Supplementary materials

Experiment 1: Population distribution and stability of state variables

Methods

In order to obtain large samples, Experiments 1 and 3 were performed outside the laboratory, on participants' own computers, inside internet browsers. The experiments, implemented using the JavaScript language and parts of the HTML5 toolset, such as the canvas element for graphics, produced high-quality, accurately timed and deterministic stimuli in a variety of internet browsers and operating systems (inadequately performing browsers, and devices such as telephones and tablets, were disallowed). Participants did not have to install any special software, because JavaScript/HTML5 interpreters are built into modern web browsers. While performing the experiments, participants were encouraged to close other programs and to put their browsers into full-screen mode; we recorded frame rate and browser window size, in order to eliminate participants in whom these parameters were too low. In these conditions, necessary to obtain large samples, we could control stimulus size in pixels, but had little control over physical stimulus size on the monitor, which depended on the size of the pixels, or optical stimulus size, which depended distance between stimulus and observer. Participants were instructed to position themselves upright and at a comfortable distance with respect to their monitors.

Each participant performed a single session that consisted of the SFM and TFM blocks performed in random order. Participants were shown written instructions for each of the two blocks, and given preliminary practice trials. The main part of each block consisted of 64 trials, with the first and second halves of each block exact replicas of one another (including the order of the trials). The duration of each stimulus was 0.5 s. Stimuli were preceded by a central fixation cross (size 30 pixels, duration 0.75 s). All stimuli consisted of 200 dots in motion; the dots were displayed as approximate antialiased black disks with diameter of 4 pixels, on a medium gray background. The median duration of each of the two experiments was about 3 minutes.

The SFM stimulus consisted of dots randomly and uniformly distributed on a virtual planar disk with diameter of 400 pixels. The disk had a slant of 45 deg and a tilt that had one of the 8 equally spaced values 11.25, 33.75, ..., 168.75 deg (for definitions of slant and tilt, see Fig. 1i). On each frame, the dots were projected onto the surface of the monitor using orthographic or parallel projection. Thus, the slanted disk appeared in projection as an ellipse whose minor axis was parallel to the tilt. The disk rotated at a constant angular speed, 30 deg/s, from −7.5 to +7.5 deg about its central orientation. The axis of rotation lay in the image plane (therefore rotations were in depth), and was oriented at either −45 or +45 deg with respect to tilt. (Thus, the instantaneous tilt during the rotation differed from its central value during the rotation, but by no more than 6.1 deg.) The reason for 45 deg difference was because it provides a particularly salient experience of 3D orientation, important for an internet experiment. Each combination of tilt and axis was repeated four times, in a randomized factorial design. Each stimulus sequence was followed by two icons at the bottom of the window, unambiguously depicting surfaces with tilts T and T + 180 deg (similar to those depicted on the rim of Fig. 1c; order randomly chosen). Participants selected the surface that corresponded most to the mean orientation that they perceived. The SFM stimulus has an exact symmetry: simultaneous reversal of relative depth (thus, changing tilt T to T + 180 deg) and the direction of motion left the

stimulus unchanged; thus, a stimulus with tilt T was also equally compatible with tilt T + 180 deg (see Fig. 1a, b for an illustration, and ref. (1)).

The TFM stimulus consisted of two subsets of 100 random dots moving linearly in opposite directions in the image plane, i.e., in directions A and A + 180 deg, at a speed of 60 pixels/s. The motion directions of one of the two subsets of dots were the 16 equally spaced values 5.625, 16.875, ..., 174.375 deg, with each direction repeated four times. The dots were constrained to lie inside a circular area of diameter 400 pixels, and when a dot exited the circle it reappeared on the opposite side, keeping the total number of dots fixed. The disk containing the stimulus was surrounded by a light gray annulus with outer radius of 250 pixels. Following the stimulus with motion directions A and A + 180 deg, participants selected one of two icons (depicted on the rim of Fig. 1f) that corresponded to the direction of motion of the layer seen in front. There were no depth cues to distinguish the two subsets of dots (intersecting dots were drawn in black as all other dots), but most observers perceive the dots as segregated into two transparent layers by their motion direction, with one of the layers seen in front or perceived as more salient (2). Thus, each stimulus was equally compatible with either direction A or A + 180 deg as being perceived in front.

The JavaScript code that ran the experiment in each participant's browser pre-rendered all the images in the animated sequence for each trial at the start of the trial, and then subsequently displayed the images as needed, using the HTML5 canvas element. We were able to empirically measure framerate on each session and were thus able to verify that we consistently achieved high framerates (around 60 Hz) on nearly all compatible internet browsers and platforms (the small number of participants with framerates below 30 Hz were excluded). At the conclusion of each experiment, data was sent from the participant's internet client to the server using an invisible HTTP request.

Unpaid volunteers were recruited by e-mail from a large French subject pool, by word-of-mouth, and through social networks. Two weeks after completing the experiment, participants were asked by e-mail to run the experiment again; about one year later, participants were asked to repeat the experiment once again. We received data for at least one experiment from 713 people in the initial experiment, 269 in the 2-week repetition, and 179 in the 1-year repetition. About 4% of these experimental sessions were eliminated from further analysis due to low frame rate (below 30 Hz), slow response (over 10 min total duration for either experiment), and small display window (height below 500 pixels, total fraction of monitor area below 50%). This left 704 participants in at least one experiment in the initial set, 235 in the 2-week repetition (median delay: 14.1 days), and 175 in the 1-year repetition (median delay from first session: 372 days). Overall, about 94% of the participants completed both the SFM and TFM experiments, while the rest completed only one of the two. Participants reported their gender, age, and manual preference (choosing from right- and left-handed). Among the participants in the initial sample, 61.6% declared that they were women (70.2% in the 2-week repetition, 66.2% in the 1-year repetition). Declared median age in the initial sample was 31 years, with ages ranging from 8 to 81 (median age 33.5 years in the 2-week repetition, 32 years in the 1-year repetition). 10.5% of the participants in the initial sample declared that they were left-handed (8.9% in the 2-week repetition, 9.1% in the 1-year repetition).

Data analysis

The calculation of biases is illustrated in Fig. 1g, h. If R_i were the responses (tilts in SFM or motion directions of the closer layer in TFM) on trial i with a total of n trials, we calculated the angular mean $\sum_{i=1}^{n}(\cos R_i, \sin R_i)\ /n$. In the case of ν equally-spaced directions around the half-circle, with the angles either parallel or anti-parallel to each of these directions, the maximum length of the angular mean is $[\nu \sin(\pi/2\nu)]^{-1}$, occurring for a pattern in which all angles lie in a 180 deg fan. We therefore normalized the angular mean by this quantity. We call the result the *bias vector.* We used the length of the bias vector, ranging from 0 for completely isotropic responses to 1 for fan-shaped patterns, as a measure of *bias strength*. For random responses, mean bias strength asymptotically approaches $\sqrt{2/\nu}$ for large ν (3). We used the direction of the bias vector as a measure of bias direction, which we also call *preferred tilt* (SFM) and *preferred motion direction* (TFM). We used the Rayleigh test to check if data from a particular session was significantly biased (anisotropic), and Watson's U^2 test to check if two sessions had significantly different biases (3, 4). Graphs of population distributions of bias directions were smoothed using a semicircle kernel with width ±15 deg. When performing multiple tests, here and elsewhere, we used the Benjamini-Hochberg procedure to control false discoveries, with the false discovery rate set to 0.05 (5).

Results

Raw data of all participants is shown in Fig. S1. The overwhelmingly common pattern is for the responses to be arranged in a 180 deg fan-shaped configuration: angles (tilts for SFM or motion directions for TFM) lying within $\pm$90 deg of a central direction are reported as having been perceived (shown as dark arrows in Fig. S1), and those lying in the opposite fan are rarely perceived (light arrows). Mean bias strengths were 0.873 in SFM, and 0.872 in TFM. However, the central directions—which we will call bias directions—varied enormously from one observer to the next. A randomly chosen pair of participants in our internet experiment would report opposite tilts in a random SFM stimulus about 43.9% of the time, and opposite motion directions in a TFM stimulus 46.7% of the time.

The simplest model, in which the observer reports the tilt or direction that has a positive projection on his or her own bias vector, accounts for 89.3% and 90.2% of all responses in SFM and TFM, respectively.

We wished to calculate the ambiguity of the stimuli with respect to the bias directions: how variable are the responses of each participant when shown the same stimulus? We used the fact that each tilt was repeated 8 times (SFM) and each motion direction 4 times (TFM) for each participant. We calculated the variance of responses within these repetitions and then averaged across participants, first adjusting the angles (tilts, motion directions) so that 0 deg corresponded to each participant's bias direction, and 180 deg to the direction opposite the bias. The results, shown in Fig. S2, show that nearly all the ambiguity is concentrated at the borders of the $\pm$90 deg fan-shaped regions. Close to the bias (and the direction opposite the bias) participants nearly always reported the same tilt or motion direction, the one having a positive projection on the participant's personal bias vector. Nearly all of the *perceptual* ambiguity was concentrated at the edges of the fan-shaped regions. Informally, many observers are surprised to learn, following the experiment, that all (or even any) of the stimuli were ambiguous.

The first and second halves of each block, for SFM and TFM, were exact replicas of one another (including the order of the trials). Analyzing the first and second halves of each block separately allowed us to estimate the reliability and stability of the bias measures. For SFM, we found that the preferred tilts calculated from the first and second halves of each block differed by a median value of 10.3 deg, with this difference below 37.6 deg for 90% of the participants. For TFM, the median difference was 6.2 deg, with 90% of the participants below 16.8 deg.

The distribution of the SFM and TFM bias directions is shown in the table below, as percentages of the population in ±45 deg quadrants centered on the cardinal orientations. The table also shows the corresponding 95% confidence intervals (bootstrap, 10^4 samples).

SFM		**TFM**	
right wall	18.4 [15.5, 21.3]	right	36.1 [32.4, 39.7]
ceiling	20.8 [17.8, 23.9]	up	18.2 [15.2, 21.7]
left wall	6.1 [4.3, 8.0]	left	14.8 [12.2, 17.6]
floor	54.7 [51.1, 58.5]	down	30.9 [27.4, 34.3]

Table T1. Percentages of the population with SFM preferred tilts and TFM preferred directions within 45 deg of the cardinal directions. 95% confidence intervals are shown in brackets.

As can be seen from the table, the following asymmetries in the population distributions are statistically significant: SFM right > left wall; SFM floor > ceiling; TFM right > left; TFM down > up.

In order to quantify the degree to which the population distributions have peaks in the cardinal directions, we calculated a *cardinality index,* defined as the fraction of the population with bias directions within ±22.5 deg of the cardinal orientations. This covers a total of 180 deg, and thus, if there were no preference for the cardinal directions, the cardinality index would be, on the average, 0.5. For SFM we find a cardinality index of 0.729 [0.696, 0.763], and for TFM 0.685 [0.650, 0.720] (95% confidence intervals in brackets, bootstrap with 10^4 samples). Thus, we exclude the 0.5 null hypothesis value: both distributions are significantly peaked in the cardinal directions.

Participants reported their gender, hand preference, and year of birth. In order to analyze the effects of these variables, we compared distributions of preferred tilts and directions different subpopulations, shown in Fig. S3. Distributions of preferred SFM tilts and TFM directions did not differ significantly between male and female participants (Watson's U^2 test, p = 0.5 and 0.4, respectively). However, men had a slightly but significantly stronger SFM bias then women (median values 0.956 and 0.934, Mann-Whitney test p = 0.005); TFM bias strengths did not differ. Preferred directions and bias strengths did not differ significantly between self-reported left- and right-handed participants. (The asymmetry between left- and right-wall preferences did seem to be weaker among left-handed participants, but this effect did not reach significance.) We split the participants by median age into younger and older groups. The distribution of preferred TFM directions differed significantly the two groups (Watson's U^2 test, p = 0.01): the down-up asymmetry was stronger in the older group, and the right-left asymmetry weaker, then in the younger group. Neither the distribution of SFM preferred tilts, nor the bias strengths, differed significantly between the older and younger groups.

In order to check whether biases were stable *within* single sessions (mean duration 3 min), we performed the following analysis. We will denote the stimulus orientations (tilts in SFM or motion

directions in TFM) on the 64 trials of the session as $\{S_1, S_2, \dots, S_{64}\}$, and the responses as $\{R_1, R_2, \dots, R_{64}\}$. Thus, $R_i = S_i$ or $R_i = S_i + 180$ deg. Within each block, we calculated the mean absolute angular difference of responses at a given lag λ trials, as a function of lag:

$$D(\lambda) = \overline{|R_i - R_{i+\lambda}|} = \frac{1}{N-\lambda} \sum_{i=1}^{N-\lambda} |R_i - R_{i+\lambda}|$$

where differences between angles are understood in the circular sense. Since the stimuli $\{S_1, S_2, \dots, S_{64}\}$ sample directions from 0 to π in random order, for any given participant the function $D(\lambda)$ will be swamped by the noise arising from the random choice of stimuli S_i and $S_{i+\lambda}$. Nevertheless, with data from many hundreds of participants available, by averaging $D(\lambda)$ over the participants, we may expect at least some signal to emerge from the noise. In particular, if the bias is constant, then lag should have no effect, and $D(\lambda)$ should be flat. Any systematic deviation from flatness, on the other hand, would signal a systematic change in bias.

The results of this analysis are shown in Fig. S4. The black curve shows mean $D(\lambda)$ over the participants, while the gray region shows the 95% between-participant confidence interval (bootstrap, 10^4 samples) of the mean. The effect is small but, given the large number of participants, very robust. The slow but significant rise in $D(\lambda)$ shows an effect of temporal proximity or order: responses on trials closer together in time tend to be closer than responses farther apart. The red dashed lines show the null-hypothesis level of D assuming no effect of order, obtained by randomly re-ordering the trials (10^3 re-orderings per participant), and taking the mean over the re-orderings. This yielded an asymptotically flat curve, which was then averaged over lags λ and over participants.

The evolution of preferred SFM tilts and TFM directions over 2 weeks and over 1 year is shown in Fig. S5. Fig. S5a shows that the bias directions in most participants undergo very little change even after a year, while a substantial minority does show change. Fig. S5b shows that the population distribution of changes in bias direction spreads or widens over one year, nevertheless retaining a peak at zero. Analyzing individual response distributions, we found that 9.1% of the repeat SFM participants and 10.2% of the repeat TFM participants had significantly different response distributions after two weeks, and 23.0% and 27.5% respectively after one year (Watson's U^2 test (3), Benjamini-Hochberg correction for multiple tests (5) with false-discovery rate 0.05).

Could the small size of the steps in the bias even over a year's time have been due to the non-uniformity of the population distribution of the biases? Consider, as a limiting case, biases distributed in a single, narrow peak. When measured a year later, biases will be very similar to their earlier values—but this similarity would be mainly due to the narrowness of the population distribution. To tease apart individual and population coherence, we calculated the difference between an initial bias and a bias measured later (2 weeks or 1 year), but instead of calculating this difference for biases in the *same* participant, we took the mean over *all pairs* of participants. If the small steps were due to population rather than individual coherence, then computing the difference for all pairs of participants should yield the same distribution as differences computed within participants. For 1-year steps, when averaging over all pairs of participants, we found a mean difference of 55.0 deg (95% between-participant confidence interval [42.1, 67.6] deg) for SFM, and 83.9 [72.9, 95.5] deg for TFM. For a completely uniform distribution, the difference would be 90 deg. The fact that the difference is significantly below 90 deg for SFM shows the effect of population-level coherence (the large peak for floor-like tilts). However, both confidence intervals exclude the 1-year steps computed within

participants, namely 22.5 [16.8, 29.3] and 26.3 [19.8, 33.8] deg for SFM and TFM respectively. Therefore, the small size of the steps in most participants, even after one year, demonstrates individual- rather than population-level coherence.

Experiment 2: Effect of biases on unambiguous stimuli

Methods

In this experiment we checked whether individual biases affected surface perception in SFM stimuli even when tilts were specified by suprathreshold horizontal disparity cues. SFM stimuli were similar to those used in the other experiments. Tilts were 32 equally spaced values 5.625, 16.875, …, 354.375 deg (the values range over 360 deg rather than 180 deg because most stimuli in this experiment no longer had the 180 deg tilt ambiguity). Horizontal disparities were applied as follows. Stimuli were displayed separately to each eye in a sequential stereoscope. We used a Sony GDM F520 CRT monitor, equipped with a full-screen active alternating circular polarization filter and passive filters worn by participants for stereoscopic separation (Z Screen, Stereographics), operating at 120 Hz (60 Hz for each eye). Stimuli were displayed in red, the monitor's fastest-decaying phosphor, in order to avoid stereoscopic ghosting. Individual frames for each eye were calculated using orthogonal projection parallel to the vector between each eye's notional position and the fixation point in the center of the stimulus. Participants' interocular spacing was measured, and on any given trial we controlled disparity by using an interocular spacing equal to 0, 0.1, 0.2, 0.3 or 0.4 times the actual spacing. With the stimulus sizes used (3.5 deg of visual angle if the surfaces had been frontoparallel), the surface slant (45 deg), and the monitor distance (participants used a chinrest to maintain monitor distance at approximately 57.3 cm), this works out to maximum horizontal disparity of 0, 1.6, 3.1, 4.7, and 6.2 deg of visual angle, respectively. Participants reported perceived surface orientation using the same images as probes as in the other experiments (giving a binary choice between tilts T and T + 180 deg, where T is the actual tilt) with unambiguous monocular depth cues and zero disparity. The experiment was a randomized factorial design, with 32 tilts and 5 disparity levels each presented once, giving 160 trials.

