stimulus (as opposed to the percept of the stimulus, which may better align with the prior response). That is, by using analysis techniques that tease apart response and stimulus effects, we failed to find an attraction to the prior stimulus in any dataset. For two of the five experiments, stimulus and response effects were teased apart successfully, and for these two experiments, we found that the response effect was an attraction to the prior response whereas the stimulus effect was repulsion from the prior stimulus. In the remaining three experiments, no conclusion could be drawn about an effect of the previous stimulus. For all datasets, when applying a model with both dependencies, the response effect appeared to be attractive while the stimulus effect was either repulsive or indeterminate.

The existence of two dependencies may help explain conflicting results in the literature. As noted in the introduction, several studies have observed positive dependencies on the prior stimulus when participants withhold responses (Fischer & Whitney, 2014; Fornaciai & Park, 2018a; Pascucci et al., 2019b). Other studies find repulsive effects even when responses are given (Bae & Luck, 2020). If there is only one kind of dependency, it is not clear from these studies why the dependency should be so changeable, including reversals in the direction of the dependency. However, all of these results used one-factor analysis techniques, which analyzed the data in terms of the prior stimulus or prior response while collapsing over the other factor. In contrast, our results suggest that most experiments encourage two (or more) dependencies, and that for a given experiment the balance may be tilted in favor of either attraction or repulsion within a one-factor analysis, or even balanced to give the appearance of no dependence. If this conclusion is correct, care must

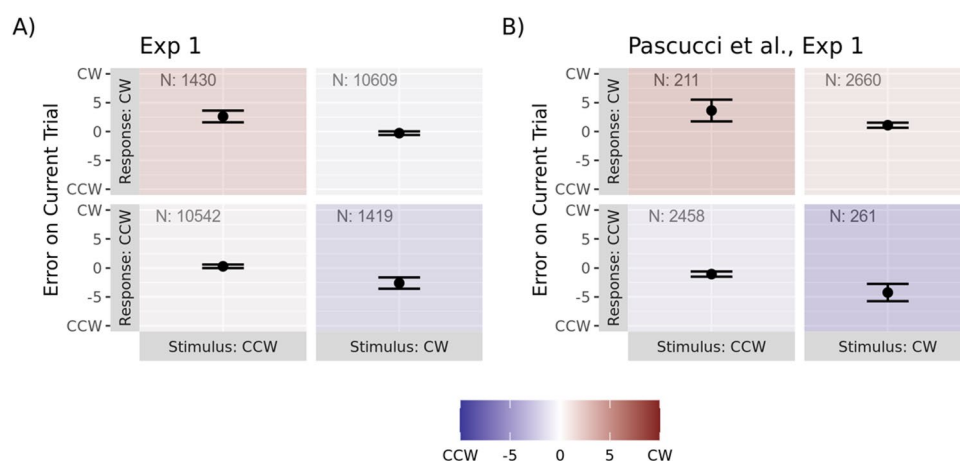


Fig. 6 Errors differ across combinations of clockwise and counterclockwise prior stimulus and prior response in **A**) Experiment 1 and **B**) Pascucci et al. Experiment 1. Datapoints and colors in each quadrant correspond to the average error for the associated combination

of clockwise/counterclockwise prior stimulus and prior response (compare with Fig. 2C). Inset numbers in each quadrant indicate how many trials fell into that quadrant. Error bars cover the 95% confidence intervals. (Color figure online)

be taken when interpreting changes in the magnitude of a dependence analyzed with a one-factor analysis. For example, a manipulation that appears to reduce or even reverse an attraction effect as inferred with a one-factor analysis may in fact reflect a relatively subtle change in the balance of power between a robust attraction to the prior response that counteracts partially or wholly a robust repulsion from the prior stimulus. Unless a two-factor analysis is used, it is not clear whether the reduction in attraction effect reflects a reduction in the attractive dependency or an increase in the repulsive dependency.

There are two important caveats to our conclusions. First, the statistical modeling technique assumed that the magnitude of stimulus and response dependencies are constant across trials (the parameters of the circular normal distribution used to capture each kind of dependency were allowed to vary between participants but were assumed to be fixed for each participant across trials). This is likely a simplification. For instance, perception of the prior stimulus is likely to vary across trials due to factors such as fluctuations in vigilance; in the extreme case, a participant may fail to look at the prior stimulus entirely, precluding any opportunity for dependence on that stimulus. By failing to acknowledge these fluctuations, the model likely underestimates the magnitude of the stimulus dependence. Such underestimation could have contributed to the ambiguous effect of the previous stimulus in the three datasets for which model comparison resulted in a tie. However, when modeling data that allowed identification of two dependencies, the stimulus effect was repulsive rather than attractive, meaning that the only positive evidence we have for any effect of the previous stimulus is for a repulsive effect, in line with the classic tilt aftereffect. Although this limitation may have led to an underestimation of effect magnitude, it is unlikely that it resulted in mistakenly identifying the direction of the stimulus dependence as negative when it was truly positive.

The second caveat is that our results, although supportive of an interpretation of the serial dependence effect as a mixture of two dependencies (Bae & Luck, 2020; Fornaciai & Park, 2019, 2020; Fritsche et al., 2017; Pascucci et al., 2019b), do not reveal the psychological processes that caused the two dependencies. Nor do they reveal whether the psychological mechanisms are distinct or shared, nor how they relate to the wider literature on sequential effects (Kiyonaga et al., 2017). When we were able to statistically isolate both dependencies, they resembled effects well-documented before the proposal of a continuity field: A repulsive visual aftereffect and an attractive response effect. We suspect that the repulsion from the prior orientation is a manifestation of the tilt-aftereffect, given that the effect is ubiquitous and automatic, it can occur after brief exposure to the inducer (e.g., under 10 ms; Sekuler & Littlejohn, 1974), does not require awareness of the inducer (Kanai et al., 2006), is

sensitive to a range of low-level features (Greenlee & Magnussen, 1987; Harris & Calvert, 1985, 1989; Morant & Mikaelian, 1960; Parker, 1972), can occur across different screen locations and spatial frequencies (Jacob et al., 2021; Morant & Mikaelian, 1960; Parker, 1972), and could arise from mechanisms that have been observed in single-cell recordings of early visual neurons (Clifford et al., 2000; Dragoi et al., 2000; Gutnisky & Dragoi, 2008; Patterson et al., 2013; Wissig & Kohn, 2012). These prior results suggest that a repulsive tilt aftereffect is caused by merely viewing the inducing stimulus, and so there is every reason to expect that such perceptual aftereffects play some role in a typical serial dependence effect study.

We further suspect that the attraction to the prior response is mediated by a process that lies “further along” the pathway than the early-stage processes implicated in the repulsive tilt aftereffect (e.g., the processes involved in making a perceptual decision). It remains possible that the process is perceptual, as implied by the continuity field; e.g., if participants “saw” an orientation that was not presented and accurately reported that illusory orientation, then subsequent attraction toward the (inaccurate) perception of the inducer would manifest as an attraction to the prior response (see also Cicchini et al., 2017; Fischer & Whitney, 2014; St John-Saaltink et al., 2016). This account of the attraction effect agrees with the characterization of serial dependence as a manifestation of perception enacting Bayesian inference (Cicchini et al., 2018; Kalm & Norris, 2018; van Bergen & Jehee, 2019). However, it is also possible that participants perceive the prior orientation accurately but respond inaccurately, and in that case, the attraction could reflect their prior decision about the inducer (Akaishi et al., 2014; Braun et al., 2018). Indeed, multiple high-level processes have been proposed as accounts of response hysteresis, including shifting beliefs about the base rate of stimuli (Zhang et al., 2014), slow changes in how much evidence participants require before initiating responses (Wagenmakers et al., 2004), or the integration of low-level sensory information into a decision (Pascucci et al., 2019b). Finally, it is possible that there are three (or more) effects, with an attraction to the prior response coexisting with both repulsion and attraction to the prior stimulus—our analyses cannot rule out this possibility.