A total of 16 volunteers participated in the experiment. Because we were interested in the decrease of bias with increasing disparity, we excluded from further analysis the data of 4 participants whose bias at zero disparity was not statistically significant, as revealed by the Rayleigh test, leaving us with data from 12 participants.

Results

We used bias strength to measure the effect of individual biases on the perception of the unambiguous SFM stimuli. Bias strength near 1 indicates that responses follow 180 deg fan-shaped patterns, and that consequently internal bias dominates response patterns. If observers randomly and independently choose between tilts T_i and T_i + 180 deg on each trial, mean bias strength should be approximately 0.247 (a number obtained through simulation, for N = 32 tilts). Finally, if the disparity depth cue dominates any individual biases, bias strength should be close to 0, because the tilts in our experiment are uniformly distributed in the 360 deg range. In general, we expected bias strength to decrease with increasing disparity. The main question is whether the individual biases continue to

influence perception event when disparity is above threshold; in other words, do there exist suprathreshold disparities at which the bias strength is significantly above zero (or above chance level)?

The results, shown in Fig. S6, demonstrate that, in most participants, individual biases persist significantly beyond disparity threshold. At the group level, mean bias strengths are significantly positive, and significantly above chance level, for all values of disparity, up to 6.2 arcmin (t test). Although estimates of stereoacuity differ, there is agreement that disparities above 2 arcmin are above threshold for a large majority of the population (6).

Experiment 3: Long-term temporal dynamics

Methods

Stimuli were similar to those used in Experiment 1. The SFM tilts were 24 equally-spaced values 3.75, 11.25, ..., 176.25 deg, with rotation axes that differed by ±45 from the tilt. The TFM directions were 48 equally-spaced values: 1.875, 5.625, ..., 178.125 deg. Each experiment consisted of 48 trials, and took about 2 minutes to complete. 120 participants were recruited by e-mail from a large French subject pool, and were asked to commit to doing a 5-minute internet experiment every day, including weekends and holidays, for 3 months. Participants were paid 1 euro per daily session. They received multiple automated e-mail reminders every day to do the experiment. Participants were asked to, inasmuch as possible, perform the experiment every day at the same time of day and in the same conditions. Thirteen participants dropped out after less than two months, and their data was excluded from further analysis. Following the experiment, participants were asked by e-mail whether they performed the experiment every day themselves, or if they ever asked anyone else to do it for them (they were told that the answer to this question would have no practical consequences for them). Ten participants admitted to asking others to do the experiment once or more, and were excluded from further data analysis. The remaining 97 participants completed a mean of 87.8 (nearly) daily sessions, with 2.5% of the daily data missing. (With 96 trials per session, the resulting data set contained approximately 0.82 million trials.) In order to evaluate the extent to which participants followed the instructions to perform the experiment at the same time every day, we calculated the standard deviation of time differences between adjacent daily sessions, excluding any 'holes'. The median value over participants was 0.176 days, whereas if participants performed the experiment at random times every day, the standard deviation would have been $1/\sqrt{6} \simeq 0.408$ days.

Of the 97 participants that were retained for analysis, 63% declared themselves to be women. Self-declared ages ranged from 18 to 66, with a median of 28. 12% declared that they were left-handed. Ten of the 97 participants had participated in Experiment 1.

Data analysis

We modeled the time series using the Box-Jenkins ARIMA framework (7). We separately analyzed the series corresponding to each of the two components (*X, Y*) of the bias vector for each of the two biases (SFM, TFM), resulting in four time series for each participant. We filled in missing values (about 3% of the daily samples) using linear interpolation applied separately to the two components of the series. In the first step of our analysis, in order to handle a large number of series, we adopted an

automatic approach to model identification using the `forecast` software package for *R* (8). The software uses the KPSS unit-root test (9) to determine the order of differencing that leads to a stationary series (i.e., the *d* parameter in ARIMA(*p, d, q*)). It then searches the *p, q* space of models for the model that best fits the series, subject to penalties from the Akaike information criterion (8). To model each of our series we searched thorugh the 8 possible models with $0 \leq p, d, q \leq 1$. We checked goodness-of-fit using the Ljung-Box test (7) for the first 25 residuals, using the criterion $p > 0.05$ for passing the test (i.e., greater than 95% of the residual being white noise). Conservatively, we performed no correction for multiple tests here because *passing* the Ljung-Box test was a common event.

In the second step of the ARIMA analysis, we fit all series with the most common model identified during the first step, the ARIMA(0, 1, 1) or IMA(1, 1) model. The fits were performed with the standard `arima()` function in *R*, using the default conditional-sum-of-squares to find starting values, then maximum likelihood to find parameters. The series were first differenced once, and then fit to an MA(1) model with an intercept term, in order to allow for drifts. As in the previous step, we checked goodness-of-fit using the Ljung-Box test.

Results

The raw time series of the 97 participants are shown in the graphs in Table T2. Mean bias strengths over all participants and sessions were 0.922 and 0.892 for SFM and TFM, respectively.

We modeled our time series using the Box-Jenkins ARIMA framework (7), which quantifies state in several different ways. The general ARIMA(*p, d, q*) model for time series $\{x_t\}$ is given by

$$\nabla^d x_t = \sum_{i=1}^{p} a_i x_{t-i} + \sum_{i=1}^{q} b_i \epsilon_{t-i} + \epsilon_t \qquad (1)$$

where $\nabla x_t = x_t - x_{t-1}$ is the differencing operator and ϵ_t are independent identically-distributed random variables drawn from a normal distribution. In this framework, the null hypothesis of white noise corresponds to the ARIMA(0, 0, 0) model—each sample is independently drawn from an unchanging distribution. If an integral component is present ($d > 0$), then the series exhibits state by summing external perturbations, like a random walk. An autoregressive component ($p > 0$) also leads to steps that depend on past states, but with influence of past states that decays over time, which distinguishes it from a random walk. A moving-average component ($q > 0$) generates correlations through a multi-sample linear filter (7).

Fig. S7 shows the autocorrelation function of the daily differences, which can be seen to go to zero after lag of 1 or at most 2 days. This behavior strongly suggests that the 1-lag differences are stationary, and that our time series can probably be modeled by ARIMA(0, 1, 1) or IMA(1, 1) processes (7).

In the first step of our ARIMA analysis, we used an automatic model identification procedure, as described in the Methods. The frequencies of each of the 8 possible models are shown for SFM and TFM series in Table T3, together with the corresponding model parameters. The fits were very good, with 95% and 94% of the SFM and TFM series, respectively, passing the Ljung-Box test. Around 21% of the time series are best modeled by ARIMA(0, 0, 0), which corresponds to the null hypothesis of daily samples independently chosen from an unchanging distribution—in other words white noise. All other models correspond to different types of cumulative change. The most common model,

chosen for 41% of the time series, is ARIMA(0, 1, 1) or IMA(1, 1), which corresponds to a random walk measured through noise. The ARIMA(1, 0, q) models, which account for about 20% of the time series, correspond to an auto-regressive process, in which random changes are accumulated over time, but in which the accumulated changes also decay—in other words, a random walk that is also attracted to a fixed point. The ARIMA(1, 1, q) processes, which model about 12% of the time series, are random walks in which the daily steps are correlated among themselves. The ARIMA(0, 0, 1) model, found for about 5% of the series, is a moving-average process, in which the daily values are correlated to previous values through a multi-day linear filter of an underlying series with independent samples. Finally, the ARIMA(0, 1, 0) model, found for about 1% of the series, is a pure random walk.

Model			**SFM**			**TFM**		
p	d	q	Freq.	a_1	b_1	Freq.	a_1	b_1
0	1	1	41.8%		−0.775 [0.143]	39.7%		−0.778 [0.197]
0	0	0	18.0%			23.7%		
1	0	0	11.9%	0.359 [0.196]		10.3%	0.267 [0.159]	
1	0	1	8.8%	0.857 [0.097]	−0.635 [0.131]	10.8%	0.600 [0.559]	−0.344 [0.528]
1	1	1	10.3%	−0.070 [0.408]	−0.621 [0.499]	8.8%	0.037 [0.394]	−0.728 [0.249]
0	0	1	5.7%		0.235 [0.077]	4.6%		0.271 [0.342]
1	1	0	2.6%	−0.499 [0.220]		1.5%	−0.268 [0.430]	
0	1	0	1.0%			0.5%		

Table T3. Summary of the fitted ARIMA models for the 3-month time series experiment. For each of the eight possible models, we give its frequency and mean auto-regressive (a_1) and moving-average (b_1) coefficients (as well as the coefficients' between-participant standard deviations in brackets), separately for the SFM and TFM experiments.

In the second step of the analysis, we fit all series to the most common model, ARIMA(0, 1, 1) or IMA(1, 1). A large majority of the fits passed the Ljung-Box goodness-of-fit test (93% for SFM, 88% for TFM). For the coefficient b_1 of the ARIMA(0, 1, 1) model (see Equation (1) above), we found a mean value of $-0.76\,[-0.79, -0.73]_{95\%\text{ conf.}}$ for SFM (95% confidence interval in brackets) and $-0.78\,[-0.81, -0.75]_{95\%\text{ conf.}}$ for TFM, with between-participant standard deviations of 0.20 and 0.21, respectively.

It can be shown that the most common identified model, IMA(1, 1), is identical to a noisy random walk. Consider a pure random walk or ARIMA(0, 1, 0) process u_t, defined by

$$u_t = u_{t-1} + \alpha_t$$

where the random walk steps α_t are independently drawn identically distributed (i.i.d.) variables from the normal distribution $N(0, \sigma_\alpha)$. Suppose that this process is not directly observable, but is measured through another process, v_t, which adds white noise to u_t:

$$v_t = u_t + \beta_t$$

where the measurement noise β_t is assumed to be iid with $N(0, \sigma_\beta)$. We then have

$$v_t - v_{t-1} = \alpha_t + \beta_t - \beta_{t-1}$$

Since the right-hand side of this equation contains only a linear combination of normally distributed variables, and its covariance vanishes for lags greater than 1, we know that $v_t - v_{t-1}$ must be an ARIMA(0, 0, 1) or MA(1) process, and therefore v_t is ARIMA(0, 1, 1) or IMA(1, 1). To reconstruct the properties of α_t and β_t from that of v_t, we express the latter as a general ARIMA(0, 1, 1) process:

$$v_t - v_{t-1} = \gamma_t - \theta\gamma_{t-1}$$

where γ_t are i.i.d. with distribution $N(0, \sigma_\gamma)$. By equating the lag-0 and lag-1 covariances of the above two equations, we find

$$\sigma_\alpha{}^2 + 2\sigma_\beta{}^2 = (1 + \theta^2)\sigma_\gamma{}^2$$
$$\sigma_\beta{}^2 = \theta\sigma_\gamma{}^2$$

If we know the parameters of the noisy series v_t—namely θ and $\sigma_\gamma{}^2$—we can obtain the variance of the measurement noise $\sigma_\beta{}^2$ from the last equation, and the variance of the random walk steps $\sigma_\alpha{}^2$ from

$$\frac{\sigma_\beta{}^2}{\sigma_\alpha{}^2} = \frac{\theta}{(1-\theta)^2}$$

The argument is taken from ref. (7), pp. 131-5, but note that their final equation, A4.3.14, has an error. The random walk model observed through white noise is also known as the *local level model* (10).

We found moderate but robust correlations between the daily *steps* in the SFM and TFM biases, both within and between participants. We calculated the squared Euclidean distance between adjacent daily measurements of the bias vectors, and, within participants, found that 80 out of 97 participants had positive Pearson correlations between the squared amplitudes of the SFM and TFM steps (82%, mean correlation 0.26, 95% confidence interval of mean by bootstrap [0.20, 0.32]). We also found a significant between-participant correlation between mean daily steps: participants who had large mean squared steps in SFM biases also tended to have large steps in TFM biases (R = 0.42, 95% confidence interval by bootstrap [0.27, 0.61]).

Experiment 4: Short-term temporal dynamics

Methods

In this experiment, we tested only the SFM stimulus. The stimulus had the same geometry as described previously in terms of pixels, but the participant's head was placed in a chinrest at a fixed distance from the monitor, so that the stimulus subtended 9.7 deg of visual angle. In order to remove visual references with edges in the cardinal direction, the fixation marker was a circle (rather than the cross used in the other experiments). The experiment was performed in blocks of 64 trials. Each block had 4 repetitions of the tilts 5.625, 16.875, ..., 174.375 deg in random order. After a single block, which lasted about 2.5 min, the participant remained in complete darkness for 30 min. Following this, a series of blocks was performed over a 30 min period, with the final incomplete block discarded. (Participants performed the experiment at different speeds, which resulted in a minimum of 9 blocks

and a maximum of 15, with mean of 12.9.) Then followed another 30 min period in the dark, and finally one last block.

The experiment was performed in conditions in which nothing other than the stimuli was visible. This was achieved by carefully sealing the experimental room for light, and displaying the stimuli on an AMOLED monitor (LG 15EL9500), which features a nearly infinite contrast; the tiny amount of backscatter light was removed using a gelatin neutral density filter fixed to the front of the monitor. Fifteen volunteers participated in this experiment.

Results

We performed a number of analyses to check whether the 30-minute series were white noise or not. Each of the series was composed of 9-15 measures of the SFM bias vector (the variability was due to differences in participants' speeds). Each measure was derived from a single 64-trial block, with the number of blocks depended on the participant's speed in performing the experimental task. First, we calculated power spectra. Dropping the DC term and performing a linear regression in log-log coordinates—in effect fitting the spectra with a power function Af^{β}—we found a mean value of $\beta = -0.92$, with a between-participant 95% bootstrap confidence interval of $[-1.29, -0.54]$. Individually, the fitted value of β was negative in 13 out of the 15 participants. Second, we calculated the mean square Euclidean distance between bias vectors as a function of lag, and performed a linear regression of this measure versus lag. We obtained a mean slope of 0.037, with a between-participant 95% bootstrap confidence interval of [0.009, 0.072]. Individually, the slope was positive in 13 out of 15 participants. Third, because these results strongly indicate that the series were non-stationary, we differenced the series and calculated the autocorrelation functions of the differences. We found that the only term in the autocorrelation functions that was significantly different from 0 was the lag-1 term, whose mean was -0.24, with a between-participant 95% bootstrap confidence interval of $[-0.34, -0.14]$. Taken together, these three results indicate that the 30-minute series are not white noise, and that they evolve in a stateful way. The last result strongly points to the series being IMA(1, 1), similar to the 90-day series.

We calculated steps (as squared Euclidean distances) in successive measurements of the bias vector. The mean step between the first two measurements, separated by 30 min of darkness, was 0.101 (between-participant variance 0.029). The mean step between successive measurements during the 30-minute sequence of continuous stimuli was 0.267 (variance 0.127). The mean step between the last two measurements (again separated by 30 min of darkness) was 0.457 (variance 0.841). The difference between the first dark step and the mean steps during the 30-min sequence was significant: a bootstrap calculation (10^4 samples) revealed that the 95% confidence interval of this difference was $[-0.35, -0.05]$. None of the other differences in the step sizes was significant, due to the large between-participant variance of the 30-min steps and the second dark steps. The difference in between-participant variance between the first dark step (0.029) and the 30-min steps (0.127) was also significant: a bootstrap calculation (10^4 samples) revealed that the 95% confidence interval of this difference was $[-0.23, -0.01]$. None of the other differences in variance was significant. Thus, we have evidence that steps in the SFM bias vectors are significantly larger during uninterrupted presentation of the brief stimulus sequences over 30 min than when two measurements are separated by 30 min of darkness. In other words, the dynamics of the bias vector seems to be 'excited' by presentation of log sequences of SFM stimuli.

We also wished to test whether the peaks in the cardinal directions of the population distributions of SFM bias directions (Fig. 2c) were due to the simultaneous presence of other visual stimuli with cardinally oriented edges, or whether this population-level preference was due to internal factors. The internet-sourced population distribution experiment was performed in a variety of conditions, but presumably at least the cardinally oriented monitor edges were usually visible during the experiment, and perhaps other edges as well. It has been shown that the distribution of visual edge orientations in typical environments is peaked in the cardinal directions, and that these inhomogeneities influence visual biases (11). In the population distribution experiment, we found that 72.9% of our participants had bias directions within ± 22.5 deg of the four cardinal directions—and the difference between this fraction and the 50% expected for a uniform distribution was significant. In this experiment, we removed all simultaneously visible edges, including the monitor edges, by performing the experiment in a dark, light-sealed room, and using an AMOLED monitor with nearly zero minimal luminance, and removing the backscattered light from the stimulus using a filter. We found that 75% of the biases measured in this experiment were within ± 22.5 deg of the four cardinal directions. This fraction was significantly greater than the null-hypothesis value of 50% (sign test, $p < 0.0001$), and not significantly different from the value in the population distribution experiment (Mann-Whitney test, $p = 0.54$). We conclude that the population-level peaks in the SFM biases in the cardinal directions also exist without any additional visual stimuli, and are therefore due to internal factors.

Experiment 5: Deliberate perturbation of state variables

Methods

In this experiment we used both SFM and TFM stimuli. The experimental blocks were of two types: perturbation and test. In the test blocks, the stimuli were the two-fold ambiguous stimuli used in the other experiments, with some small modifications detailed below. The perturbation stimuli used the same underlying objects (planar surfaces rotating in depth for SFM, two layers of dots moving in opposite directions for TFM), but in which binocular disparity and other cues disambiguated depth structure, with perspective rather than orthographic projection used throughout. For SFM, the resulting binocular disparity, texture, and second-order motion parallax cues disambiguated tilt. For TFM, the two sets of dots moving in opposite directions were separated by a simulated depth of 1 cm; thus, binocular parallax unambiguously specified depth order. Stimuli subtended 8 deg of visual angle, and were displayed on a Sony GDM F520 CRT monitor, using a full-screen alternating circular polarization filter and passive filters worn by participants for stereoscopy (Z Screen, Stereographics). Stimuli were displayed in red, the monitor's fastest-decaying phosphor, in order to avoid stereoscopic ghosting.

The methodological difficulty of this experiment is that a typical bias measure takes at least a few tens of seconds, while pilot studies convinced us that the largest effect of the perturbation was very short-lived. We therefore devised a method that allowed us to average trials from multiple blocks. Since our general technique for measuring biases is based on uniformly sampling directions in random order, we used a latin-square design, in which both the trials across one block, and the *n*-th trials over different blocks, uniformly and randomly sampled directions. This design effectively increased temporal resolution of the bias measurement, from the duration of a block down to the

duration of a single trial. The detailed design of this experiment is illustrated in Fig. S10. Each participant completed 16 interleaved sessions in both SFM and TFM experiments, completing one or more sessions per day; when completing more than one session on a given day, at least 15 minutes elapsed between sessions. Each session consisted of 5 blocks, in the following order: test, perturbation, test, test, test. The transitions between blocks were seamless. Each block consisted of 16 trials, and lasted about 30 s. In SFM, the 16 trials in each block corresponded to the 16 equally spaced tilts 5.625, 16.875, ..., 174.375 deg, in random order; surfaces with tilt T rotated about an axis oriented at angle T + 45 deg in the image plane. In TFM, the motion directions were the same 16 equally-spaced values.

Following the initial test block, the participant's bias direction was determined (see Data analysis, below). On the following perturbation block, participants were shown surfaces with unambiguous tilts (SFM) or displays with front-layer motion direction (TFM) that lay in a 45 deg arc between 90-135 deg from the preferred tilt or direction. On alternate sessions the arc was offset clockwise or counter-clockwise from the bias direction. Each perturbation block consisted of 16 trials, with tilts or directions equally spaced between 90 and 135 deg. Thus, the mean value of the perturbation tilt or direction was 112.5 deg from the preferred tilt or direction, alternately clockwise and counter-clockwise.

Immediately following the perturbation block in each session, participants performed the three test blocks. Our technique for calculating bias (see below) is based on analyzing sets of trials with equally spaced tilts or directions, with each direction included the same number of times. The spatial resolution increases with the number of directions, so we cannot have too few. Including entire sets of trials, however, limits temporal resolution. In this experiment, we raised both spatial and temporal resolutions by analyzing trials from multiple sessions, as follows (see Fig. S10). In addition to the 16 tilts or directions being randomized within each block, the values were arranged in 3 Latin squares over the 16 sessions. In other words, the first tilt following the perturbation assumed each of the 16 possible values once and only once over the 16 sessions, so did the second tilt, and so on. This allowed us to analyze the 16 trials immediately following the perturbation in the 16 sessions as we would have analyzed the 16 trials of a single block. Thus, by analyzing all the first trials over the 16 sessions, all the second trials, etc.—rather than all the trials from a single block—we were able to raise temporal resolution from 30 s (the duration of a block) to 2 s (the duration of a trial), provided that the dynamic properties of the bias were invariant over the different sessions.

Eight volunteers participated in the SFM condition of the experiment, including one of the authors. The data of two participants were excluded from further analysis because one showed little sensitivity to the binocular cues used in the perturbation stimulus, and the other had very weak bias in the pretests. Seven volunteers participated in the TFM condition of the experiment, including two of the authors. The data of one of the participants was excluded because he reported, paradoxically, perturbation stimuli as having motion directions opposite to those specified by disparity. Three of the participants took part in both the SFM and the TFM conditions.

Data analysis

We analyzed all of the first trials after the perturbation across the 16 sessions, all the second trials, etc. Due to the Latin square design of the experiment, this was formally identical to analyzing the responses of all the trials in a single block. On each session we defined the angular zero as the as the

preferred tilt or direction as measured in the first test block of the session, and angles defined as positive in the direction of the perturbation, which alternated between clockwise and counterclockwise in successive sessions.

Results

Individual data and exponential fits are shown in Fig. S11. Biases, as measured in the pretests, were strong: mean bias strength was 0.98 for SFM and 0.96 for TFM. During the perturbation blocks, the binocular disparity was a strong enough cue so that 98% of all reported tilts in SFM were compatible with disparity, and so were 97% of reported motion directions in TFM. The exponential fits to the post-test series were reasonably good, with mean adjusted $R^2 = 0.29$ for SFM and 0.51 for TFM. The population-level 95% confidence intervals for the median initial strength of the perturbation effect were $[30.5, 45.2]$ deg (and thus $[27.1, 40.2]$% of the mean perturbation amplitude of 112.5 deg) in SFM, and $[14.7, 56.7]$ deg ($[13.1, 50.4]$%) for TFM (bootstrap with 10^4 iterations). The population-level 95% confidence intervals for the median half-lives of the perturbation were $[15.1, 40.7]$ s for SFM and $[28.6, 165.7]$ s for TFM.

Experiment 6: Effect of bistable fluctuations on biases

Methods

In this experiment we used only SFM stimuli. Each measurement block consisted of 16 trials, with equally-spaced tilts $5.625, 16.875, \ldots, 174.375$ deg shown in random order. The rotation axis was always offset by 45 deg counterclockwise. In all other respects, the stimuli were identical to the SFM stimuli in the other experiments. Stimuli were shown on a laptop monitor, with the participant placed at a self-chosen comfortable distance from the monitor, typically around 50 cm. Instead of using visual probes to respond, participants pushed a knob (SpaceMouse Pro, 3DConnexion) in a direction corresponding to the perceived tilt, having previously learned the relationship between tilts and knob directions. Measurement blocks had a mean duration of 28 s.

We performed measurement blocks until we found stable bias values in two successive measurements. Participants then completed one or more spontaneous fluctuation blocks. In these blocks, the same stimulus was shown on every trial, with tilt that corresponded to the participant's bias tilt, as measured in the final measurement block. In the fluctuation blocks there was a fixed interstimulus interval, with participants having to respond within this interval. (This was in contrast to the measurement blocks and all other experiments reported here, where participants had an arbitrary amount of time to respond following the presentation of each stimulus, and where the interstimulus interval was therefore variable.) If no response was given during the fixed interstimulus interval, data from the trial was dropped from subsequent analysis. The initial interstimulus interval was set at 1 s. If no spontaneous reversals were obtained after several minutes, the interstimulus interval was reduced to 0.5 s. (It has been shown that briefer interstimulus intervals lead to more frequent spontaneous fluctuations (*18*).) During the fluctuation blocks, the experimenter monitored the participant's responses, and in particular whether a reversal had occurred, using audio signals received through headphones. After a spontaneous reversal, with the participant reporting the tilt opposite to her or her own initial bias, and with the reversed tilt being reported for 10-20 consecutive

trials, the fluctuation block was stopped, and a measurement block immediately performed. Following the first post-fluctuation measurement, subsequent measurement blocks were performed over an interval of several minutes, until stable values of the bias were obtained. Seven volunteers participated in this experiment; the data of one participant was excluded from further analysis because he confused the convention between tilts and knob presses.

Results

Prior to the spontaneous fluctuation block, all subjects presented a consistent bias: bias directions in the final two measurement blocks differed by at most 14.6 deg, and mean bias strength was 0.89. During the spontaneous fluctuation block, 133 – 1283 s (mean 441 s) were required until a stable tilt reversal occurred (lasting at least 10 trials). After this reversal, we continued showing the participants the same stimulus (now consistently perceived with a reversed tilt) for a period lasting 24.4 – 79.7 s (mean 43.9 s). The bias directions measured in the blocks following the spontaneous reversal are shown in Fig. S12, as a function of time after the reversal. Bias direction in Fig. S12 is adjusted so that the original, pre-fluctuation bias for each participant is at 0 (and therefore the tilt perceived after the spontaneous reversal is 180 deg). As can be seen in Fig. S12, the state variables are measured near their pre-fluctuation values, either at once (i.e., on the first measurement), or at most after a few minutes. The final measurement shows a bias direction that differs between 6.0 and 17.6 deg (mean 10.7 deg) from its pre-fluctuation value—and thus is in all cases much closer to the pre-fluctuation value than the reversed value 180 deg away. Thus, in all 6 participants, the spontaneous fluctuations do not have a durable effect on the state variables. In four participants the effects of the fluctuation are either non-existent or are too brief to measure, while in the remaining two the state variables seems to decay to its pre-fluctuation value over several minutes.

Experiment 7: Existence of the bias before the first stimulus

Methods

In this experiment we used only SFM stimuli. Following very careful instructions and without any practice trials, each participant performed a block of 24 trials, with equally-spaced tilts 3.75, 11.25, …, 176.25 deg presented in random order. The rotation axis was always offset by 45 deg counterclockwise. In all other respects, the stimuli were identical to the SFM stimuli in the other experiments. Stimuli were shown on a laptop monitor, with the participant placed at a self-chosen comfortable distance from the monitor, typically around 50 cm. Thirty volunteers participated in the experiment, none of whom had participated in any of the other experiments described here, or had even seen the SFM stimuli.

Results

We calculated the *partial bias* vector, based on the first *i* trials:

$$\mathbf{B}_i = \frac{1}{N_i}\sum_{j=1}^{i}(\cos R_i\,, \sin R_i)$$

where R_i is the reported tilt on trial i, and $N_i = 1/[(i/2)\sin(\pi/i)]$ is the normalization factor, equal to the maximum length of the bias vector. Let β_i be the direction (angle) of $\mathbf{B}_i$; we the define the absolute difference between the direction of the partial bias and the full bias calculated over the entire session of 24 trials:

$$D_i = |\beta_i - \beta_{24}|$$

where the difference is understood to be angular. By definition, the final difference $D_{24} = 0$. As we shall see below, D_i is useful to calculate because it allows one to test two competing hypotheses: on one hand that there is a pre-existing bias before the start of the experiment, and participants respond in accordance with this bias; or, on the other hand, that the first perceived tilt (or the vector average of the first few perceived tilts) generates a sensory memory that in effect becomes what we call the bias.

The difference between partial and full biases, averaged over all 30 participants, is shown in Fig. S13, with the black curve showing a mean over participants, and the gray area the 95% between-participant confidence intervals (bootstrap with 10^4 iterations). The mean over participants is necessary because partial biases are calculated over a randomly chosen subset of tilts, and in each participant are heavily contaminated by noise arising from the random choice—in contrast to full biases, which is calculated over an evenly spaced set of tilts sampling all directions.

The red dotted curve in Figure S13 shows the predicted data, provided that the biases pre-exist the experiment. If the biases do pre-exist the experiment, then the order of trials should not matter. The predicted curve was therefore calculated by computing the partial biases over trials rearranged in random order (10^3 randomizations/participant). As can be seen in Figure S13, the predicted curve lies within the confidence region of the data.

The green dashed curve in Figure S13 shows the predicted data, provided that an unbiased random choice of tilt is made on the first trial, and that this perceived tilt becomes, through a mechanism of sensory memory, the bias for subsequent trials. The curve was calculated by fitting a logistic response model to mean data over participants:

$$p(R_i = T) = \begin{cases} a + (1-2a)/[1 + e^{-k(T-T_0+\pi/2)} & \text{if } T - T_0 < 0 \\ a + (1-2a)/[1 + e^{+k(T-T_0-\pi/2)} & \text{if } T - T_0 > 0 \end{cases}$$

where T_0 is the participant's full bias direction, and differences of angles are understood to be in the circular sense. Using a least-squares fit we found the slope $k = 4.39\ \text{rad}^{-1}$ and the lapse parameter $a = 0.046$. Using this model, we simulated data from 10^4 participants by randomly choosing one of the two possible tilts on the first trial, taking this tilt as the bias, T_0, and applying the above model to the remaining tilts (in random order) to calculate the predicted green dashed curve in Figure S13.

The predicted green curve has a minimum at trial 1, because by definition of the model, the partial bias on the first trial—the response tilt—*is* the full bias, and so D_1 should be close to 0. (Not exactly zero, because the bias is not all-or-none, and the model is therefore probabilistic.) The curve then rises, because the random stimuli in subsequent trials are not equal to the bias tilt; and then falls back to zero.

The responses predicted by this model disagree with the data, as can be seen in Figure S13. More complex sensory-memory models, in which the bias is not the first perceived tilt but the vector average of the first n perceived tilts, have predicted D_i curves with minima at n, which do not show up in the data.

Supplementary references

1. Ullman S (1979) *The interpretation of visual motion* (MIT Press, Cambridge, Mass.).

2. Mamassian P, Wallace JM (2010) Sustained directional biases in motion transparency. *J Vis* 10(13):23.

3. Mardia KV, Jupp PE (1999) *Directional Statistics* (Wiley, Chichester ; New York). 1 edition.

4. Watson GS (1962) Goodness-of-Fit Tests on a Circle. II. *Biometrika* 49(1/2):57.

5. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B Methodol*:289–300.

6. Howard IP, Rogers BJ (2008) *Seeing in depth* (Oxford).

7. Box GEP, Jenkins GM (2008) *Time Series Analysis: Forecasting and Control* (Wiley, Hoboken, N.J). 4 edition.

8. Hyndman RJ, Khandakar Y (2008) Automatic time series forecasting: The forecast package for R. *J Stat Softw* 27(3).

9. Kwiatkowski D, Phillips PCB, Schmidt P, Shin Y (1992) Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *J Econom* 54(1-3):159–178.

10. Durbin J, Koopman SJ (2012) *Time Series Analysis by State Space Methods* (Oxford Univ Pr, Oxford).

11. Girshick AR, Landy MS, Simoncelli EP (2011) Cardinal rules: visual orientation perception reflects knowledge of environmental statistics. *Nat Neurosci* 14(7):926–932.

Supplementary figures

S1. Raw data for all participants (N = 691 for SFM and 682 for TFM) in the population distribution experiment (Experiment 1). Each trial is represented as two opposite-facing arrows, corresponding to the tilt for the SFM stimuli and motion direction of closer layer for the TFM stimuli—both of which had a 180 deg ambiguity. The darker gray arrow shows the reported tilt (SFM) or reported direction of motion of the closer layer (TFM), whereas the lighter arrow shows the tilt or motion direction not reported. The thick red arrow shows the resulting bias vector. For clarity, only one-quarter of the trials for each participant are shown.

S2. Variability of responses to the same stimulus by the same observer, in the SFM and TFM population distribution experiment (Experiment 1). Each observer was shown the same tilt 8 times (SFM) and the same motion direction 4 times (TFM). We calculated the variance over these repetitions as a function of angle (tilt or motion direction), and then averaged across observers. Before averaging, we adjusted the angles by subtracting each observer's bias direction, so that 0 deg corresponds to the bias direction, and 180 deg to the direction opposite the bias. The black curves show the mean across observers as a function of deviation with respect to bias (variance in arbitrary units), and the gray regions the between-observer 95% confidence intervals of the mean (bootstrap). These graphs show that nearly all the variability of response to the same stimulus—in other words the ambiguity of the stimulus—is concentrated at the edges of the $\pm$90 deg fan-shaped bias regions.

S3. Effect of self-reported gender, hand preference, and age on distributions of SFM and TFM bias directions in the population distribution experiment (Experiment 1).

S4. Mean absolute difference of responses as a function of lag (Experiment 1). Responses in degrees, lag in number of trials. The black curves show means of participants, the gray region 95% between-participant confidence intervals (bootstrap, 10^4 samples). The red dashed lines show the null hypothesis predictions of no bias change, obtained by randomizing trial order.

S5. Evolution of preferred SFM tilts and TFM directions over two weeks and over one year (Experiment 1). (a) Bias directions for each repeat participant (participants shown as dots), for SFM (left) and TFM (right), two weeks later versus initial value (top) and one year later versus initial (bottom). Comparisons where Watson's U^2 test revealed a significant difference (with Benjamini-Hochberg correction for multiple tests with false discovery rate 0.05) are shown as red, and the rest as black. (b) Smoothed population distribution of absolute preferred tilt or direction differences over two weeks (dashed line) and over one year (solid line).