Our development of techniques for untangling two dependencies based on prior responses and prior stimuli indicates that most studies are statistically underpowered to address this problem. Our power analysis indicated that more data are needed to separate these effects, which is why we ran a study with five times as many trials than are typically collected. As applied to our more powerful dataset, these techniques revealed the clear answer that both dependencies exist and work in opposite directions. As applied to previously collected datasets, which are underpowered for this question, the results were somewhat equivocal, but were

consistent with attraction to the prior response and repulsion from the prior stimulus. Our recommendation for the serial dependence literature is to run more highly powered studies that can tease apart the two dependencies based on natural variation across trials. Furthermore, in light of mounting evidence that both dependencies exist and evidence that the two dependencies can counteract each other, studies that seek to manipulate serial dependence should take a multi-factor manipulation approach rather than manipulating only one experimental variable. Single factor manipulations may affect the two dependencies in opposite ways, the same way, or may affect just one of the two dependencies, with these various possibilities complicating interpretation of the results. By crossing a manipulation aimed at affecting one of the dependencies with a different manipulation aimed at affecting the other dependency, it may be possible to directly identify the two dependencies through manipulations. But in the absence of such multifactor manipulation studies, our analysis techniques could be used to measure the dependencies, provided that the study collects enough data.

In sum, we suggest that studies of the serial dependence effect ought to be analyzed with two-factor analyses that consider combinations of prior stimulus and prior response. We developed two new techniques for doing so, with the first assessing the reliability of separate dependencies using model comparison, and the second producing a way of visualizing the separate dependencies in a quadrant plot. As applied to several datasets, these techniques failed to find evidence of an attraction toward the prior stimulus. Furthermore, when they were able to identify separate dependencies, they revealed a repulsive stimulus effect and an attractive response effect. Now that we have successfully applied this two factor-modeling approach to the study of serial dependence effects in orientation tasks, a similar approach could be taken in addressing other kinds of serial dependence paradigms that involve trial-by-trial presentations of stimuli that must be perceived and identified with continuous judgments (e.g., line length, approximate number judgments, auditory pitch, duration judgments, etc.).

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.3758/s13423-023-02320-3>.

Acknowledgments We thank Dr. Fischer for providing us with their data, as well as Dr. Samaha and Dr. Pascucci for uploading their data publicly. This work was funded by the National Institute of Mental Health grant 1RF1MH114277 awarded to R.A.C. and D.E.H.

Funding This work was funded by the National Institute of Mental Health grant 1RF1MH114277 awarded to R.A.C. and D.E.H.

Data Availability The datasets generated during and analyzed during the current study are available in the Open Science Framework repository, <https://osf.io/w4kcj/>.

Declarations

Conflicts of interest The authors have no relevant financial or non-financial interests to disclose.

Ethics approval This research was done in accordance with the 1964 Declaration of Helsinki and was approved by the University of Massachusetts Amherst Institutional Review Board.

Consent to participate Informed consent was obtained from all individual participants included in the study.

Consent to publish The authors affirm that human research participants provided informed consent for the publication of the results of Experiment 1.