S6. Results of the unambiguous stimulus experiment (Experiment 2), showing the effect of individual SFM biases on perceived surface tilts, even when these are pitted against tilts unambiguously specified by suprathreshold binocular disparity. (a) Group data, showing bias strength as a function of horizontal binocular disparity (the curve shows means, the error bars between-participant standard errors). Bias strength 1 corresponds to disambiguation by

individual biases, 0.25 (dashed line) to random responses, and 0 to responses based on disparity rather than individual bias. (b) Individual data, with each ring corresponding to one participant. In each participant's data, the polar angle corresponds to stimulus tilt, and the radial variable to disparity, with the innermost ring corresponding to minimum (zero) disparity and therefore to perfect two-fold ambiguity (as in the other experiments in this study), and the outermost ring corresponding to maximum disparity (about 6.2 arcmin of visual angle). As usual, dark gray cells correspond to reported tilts, and light gray cells to unreported tilts. A ring with a 180 deg fan of reported (dark) and the opposite fan of unreported (light) tilts corresponds to a full effect of individual bias; a ring with all tilts reported (dark) corresponds to responses in full accord with the unambiguous stimulus, i.e., no effect of individual bias.

S7. The full and partial autocorrelation functions of the daily differences (steps) in the SFM and TFM bias vectors, in the long-term temporal dynamics experiment (Experiment 3). The curves show the means over all participants, the gray regions the between-participant 95% confidence intervals of the mean. The functions are only shown for lags up to 30 days, as they remain near zero for all larger lags. These curves strongly suggest that the daily differences are MA(1), so that the full series are IMA(1, 1).

S8. Distributions of daily bias directions (Experiment 3). The circular histograms have been smoothed using 10 degree kernels. These distributions are similar to those obtained in Experiment 1 (see Figs. 2c, f) on an almost completely difference sample (10 of the 97 participants here had also participated in Experiment 1).

S9. Timelines of SFM biases for the 15 participants in the short-term temporal dynamics experiment (Experiment 4). Biases were measured once, followed by 30 min of darkness, followed by 30 min of continuous SFM stimuli and bias measurements, followed by another 30 min of darkness, followed by one last measurement. The bias vector measured on each block (64 trials, mean duration 2.5 min) is shown as an arrow. Arrows are colored red for each pair of consecutive blocks whose biases differed significantly (Watson's U^2 test, Benjamini-Hochberg correction for multiple tests with false-discovery rate 0.05). The three cases of significant bias change in darkness are marked by blue asterisks.

S10. The design of the perturbation experiment (Experiment 5). The design allowed us to improve the temporal resolution of the measurements of the perturbation effect, from one block (approximately 30 s) down to a single trial (approximately 2 s). The experiment consisted of 16 session, with participants performing one or more session per day. Each session consisted of 3 phases: a preliminary block to measure the bias direction (O), followed by a block of perturbation stimuli (P), followed by 3 blocks to measure the effect of the perturbation on the bias direction (A, B, C). The goal was to be able to calculate a bias $\mathbf{M_1}$ from the ensemble of the first post-perturbation trials ($A_1^1, A_1^2, \ldots, A_1^{16}$), a bias $\mathbf{M_2}$ from the second trials ($A_2^1, A_2^2, \ldots, A_2^{16}$), and so on, thereby improving the temporal resolution of the time series down to the duration of a single trial. In order to carry out this kind of transverse bias measurement, the first trials, second trials, etc., needed to sample over all the stimulus angles (tilts for SFM, motion directions

for TFM). Additionally, we wished to calculate the bias from individual blocks (for example $A_1^1, A_2^1, \ldots, A_{16}^1$; $B_1^1, B_2^1, \ldots, B_{16}^1$; etc.), so they also needed to sample over all the stimulus angles. We accomplished this by using a design in which the post-perturbation blocks over all sessions, A_j^i, B_j^i, and C_j^i, were randomized 16 × 16 Latin squares.

S11. The effect of perturbation on bias direction, in individual participants (Experiment 5). Each graph shows the mean deviation of bias direction from its initial value (set to 0) in degrees, as a function of post-perturbation trial number. Data is averaged over 16 sessions, with the zero always set to the pre-perturbation bias direction measured at the start of each session, and the sign normalized so that positive deviations are in the direction of the perturbation (which was between +90 and +135 deg). The mean duration of each trial is 2.0 s, so the time scale in the graphs is about 100 s. The red dashed curves show exponential fits. The graphs on the left show data from the 6 participants in the SFM condition, and the graphs on the right the 6 participants in TFM.

S12. The effect of spontaneous fluctuations on the SFM bias variable (Experiment 6). In each participant we first measured the bias direction. We then repeatedly exposed the participant to the SFM stimulus with this tilt—but that could also be perceived as having reverse tilt (i.e., one that differed by 180 deg). At first all participants perceived the tilt corresponding to their bias, but all eventually underwent a spontaneous fluctuation, perceiving the reverse tilt in a stable way. We then re-measured their SFM bias direction over the subsequent minutes, and these are the measurements shown the graph. The time evolution of the bias is shown as a separate curve for each participant, normalized so that 0 corresponds to the initial bias direction (before the spontaneous fluctuation) and 180 deg to the bias direction after the fluctuation. Each point corresponds to a single block measuring the bias direction, with the horizontal axis corresponding to the time (minutes) following the spontaneous reversal. In all 6 participants, the bias direction returns—either immediately or gradually—to near its pre-fluctuation value.

S13. Difference between partial and full bias, as a function of trial number (Experiment 7). The black curve shows the mean of absolute difference over participants, in degrees. The gray region shows the 95% between-participant confidence interval (bootstrap with 10^4 iterations). The red dotted curve shows the predicted data, if the bias pre-existed the experiment. The green dashed curve shows the predicted data, if the bias were generated as a sensory memory by the perceptual decision on the first trial.

Supplementary tables

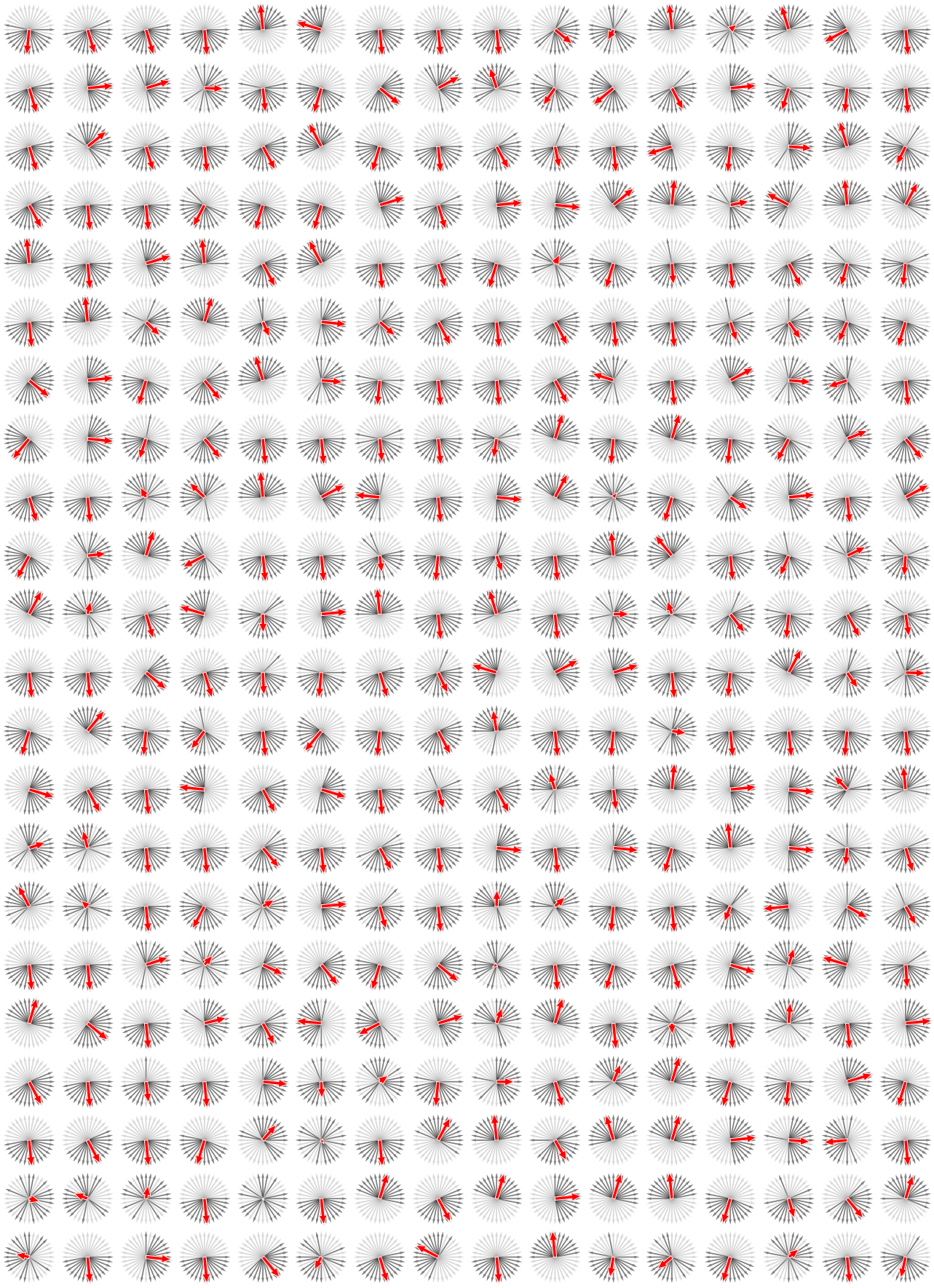
T1. [Inserted in text] Percentages of the population with SFM preferred tilts and TFM preferred directions within 45 deg of the cardinal directions. 95% confidence intervals are shown in brackets.

T2. Raw data and individual fitted ARIMA models for the 3-month time series experiment. Each line represents one of the 97 participants. Data for the SFM experiment is shown on the left, for the TFM experiment on the right. The graphs show the time evolution of the biases. The top graph (black curve) shows the bias direction time series (preferred tilt or motion direction, ordinate in degrees, abscissa in days). The bottom graph (blue curve) shows the corresponding bias strength. Periods with pink background correspond to days when responses were *not* significantly anisotropic, as measured using Watson's U^2 test with Benjamini-Hochberg correction for multiple tests at false discovery rate 0.05. At all other times (white background) responses were significantly anisotropic. Next to each graph are details of the chosen ARIMA model. The model for each series was selected from among eight possible models using an AIC criterion and unit-root tests; then the model's coefficients were estimated using maximum likelihood. The x and y components of the bias vector ($b \cos \theta$ and $b \sin \theta$ respectively, where b is bias strength and θ the preferred tilt or direction) were analyzed separately. The top line shows the model for the x component, the bottom line for the y component. The p, d, and q columns give the degrees of the auto-regressive component, the differencing operator, and the moving-average component, respectively (each could be 0 or 1). The next three columns show the appropriate coefficients of the model: a_1 is the coefficient of the auto-regressive term (if any), b_1 the coefficient of the moving-average term, and "int." the intercept or constant term.

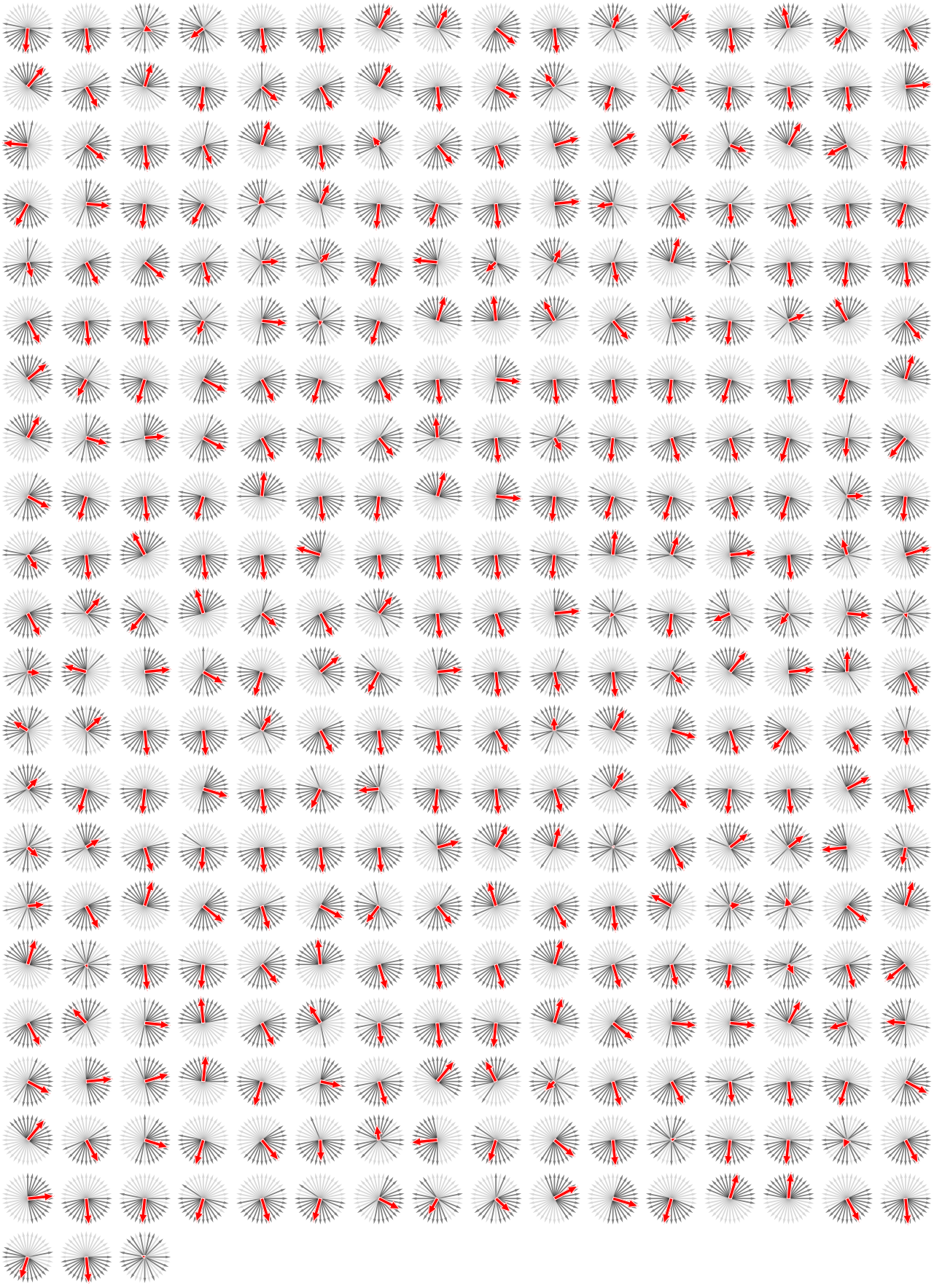
T3. [Inserted in text] Summary of the fitted ARIMA models for the 3-month time series experiment. For each of the eight possible models, we give its frequency and mean auto-regressive (a_1) and moving-average (b_1) coefficients (as well as the coefficients' between-participant standard deviations in brackets), separately for the SFM and TFM experiments.

Figure S1 (extends over next 4 pages)

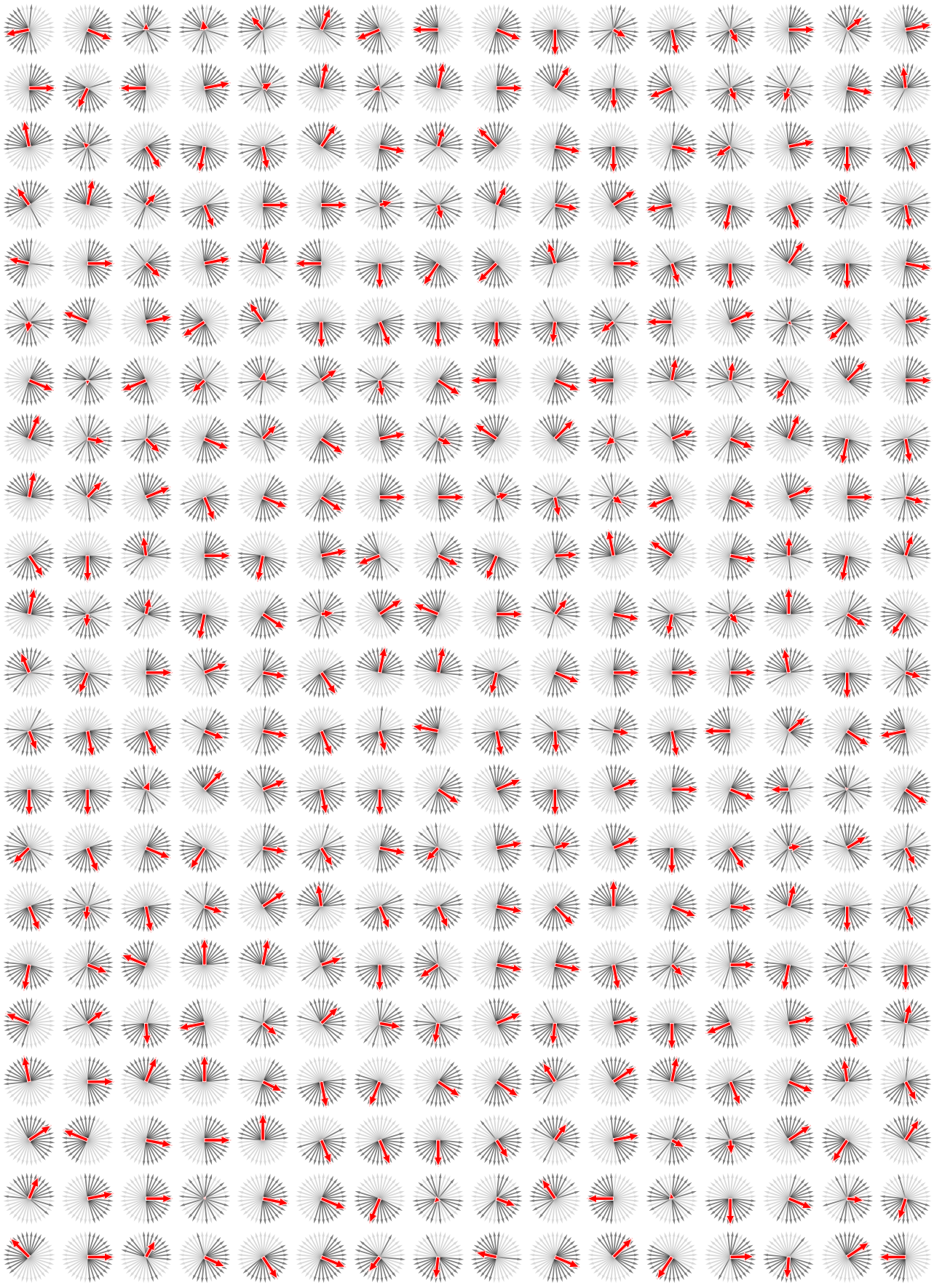
SFM



SFM



TFM



TFM

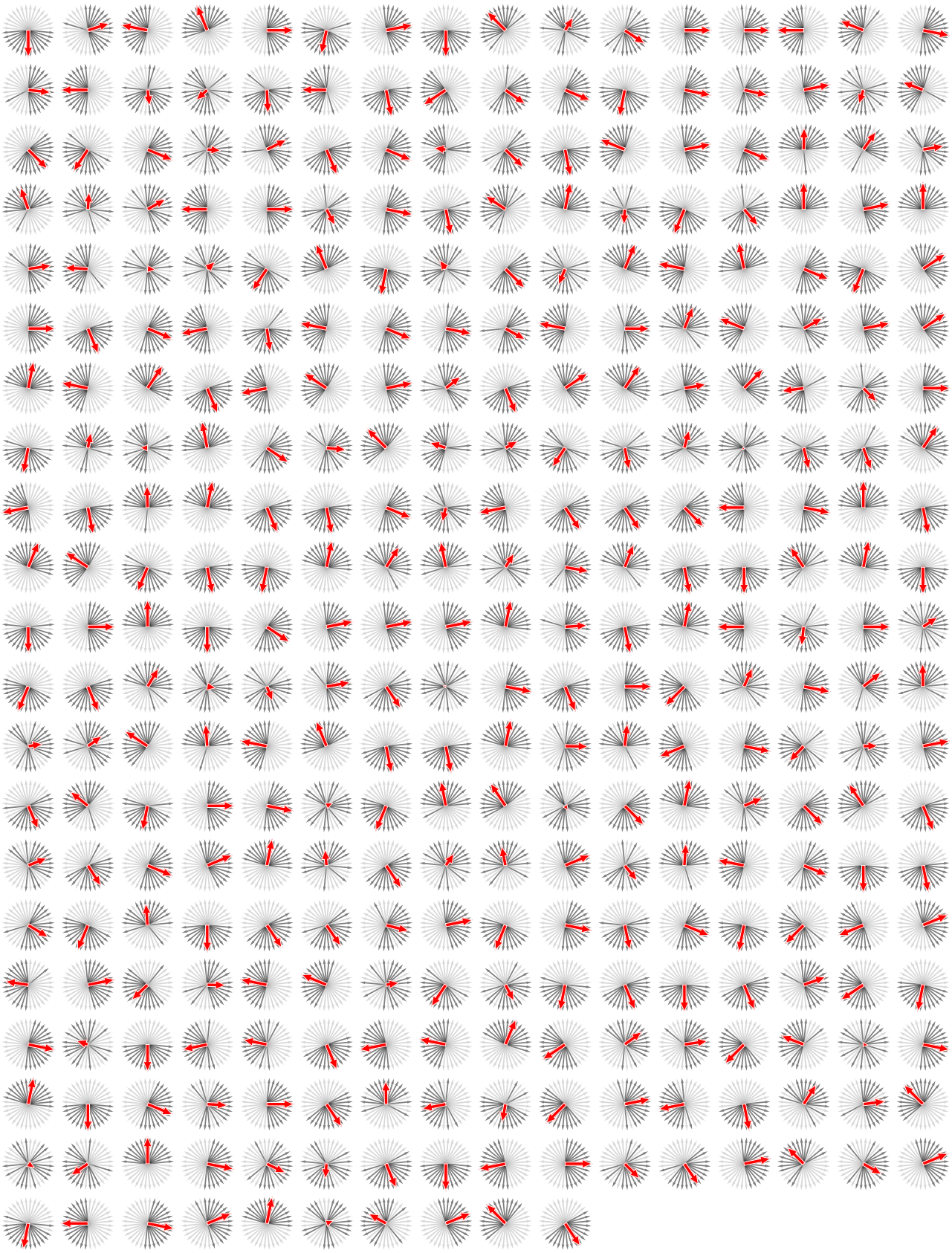


Figure S2

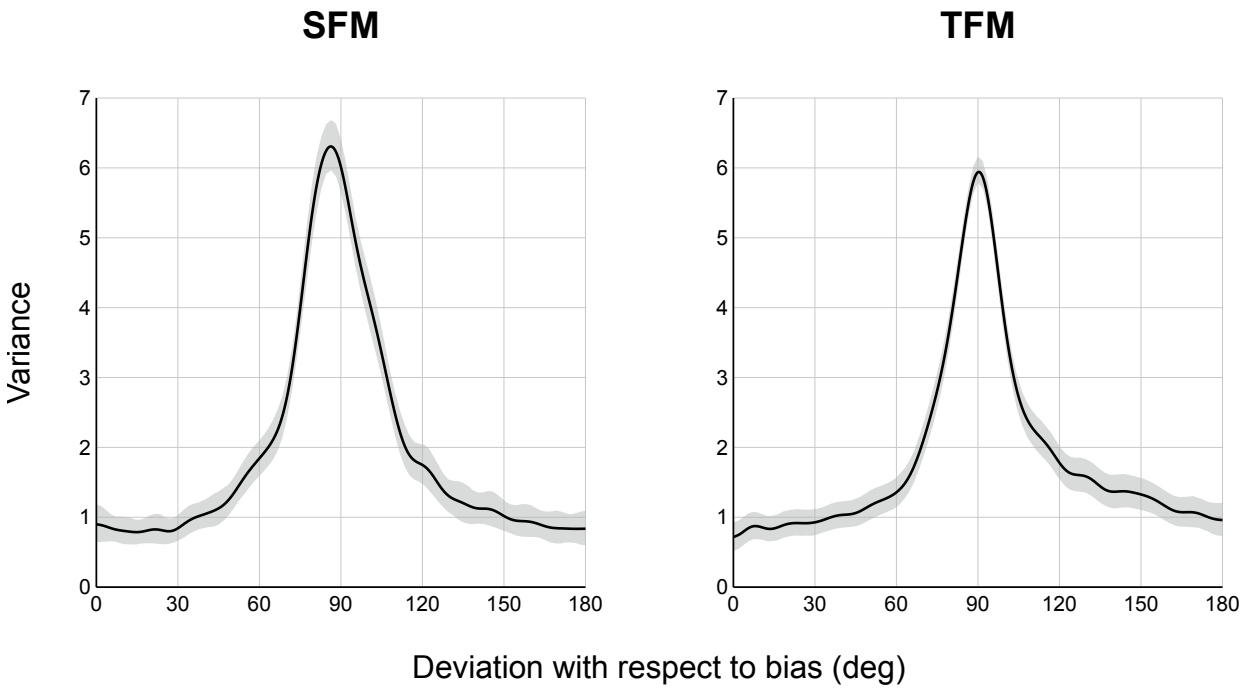

SFM
TFM
Variance
0
1
2
3
4
5
6
7
0
30
60
90
120
150
180
Deviation with respect to bias (deg)

Figure S3

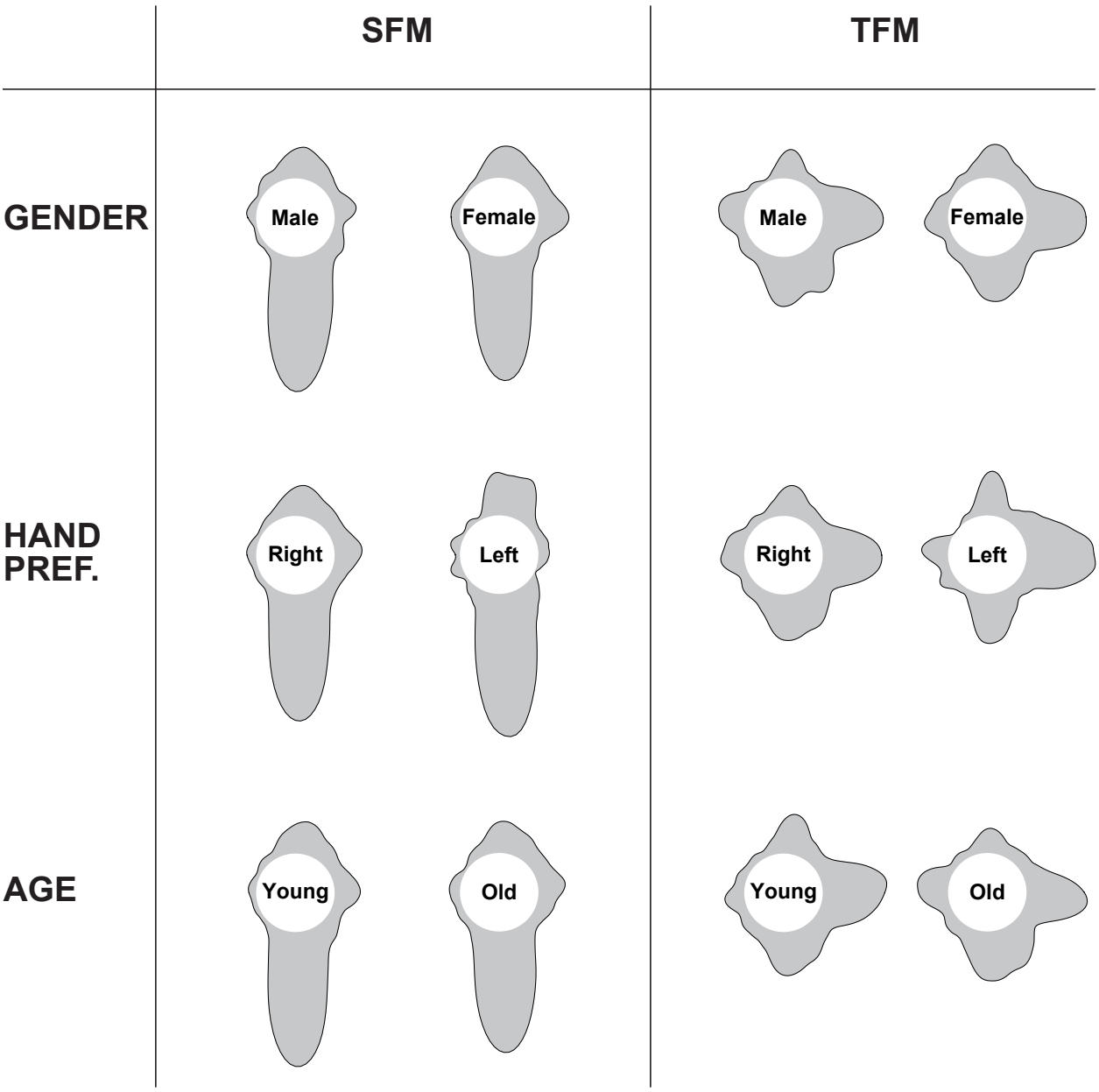

SFM
TFM
GENDER
HAND PREF.
AGE
Male
Female
Male
Female
Right
Left
Right
Left
Young
Old
Young
Old

Figure S4

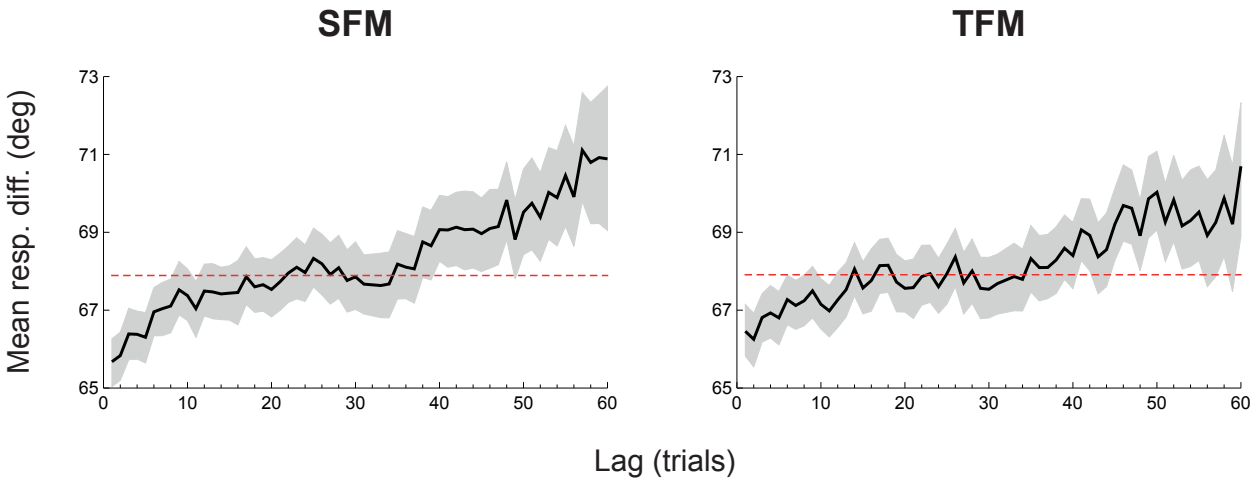

SFM
TFM
Mean resp. diff. (deg)
73
71
69
67
65
0
10
20
30
40
50
60
Lag (trials)

Figure S5

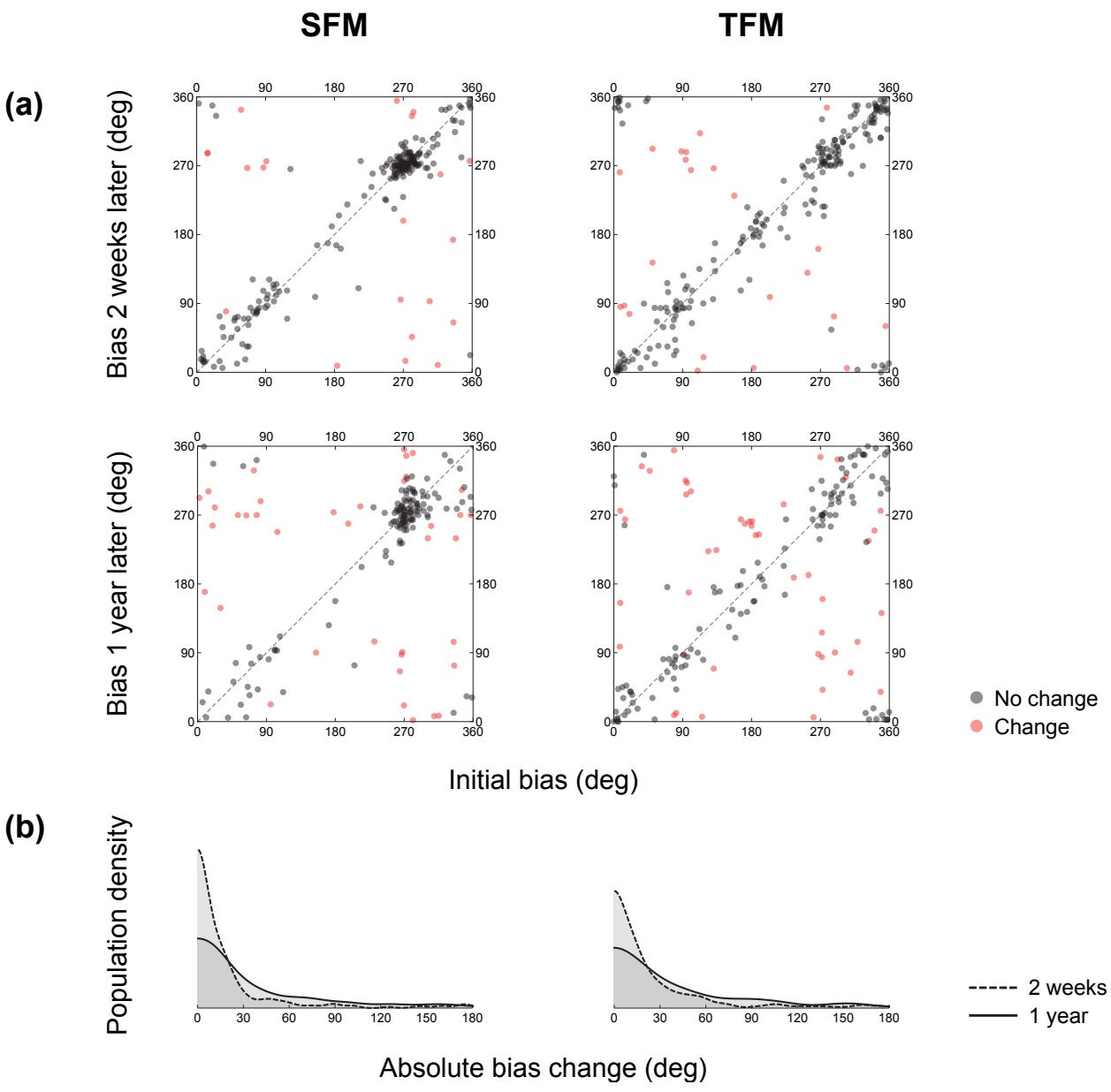

SFM
TFM
(a)
Bias 2 weeks later (deg)
Bias 1 year later (deg)
0
90
180
270
360
No change
Change
Initial bias (deg)
(b)
Population density
0
30
60
90
120
150
180
2 weeks
1 year
Absolute bias change (deg)