References

- Akaishi, R., Umeda, K., Nagase, A., & Sakai, K. (2014). Autonomous mechanism of internal choice estimate underlies decision inertia. *Neuron*, 81(1), 195–206. <https://doi.org/10.1016/j.neuron.2013.10.018>
- Alais, D., Kong, G., Palmer, C., & Clifford, C. (2018). Eye gaze direction shows a positive serial dependency. *Journal of Vision*, 18(4), 11. <https://doi.org/10.1167/18.4.11>
- Annis, J., Dubé, C., & Malmberg, K. J. (2018). A Bayesian approach to discriminating between biased responding and sequential dependencies in binary choice data. *Decision*, 5(1), 16–41. <https://doi.org/10.1037/dec0000060>
- Bae, G.-Y., & Luck, S. J. (2020). Serial dependence in vision: Merely encoding the previous-trial target is not enough. *Psychonomic Bulletin & Review*, 27(2), 293–300. <https://doi.org/10.3758/s13423-019-01678-7>
- Betancourt, M. (2017). A conceptual introduction to Hamiltonian Monte Carlo. *ArXiv*. [arXiv:1701.02434v2](https://arxiv.org/abs/1701.02434v2)
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, 10(4), 433–436. <https://doi.org/10.1163/156856897X00357>
- Braun, A., Urai, A. E., & Donner, T. H. (2018). Adaptive history biases result from confidence-weighted accumulation of past choices. *The Journal of Neuroscience*, 38(10), 2418–2429. <https://doi.org/10.1523/JNEUROSCI.2189-17.2017>
- Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M. A., Guo, J., Li, P., & Riddell, A. (2016). Stan: A probabilistic programming language. *Journal of Statistical Software*, 76(1), 1–43. <https://doi.org/10.18637/jss.v076.i01>
- Cicchini, G. M., Anobile, G., & Burr, D. C. (2014). Compressive mapping of number to space reflects dynamic encoding mechanisms, not static logarithmic transform. *Proceedings of the National Academy of Sciences*, 111(21), 7867–7872. <https://doi.org/10.1073/pnas.1402785111>
- Cicchini, G. M., Mikellidou, K., & Burr, D. (2017). Serial dependencies act directly on perception. *Journal of Vision*, 17(14), 6. <https://doi.org/10.1167/17.14.6>
- Cicchini, G. M., Mikellidou, K., & Burr, D. C. (2018). The functional role of serial dependence. *Proceedings of the Royal Society B: Biological Sciences*, 285(1890), 20181722. <https://doi.org/10.1098/rspb.2018.1722>
- Clifford, C. W. G., Wenderoth, P., & Spehar, B. (2000). A functional angle on some after-effects in cortical vision. *Proceedings of the Royal Society of London Series B: Biological Sciences*, 267(1454), 1705–1710. <https://doi.org/10.1098/rspb.2000.1198>