Figure S6

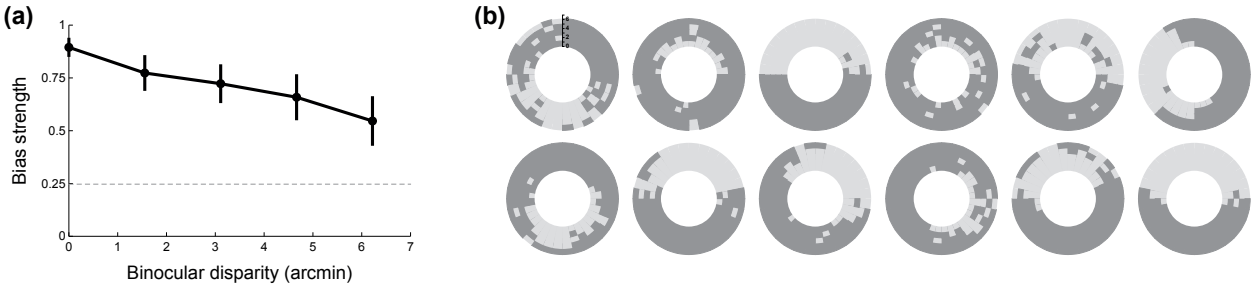

(a)
Bias strength
1
0.75
0.5
0.25
0
0
1
2
3
4
5
6
7
Binocular disparity (arcmin)
(b)

Figure S7

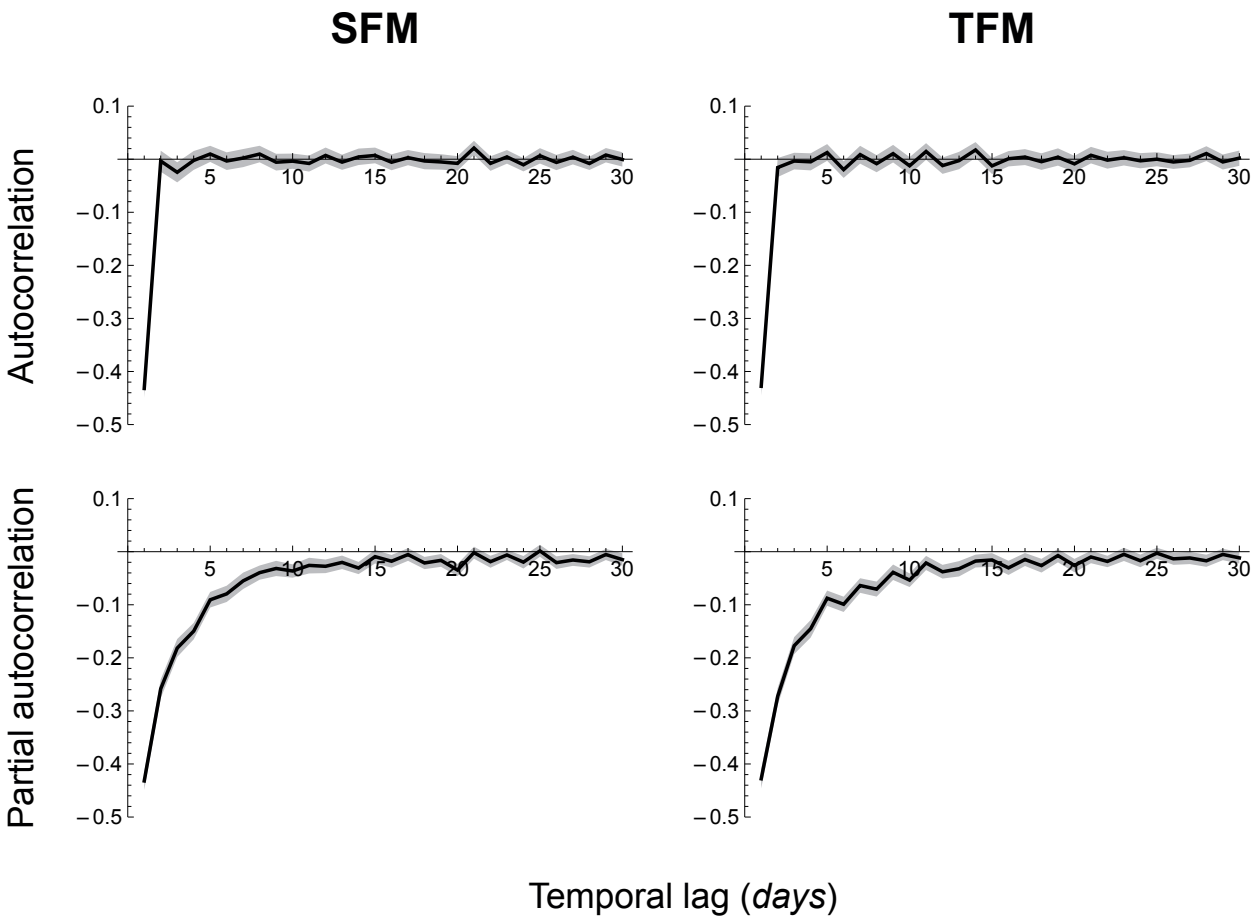

SFM
TFM
Autocorrelation
Partial autocorrelation
0.1
−0.1
−0.2
−0.3
−0.4
−0.5
5
10
15
20
25
30
Temporal lag (*days*)

Figure S8

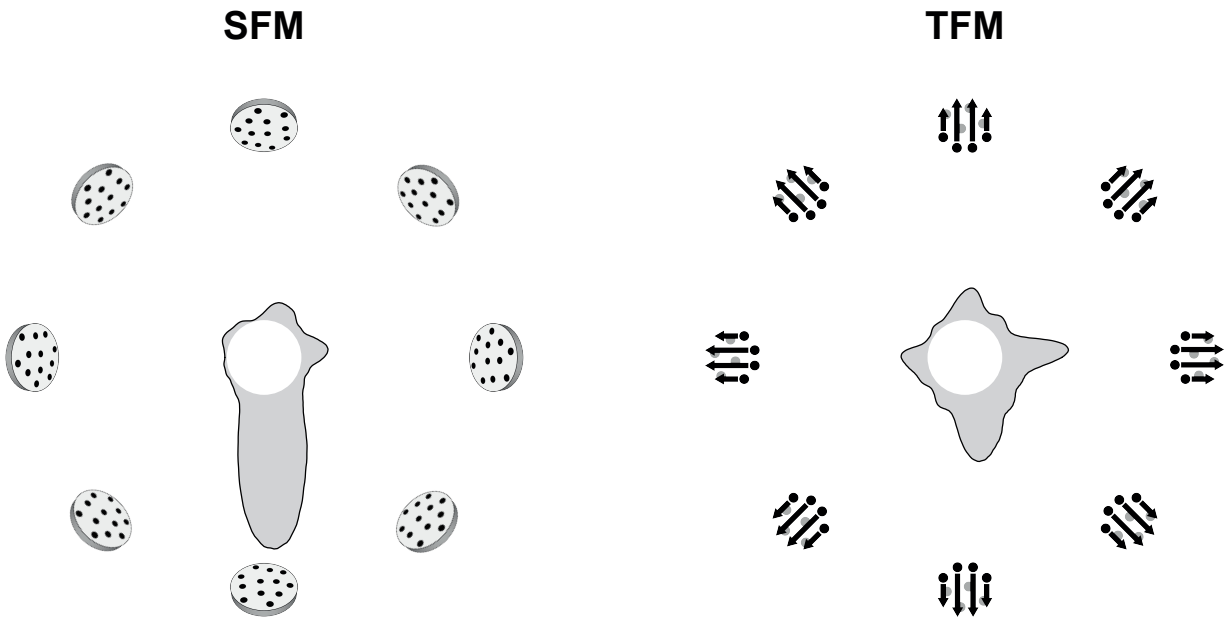


Figure S9

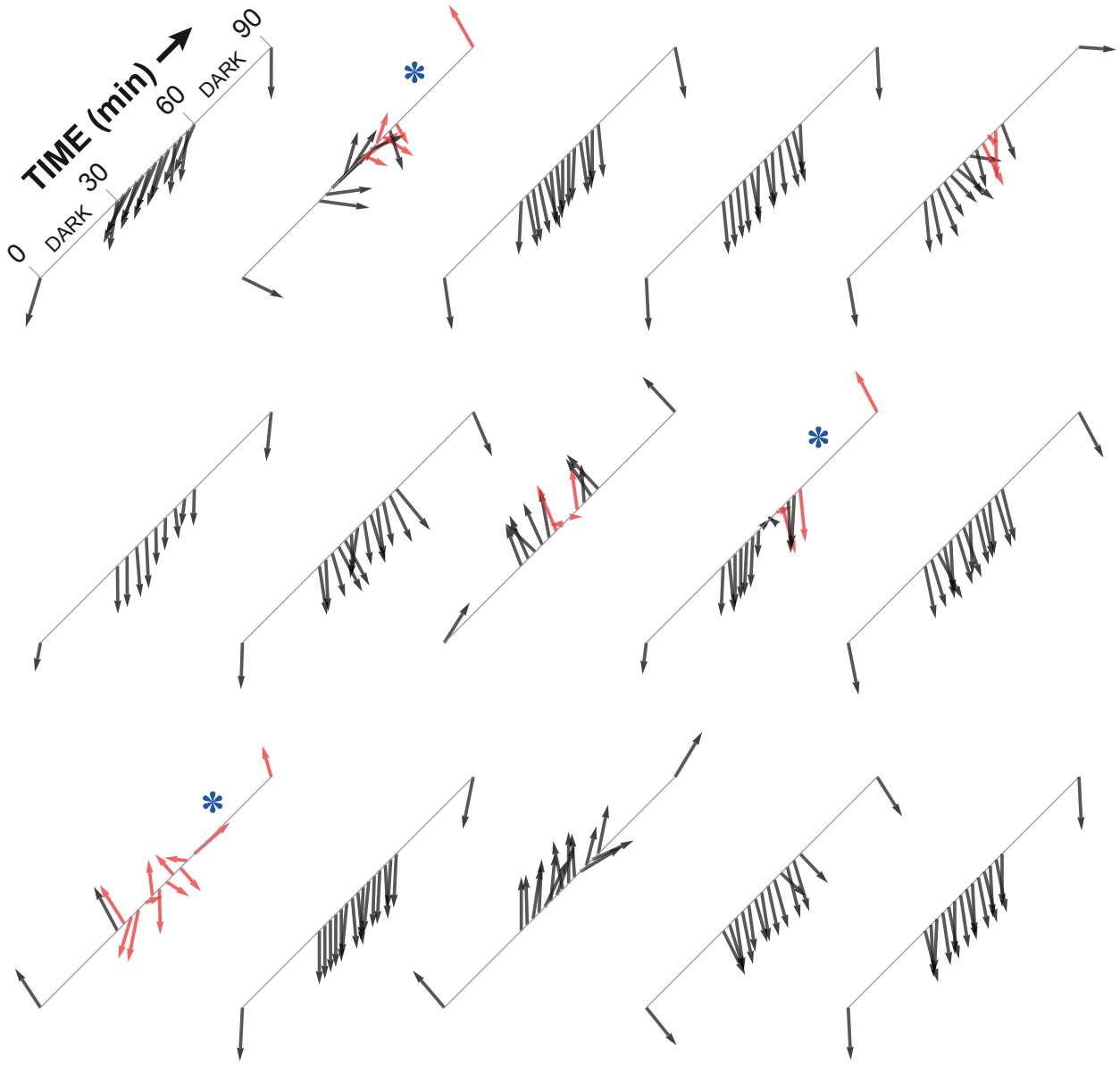


Figure S10

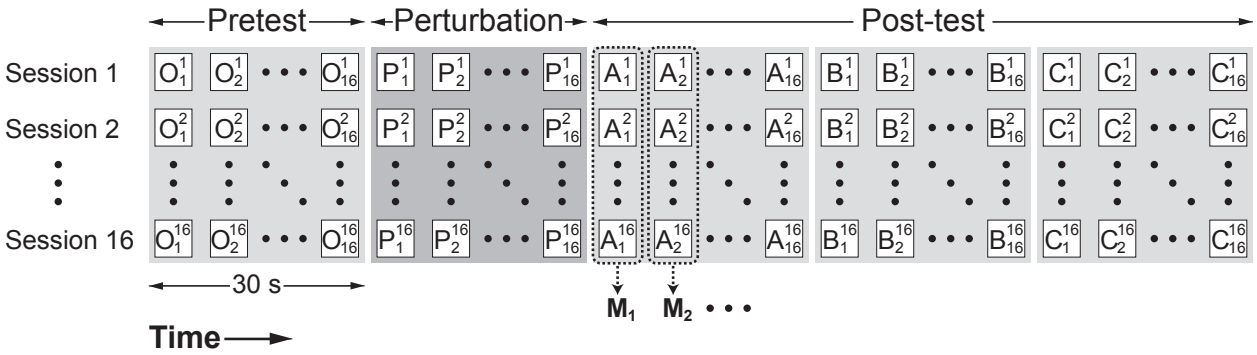

Pretest
Perturbation
Post-test
Session 1
Session 2
Session 16
30 s
Time
M1
M2

Figure S11

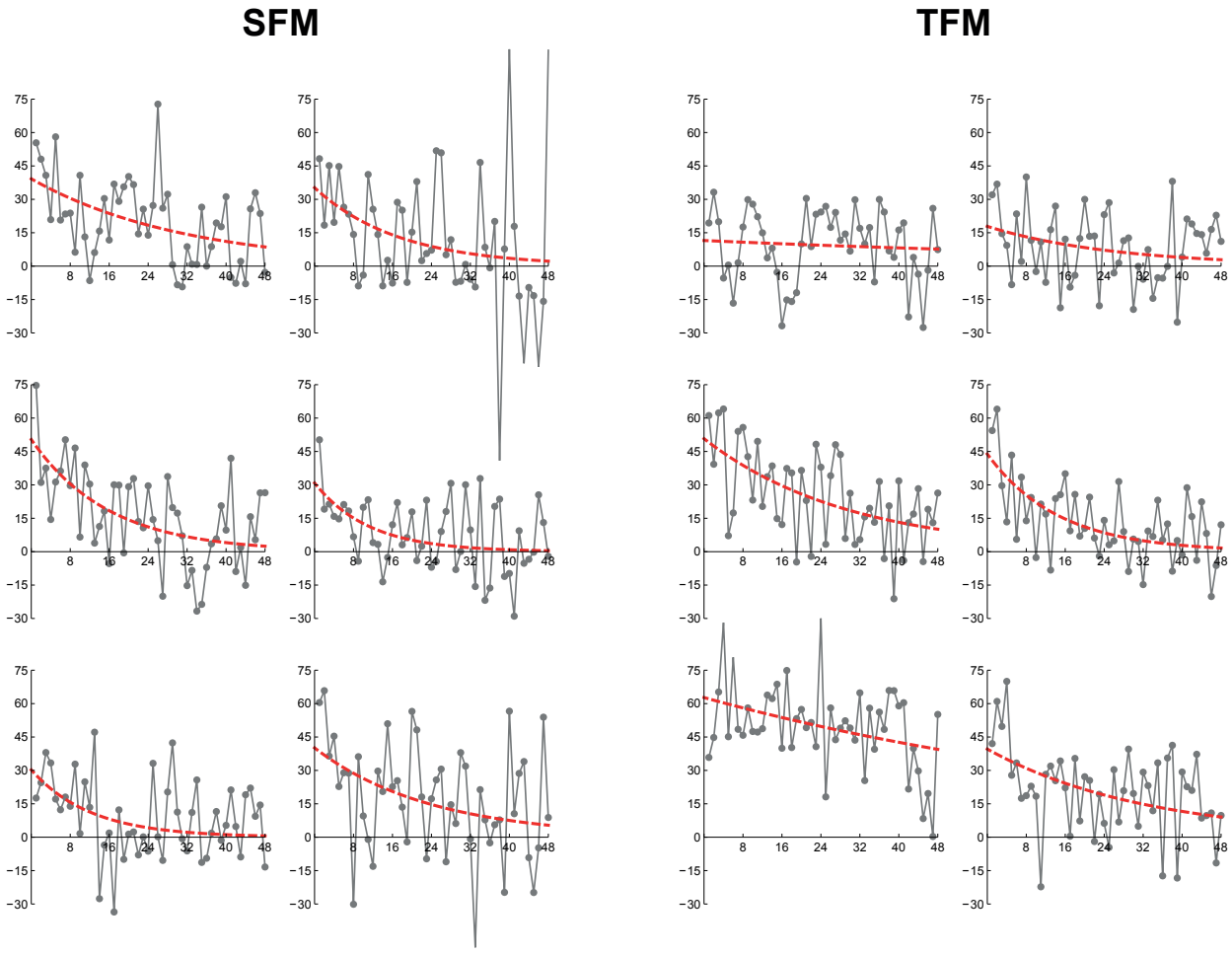


Figure S12

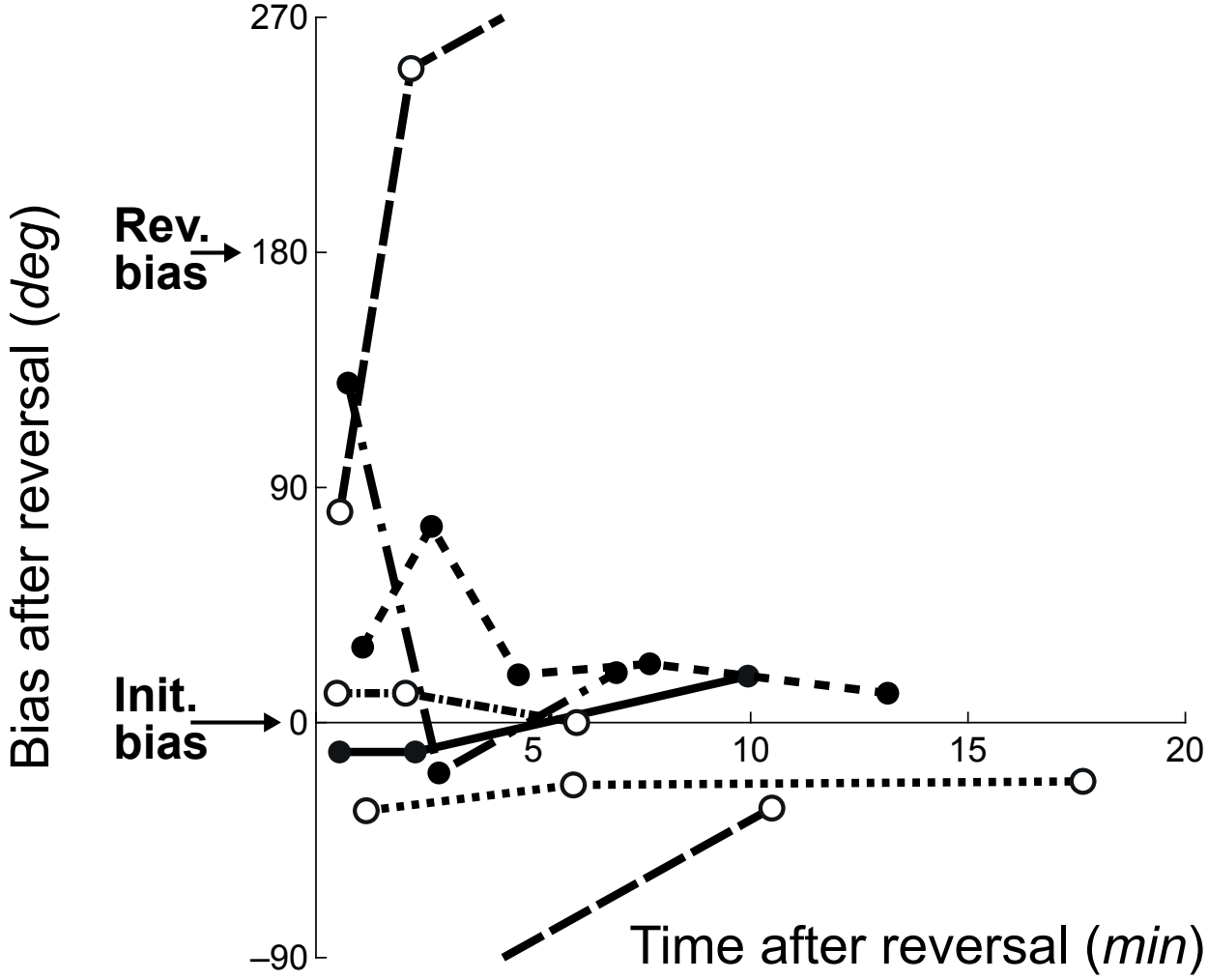

270
180
90
0
–90
Rev.
bias
Init.
bias
Bias after reversal (deg)
5
10
15
20
Time after reversal (min)

Figure S13

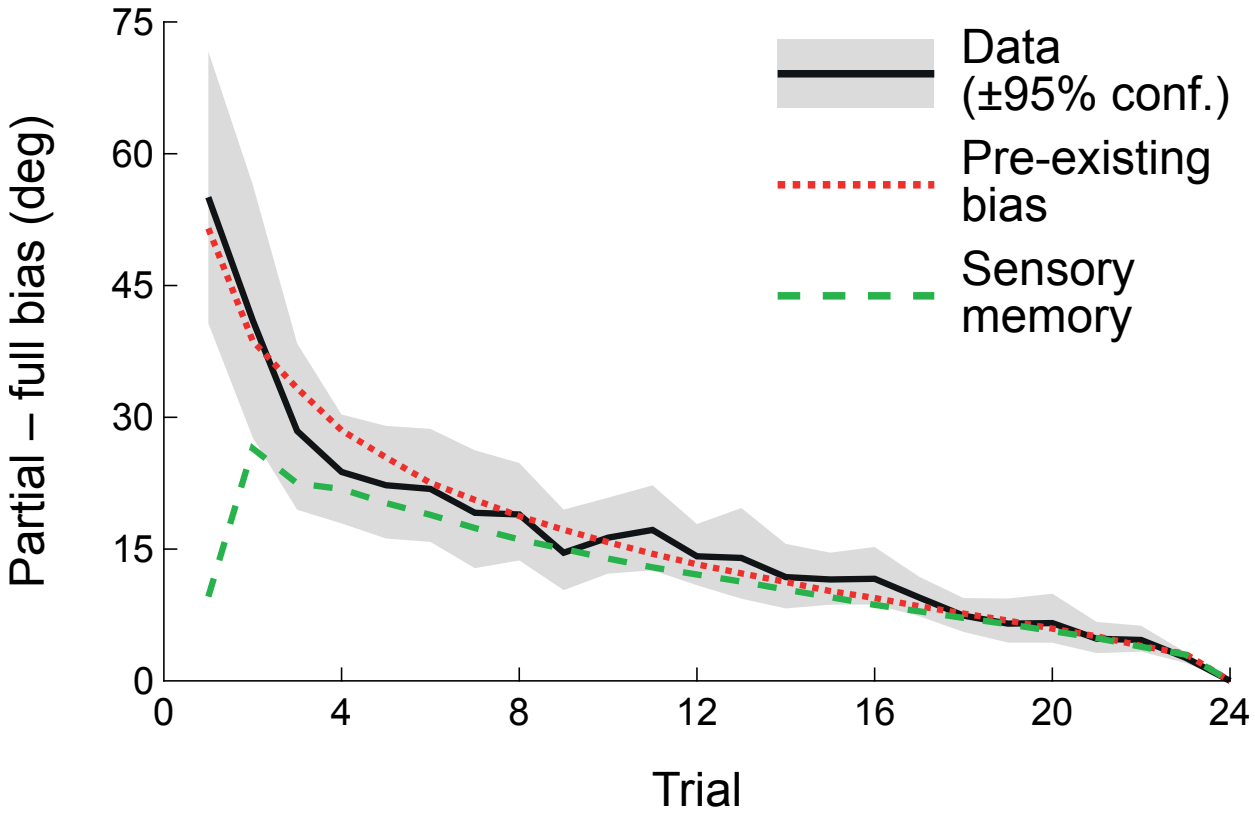
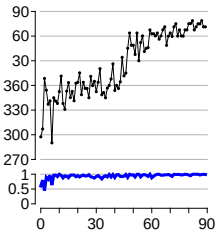
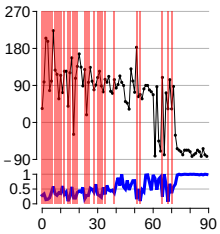
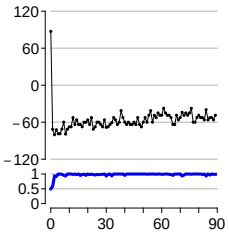
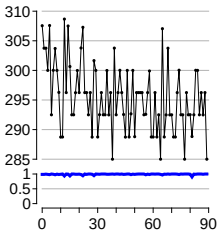
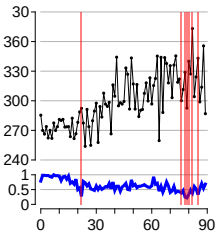
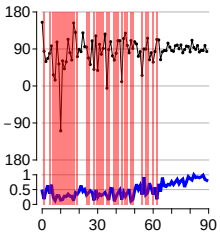
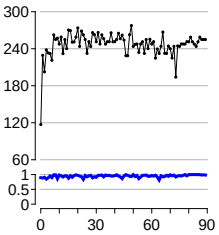
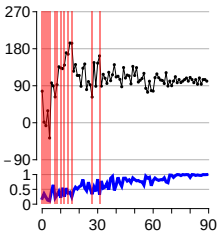
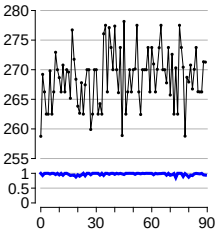
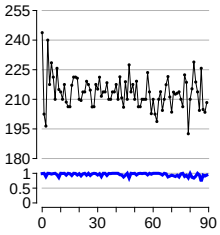
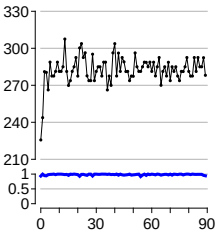
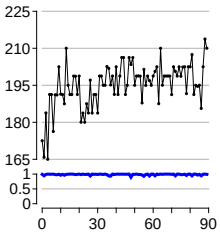
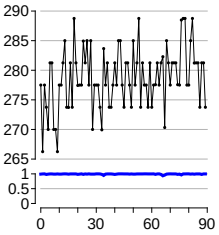
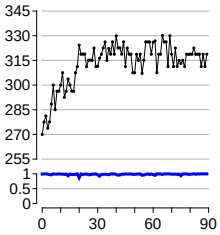
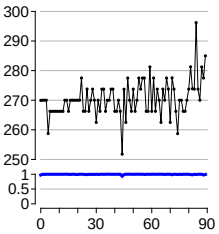
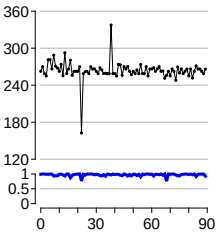
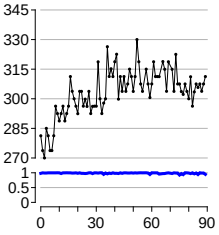
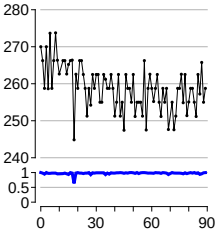
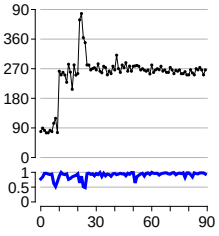
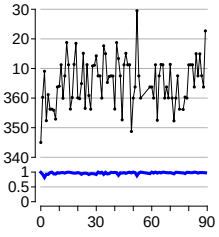
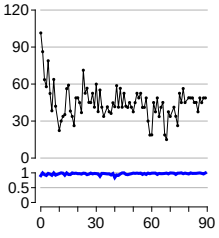

Partial – full bias (deg)
75
60
45
30
15
0
0
4
8
12
16
20
24
Trial
Data
(±95% conf.)
Pre-existing
bias
Sensory
memory

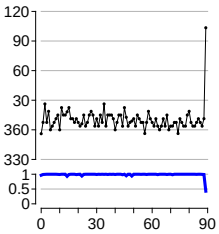
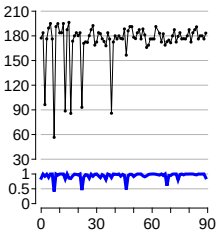
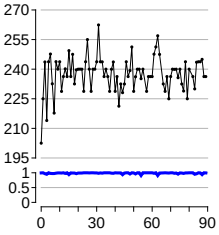
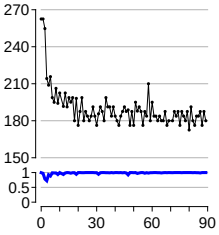
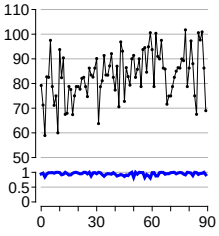
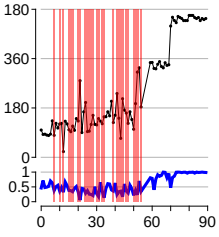
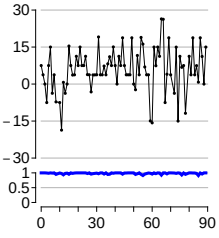
Table T2 (extends over next 14 pages)

SFM							TFM						
	p	*d*	*q*	a_1	b_1	int.		*p*	*d*	*q*	a_1	b_1	int.
	0	1	1		−0.646			0	1	1		−0.876	
	0	1	1		−0.730			1	1	0	−0.619		
	0	1	1		−0.642			0	1	1		−0.912	
	1	0	0	0.812		0.745		0	1	1		−0.908	
	1	1	1	−0.203	−0.770			0	0	0			
	1	1	1	−0.243	−0.620			0	1	1		−0.839	
	1	0	1	0.900	−0.745	0.345		1	0	1	0.819	−0.606	−0.183
	1	0	1	0.994	−0.578			0	1	1		−0.942	
	0	0	0			0.022		0	0	0			−0.791
	1	0	1	0.826	−0.669	0.970		0	1	1		−0.935	
	1	0	0	0.344		−0.206		0	1	1		−0.900	
	0	0	0			0.943		0	1	1		−0.698	
	0	1	1		−0.906			0	1	1		−0.483	
	1	0	0	0.159		0.978		0	1	1		−0.637	

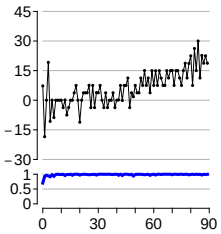
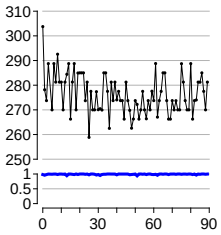
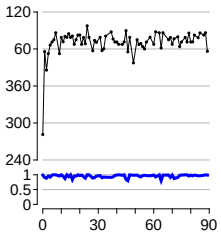
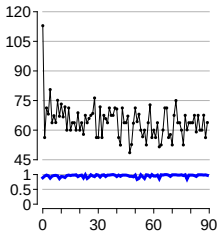
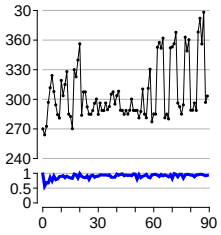
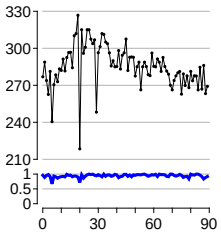
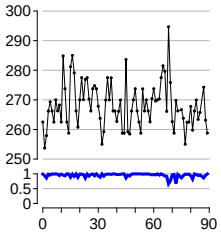
SFM

	p	d	q	a_1	b_1	int.
	0	1	1		−0.841	
	0	1	1		−0.933	
	0	0	0			0.097
	0	0	0			0.942
	0	1	1		−0.643	
	0	1	1		−0.753	
	0	1	1		−0.853	
	0	0	0			0.953
	1	0	1	0.808	−0.599	
	0	1	0			
	0	1	1		−0.944	
	0	0	1		0.175	−0.084
	0	1	1		−0.479	
	1	0	0	0.409		−0.664