- Dragoi, V., Sharma, J., & Sur, M. (2000). Adaptation-induced plasticity of orientation tuning in adult visual cortex. *Neuron*, 28(1), 287–298. [https://doi.org/10.1016/S0896-6273\(00\)00103-3](https://doi.org/10.1016/S0896-6273(00)00103-3)
- Duane, S., Kennedy, A. D., Pendleton, B. J., & Roweth, D. (1987). Hybrid Monte Carlo. *Physics Letters B*, 195(2), 216–222. [https://doi.org/10.1016/0370-2693\(87\)91197-X](https://doi.org/10.1016/0370-2693(87)91197-X)
- Fischer, J., & Whitney, D. (2014). Serial dependence in visual perception. *Nature Neuroscience*, 17(5), 738–743. <https://doi.org/10.1038/nn.3689>
- Fischer, C., Czoschke, S., Peters, B., Rahm, B., Kaiser, J., & Bleidowski, C. (2020). Context information supports serial dependence of multiple visual objects across memory episodes. *Nature Communications*, 11(1), 1932. <https://doi.org/10.1038/s41467-020-15874-w>
- Fornaciai, M., & Park, J. (2018). Attractive serial dependence in the absence of an explicit task. *Psychological Science*, 29(3), 437–446. <https://doi.org/10.1177/0956797617737385>
- Fornaciai, M., & Park, J. (2018). Serial dependence in numerosity perception. *Journal of Vision*, 18(9), 15. <https://doi.org/10.1167/18.9.15>
- Fornaciai, M., & Park, J. (2019). Spontaneous repulsive adaptation in the absence of attractive serial dependence. *Journal of Vision*, 19(5), 21. <https://doi.org/10.1167/19.5.21>
- Fornaciai, M., & Park, J. (2020). Attractive serial dependence between memorized stimuli. *Cognition*, 200, 104250. <https://doi.org/10.1016/j.cognition.2020.104250>
- Fritsche, M. (2016). *To smooth or not to smooth: Investigating the role of serial dependence in stabilising visual perception*. Radboud University.
- Fritsche, M., Mostert, P., & de Lange, F. P. (2017). Opposite effects of recent history on perception and decision. *Current Biology*, 27(4), 590–595. <https://doi.org/10.1016/j.cub.2017.01.006>
- Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, 7(4), 457–472. <https://doi.org/10.1214/ss/1177011136>
- Gelman, A., Carlin, J., Stern, H., Dunson, D., Vehtari, A., & Rubin, D. (2013). *Bayesian data analysis* (3rd ed.). Chapman and Hall/CRC.
- Gibson, J. J., & Radner, M. (1937). Adaptation, after-effect and contrast in the perception of tilted lines. I. quantitative studies. *Journal of Experimental Psychology*, 20(5), 453–467. <https://doi.org/10.1037/h0059826>
- Greenlee, M. W., & Magnussen, S. (1987). Saturation of the tilt aftereffect. *Vision Research*, 27(6), 1041–1043. [https://doi.org/10.1016/0042-6989\(87\)90017-4](https://doi.org/10.1016/0042-6989(87)90017-4)
- Gutnisky, D. A., & Dragoi, V. (2008). Adaptive coding of visual information in neural populations. *Nature*, 452(7184), 220–224. <https://doi.org/10.1038/nature06563>
- Harris, J. P., & Calvert, J. E. (1985). The tilt after-effect: Changes with stimulus size and eccentricity. *Spatial Vision*, 1(2), 113–129. <https://doi.org/10.1163/156856885X00134>
- Harris, J. P., & Calvert, J. E. (1989). Contrast, spatial frequency and test duration effects on the tilt aftereffect: Implications for underlying mechanisms. *Vision Research*, 29(1), 129–135. [https://doi.org/10.1016/0042-6989\(89\)90179-X](https://doi.org/10.1016/0042-6989(89)90179-X)
- Hoffman, M., & Gelman, A. (2014). The no-U-turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research*, 15, 1593–1623.
- Jacob, L. P. L., Potter, K. W., & Huber, D. E. (2021). A neural habituation account of the negative compatibility effect. *Journal of Experimental Psychology: General*, 150(12), 2567. <https://doi.org/10.1037/xge0001055>
- Jastrow, J. (1892). Studies from the University of Wisconsin: On the judgment of angles and positions of lines. *The American Journal of Psychology*, 5(2), 214. <https://doi.org/10.2307/1410867>
- Jepma, M., Jones, M., & Wager, T. D. (2014). The dynamics of pain: Evidence for simultaneous site-specific habituation and site-nonspecific sensitization in thermal pain. *The Journal of Pain*, 15(7), 734–746. <https://doi.org/10.1016/j.jpain.2014.02.010>
- Jesteadt, W., Luce, R. D., & Green, D. M. (1977). Sequential effects in judgments of loudness. *Journal of Experimental Psychology: Human Perception and Performance*, 3(1), 92–104. <https://doi.org/10.1037/0096-1523.3.1.92>
- Jones, M., Curran, T., Mozer, M. C., & Wilder, M. H. (2013). Sequential effects in response time reveal learning mechanisms and event representations. *Psychological Review*, 120(3), 628–666. <https://doi.org/10.1037/a0033180>
- Kalm, K., & Norris, D. (2018). Visual recency bias is explained by a mixture model of internal representations. *Journal of Vision*, 18(7), 1. <https://doi.org/10.1167/18.7.1>
- Kanai, R., Tsuchiya, N., & Verstraten, F. A. J. (2006). The scope and limits of top-down attention in unconscious visual processing. *Current Biology*, 16(23), 2332–2336. <https://doi.org/10.1016/j.cub.2006.10.001>
- Kiyonaga, A., Scimeca, J. M., Bliss, D. P., & Whitney, D. (2017). Serial dependence across perception, attention, and memory. *Trends in Cognitive Sciences*, 21(7), 493–497. <https://doi.org/10.1016/j.tics.2017.04.011>
- Kruschke, J. K. (2015). *Doing Bayesian data analysis* (2nd ed.). Elsevier. <https://doi.org/10.1016/B978-0-12-405888-0.00001-5>
- Liberman, A., Fischer, J., & Whitney, D. (2014). Serial dependence in the perception of faces. *Current Biology*, 24(21), 2569–2574. <https://doi.org/10.1016/j.cub.2014.09.025>
- Moon, J., & Kwon, O.-S. (2022). *Dissecting the effects of adaptive encoding and predictive inference on a single perceptual estimation* [Preprint]. Neuroscience. <https://doi.org/10.1101/2022.02.24.481765>
- Morant, R. B., & Mikaelian, H. H. (1960). Inter-field tilt after-effects. *Perceptual and Motor Skills*, 10(2), 95–98.
- Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorials in Quantitative Methods for Psychology*, 4(2), 61–64. <https://doi.org/10.3758/s13414-012-0291-2>
- Parker, D. M. (1972). Contrast and size variables and the tilt after-effect. *Quarterly Journal of Experimental Psychology*, 24(1), 1–7. <https://doi.org/10.1080/14640747208400260>
- Pascucci, D., & Plomp, G. (2021). Serial dependence and representational momentum in single-trial perceptual decisions. *Scientific Reports*, 11(1), 9910. <https://doi.org/10.1038/s41598-021-89432-9>
- Pascucci, D., Mancuso, G., Santandrea, E., Della Libera, C., Plomp, G., & Chelazzi, L. (2019a). *Laws of concatenated perception: Vision goes for novelty, decisions for perseverance* [Data set]. Zenodo. <https://doi.org/10.5281/ZENODO.2544946>
- Pascucci, D., Mancuso, G., Santandrea, E., Della Libera, C., Plomp, G., & Chelazzi, L. (2019b). Laws of concatenated perception: Vision goes for novelty, decisions for perseverance. *PLOS Biology*, 17(3), e3000144. <https://doi.org/10.1371/journal.pbio.3000144>
- Patterson, C. A., Wissig, S. C., & Kohn, A. (2013). Distinct effects of brief and prolonged adaptation on orientation tuning in primary visual cortex. *Journal of Neuroscience*, 33(2), 532–543. <https://doi.org/10.1523/JNEUROSCI.3345-12.2013>
- 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>
- Samaha, J., Switzky, M., & Postle, B. R. (2019). Confidence boosts serial dependence in orientation estimation. *Journal of Vision*, 19(4), 25. <https://doi.org/10.1167/19.4.25>
- Samaha, J., Barrett, J. J., Sheldon, A. D., LaRocque, J. J., & Postle, B. R. (2016). Dissociating perceptual confidence from discrimination accuracy reveals no influence of metacognitive awareness on working memory. *Frontiers in Psychology*, 7, 851. <https://doi.org/10.3389/fpsyg.2016.00851>

- Schwiedrzik, C. M., Ruff, C. C., Lazar, A., Leitner, F. C., Singer, W., & Melloni, L. (2014). Untangling perceptual memory: Hysteresis and adaptation map into separate cortical networks. *Cerebral Cortex*, 24(5), 1152–1164. <https://doi.org/10.1093/cercor/bhs396>
- Sekuler, R., & Littlejohn, J. (1974). Tilt aftereffect following very brief exposures. *Vision Research*, 14(1), 151–152. [https://doi.org/10.1016/0042-6989\(74\)90133-3](https://doi.org/10.1016/0042-6989(74)90133-3)
- St John-Saaltink, E., Kok, P., Lau, H. C., & de Lange, F. P. (2016). Serial dependence in perceptual decisions is reflected in activity patterns in primary visual cortex. *The Journal of Neuroscience*, 36(23), 6186–6192. <https://doi.org/10.1523/JNEUROSCI.4390-15.2016>
- Suárez-Pinilla, M., Seth, A. K., & Roseboom, W. (2018). Serial dependence in the perception of visual variance. *Journal of Vision*, 18(7), 4. <https://doi.org/10.1167/18.7.4>
- Taubert, J., Alais, D., & Burr, D. (2016). Different coding strategies for the perception of stable and changeable facial attributes. *Scientific Reports*, 6(1), 32239. <https://doi.org/10.1038/srep32239>
- Taubert, J., Van der Burg, E., & Alais, D. (2016). Love at second sight: Sequential dependence of facial attractiveness in an on-line dating paradigm. *Scientific Reports*, 6(1), 22740. <https://doi.org/10.1038/srep22740>
- The MathWorks Inc. (2015). MATLAB version: 8.5.0 (R2015a), Natick, Massachusetts: The MathWorks Inc. <https://www.mathworks.com>
- Urai, A. E., de Gee, J. W., Tsetsos, K., & Donner, T. H. (2019). Choice history biases subsequent evidence accumulation. *eLife*, 8, e46331. <https://doi.org/10.7554/eLife.46331>
- van Bergen, R. S., & Jehee, J. F. M. (2019). Probabilistic representation in human visual cortex reflects uncertainty in serial decisions. *The Journal of Neuroscience*, 39(41), 8164–8176. <https://doi.org/10.1523/JNEUROSCI.3212-18.2019>
- Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, 27(5), 1413–1432. <https://doi.org/10.1007/s11222-016-9696-4>
- Vehtari, A., Simpson, D., Gelman, A., Yao, Y., & Gabry, J. (2015). Pareto smoothed importance sampling. *arXiv*. [arXiv:1507.02646](https://arxiv.org/abs/1507.02646)
- Vehtari, A., Gabry, J., Magnusson, M., Yao, Y., Bürkner, P., Paananen, T., Gelman, A. (2020). loo: Efficient leave-one-out cross-validation and WAIC for Bayesian models. *R package version 2.4.1*. <https://mc-stan.org/loo/>
- Wagenmakers, E.-J., Farrell, S., & Ratcliff, R. (2004). Estimation and interpretation of 1/fⁿ noise in human cognition. *Psychonomic Bulletin & Review*, 11(4), 579–615. <https://doi.org/10.3758/BF03196615>
- Webster, M. A. (2015). Visual adaptation. *Annual Review of Vision Science*, 1(1), 547–567. <https://doi.org/10.1146/annurev-vision-082114-035509>
- Wei, X.-X., & Stocker, A. A. (2015). A Bayesian observer model constrained by efficient coding can explain “anti-Bayesian” percepts. *Nature Neuroscience*, 18(10), 1509–1517. <https://doi.org/10.1038/nn.4105>
- Wissig, S. C., & Kohn, A. (2012). The influence of surround suppression on adaptation effects in primary visual cortex. *Journal of Neurophysiology*, 107(12), 3370–3384. <https://doi.org/10.1152/jn.00739.2011>
- Xia, Y., Leib, A. Y., & Whitney, D. (2016). Serial dependence in the perception of attractiveness. *Journal of Vision*, 16(15), 28. <https://doi.org/10.1167/16.15.28>
- Zhang, S., Huang, H. C., & Yu, A. J. (2014). Sequential effects: A Bayesian analysis of prior bias on reaction time and behavioral choice. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 36(36). <https://escholarship.org/uc/item/9jp3j5c7>

Open practices statement The datasets generated during and analyzed during the current study are available in the Open Science Framework repository (<https://osf.io/w4kcj/>). Experiment 1 was not preregistered.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.