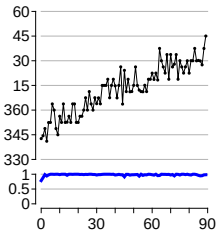
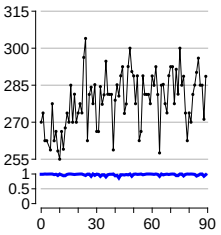
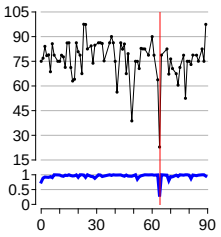
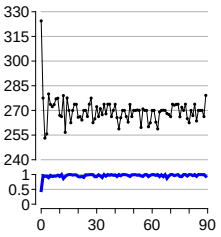
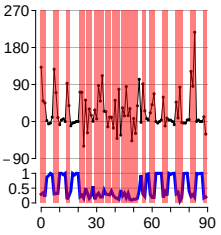
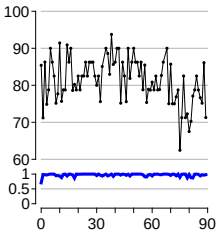
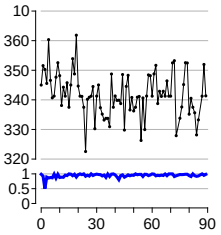
TFM

	p	d	q	a_1	b_1	int.
	0	0	0			0.963
	0	0	0			0.163
	1	1	1	−0.161	−0.933	
	0	0	0			0.041
	1	0	0	0.195		−0.525
	0	0	0			−0.823
	0	1	0			
	1	1	1	−0.404	−0.318	
	0	1	1		−0.904	
	0	0	0			0.932
	1	0	1	0.922	−0.270	
	0	1	1		−0.494	
	0	1	1		−0.946	
	1	0	0	0.156		0.095

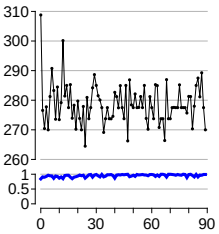
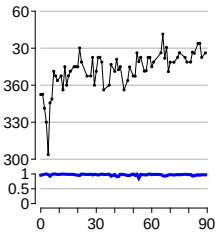
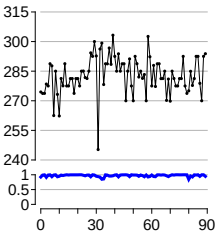
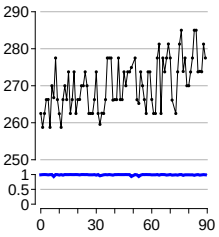
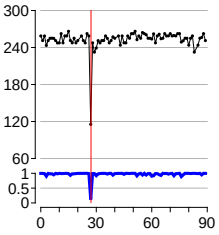
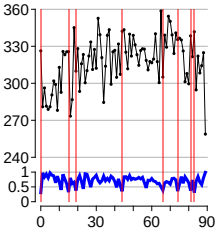
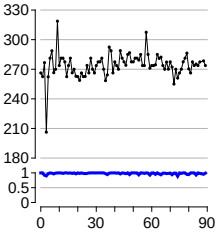
SFM

	p	*d*	*q*	a_1	b_1	int.
	1	0	1	0.966	−0.628	−0.937
	1	1	1	−0.288	−0.878	
	0	1	1		−0.888	
	0	1	1		−0.888	
	1	0	0	0.325		−0.269
	1	1	0	−0.843		
	0	1	1		−0.901	
	0	0	0			−0.853
	1	1	1	0.459	−0.955	
	1	1	1	0.445	−0.938	
	0	1	1		−0.808	
	1	1	1	0.436	−0.860	
	0	0	1		0.352	0.033
	1	0	0	0.202		0.947

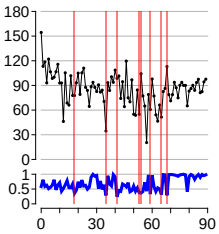
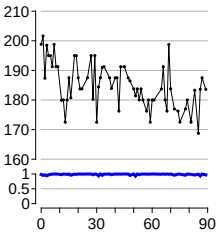
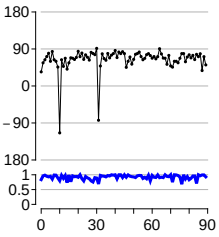
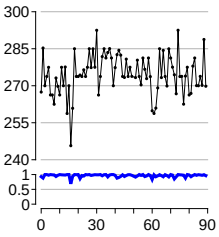
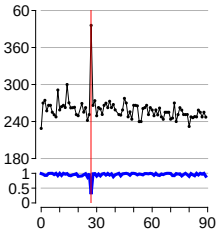
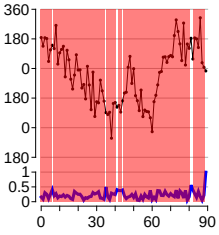
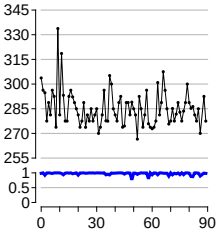
TFM

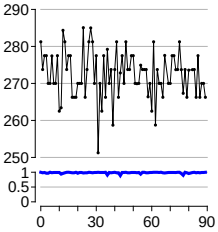
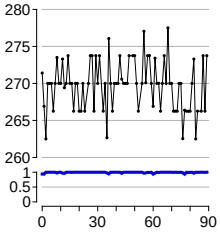
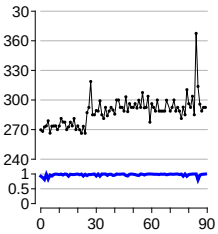
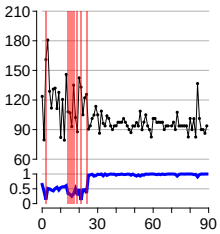
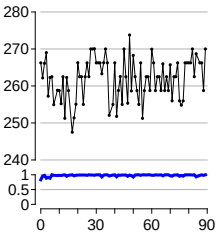
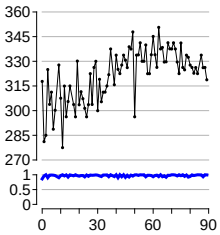
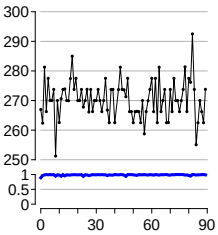
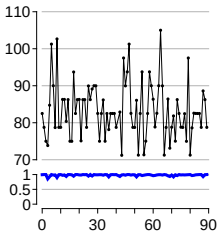
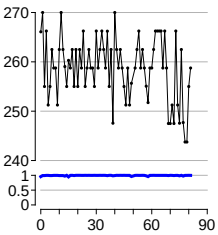
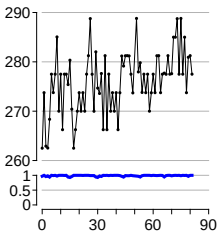
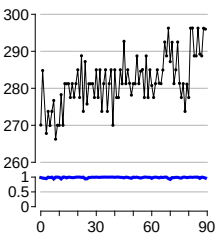
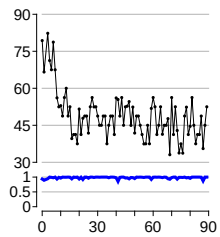
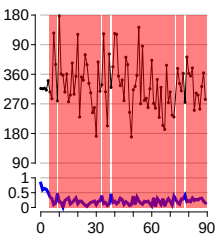
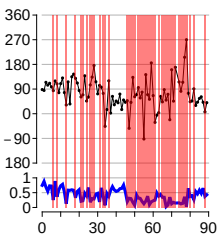
	p	*d*	*q*	a_1	b_1	int.
	1	1	1	−0.379	−0.336	
	0	1	1		−0.873	
	1	1	1	0.182	−0.884	
	1	0	0	0.210		−0.949
	1	0	0	0.402		0.208
	0	0	0			0.918
	0	0	0			
	0	1	1		−0.768	
	1	0	0	0.467		0.424
	0	0	0			0.063
	0	1	1		−0.834	
	1	0	0	0.294		0.954
	0	0	0			0.891
	1	0	0	0.327		−0.296

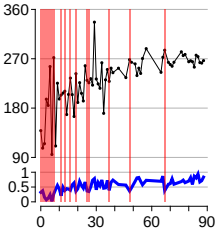
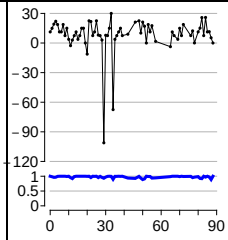
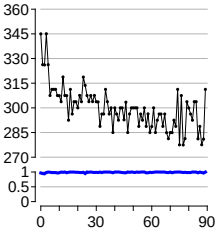
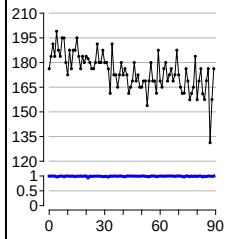
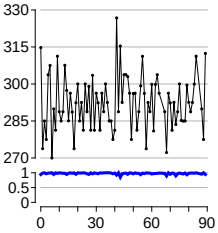
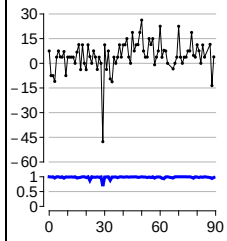
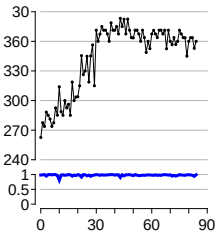
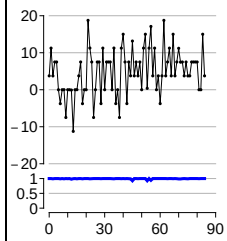
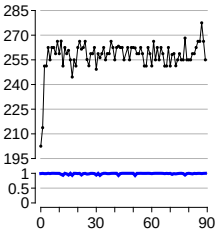
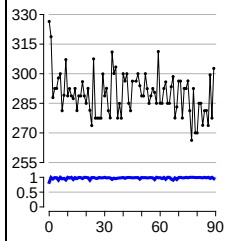
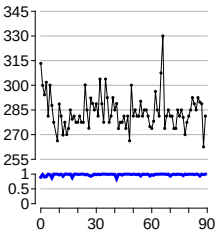
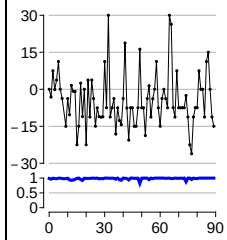
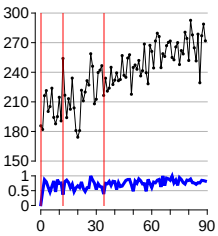
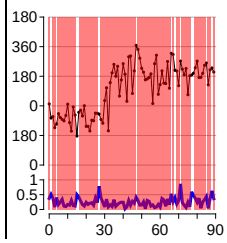
SFM

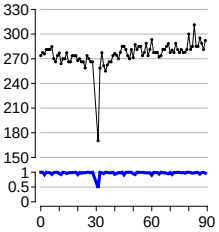
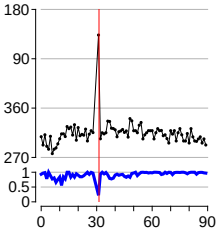
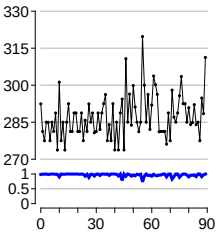
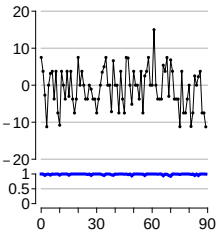
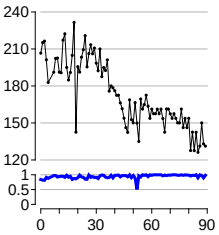
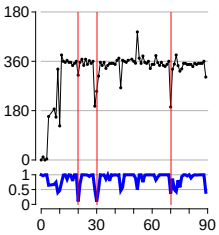
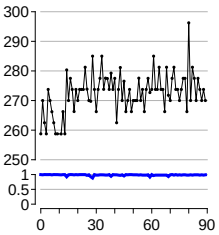
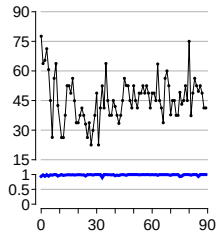
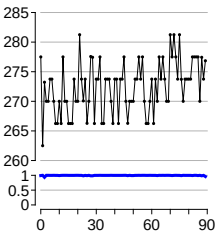
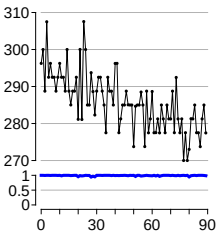
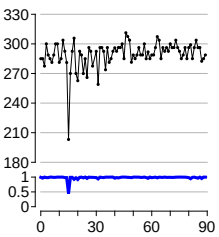
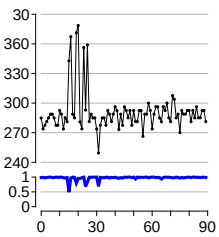
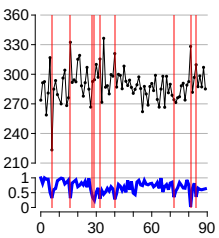
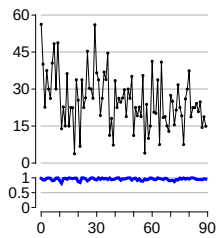
	p	d	q	a_1	b_1	int.
	0	0	0			−0.141
	0	1	1		−0.884	
	1	1	1	0.374	−0.943	
	1	1	0	−0.294		
	0	0	0			−0.214
	1	0	1	0.861	−0.631	0.937
	1	1	1	0.312	−0.916	
	0	0	1		0.190	0.984
	1	0	0	0.294		0.262
	0	0	0			0.926
	0	1	1		−0.834	
	0	1	1		−0.827	
	0	0	1		0.242	−0.074
	0	0	0			0.959

TFM

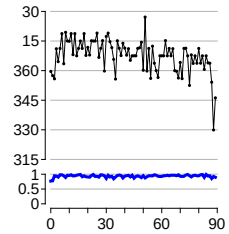
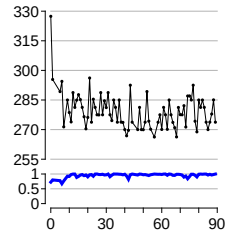
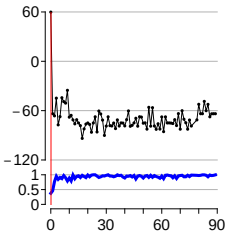
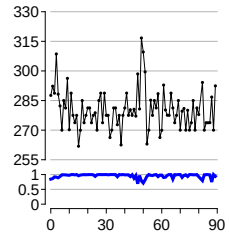
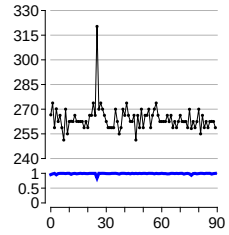
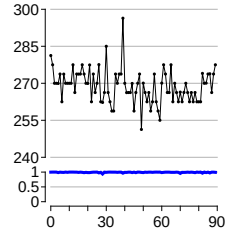
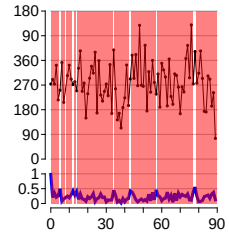
	p	d	q	a_1	b_1	int.
	0	1	1		−0.772	
	0	1	1		−0.803	
	1	1	1	−0.416	−0.425	
	0	1	1		−0.614	
	0	0	0			0.325
	0	0	0			0.807
	1	0	1	0.856	−0.726	0.078
	0	0	0			−0.952
	0	1	1		−0.915	
	0	0	0			−0.910
	1	0	0	0.219		
	0	0	0			0.051
	0	0	0			0.250
	0	0	0			−0.925

SFM							TFM						
	p	*d*	*q*	a_1	b_1	int.		*p*	*d*	*q*	a_1	b_1	int.
	0	0	0			−0.042		0	0	0			
	0	0	0			0.977		1	0	1	−0.587	0.933	−0.990
	0	1	1		−0.764			1	1	1	−0.182	−0.926	
	0	1	1		−0.909			0	1	1		−0.363	
	1	0	1	0.787	−0.574	0.136		0	1	1		−0.878	
	0	1	1		−0.826			0	1	1		−0.817	
	0	0	0					0	0	0			0.114
	0	0	0			0.981		0	0	0			0.966
	0	1	1		−0.836			0	1	1		−0.724	
	0	1	1		−0.833			1	0	0	0.238		−0.975
	1	1	1	−0.373	−0.478			0	1	1		−0.547	
	0	1	1		−0.771			0	1	1		−0.728	
	1	0	0	0.273		−0.098		0	1	1		−0.710	
	1	0	1	0.938	−0.806	0.125		0	1	1		−0.806	

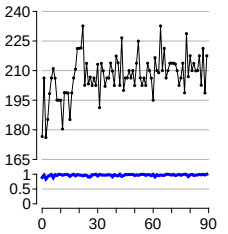
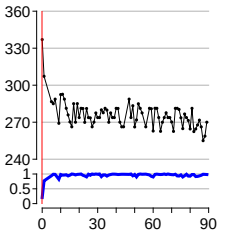
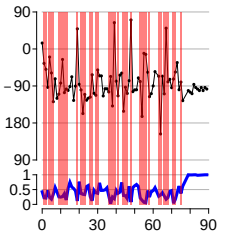
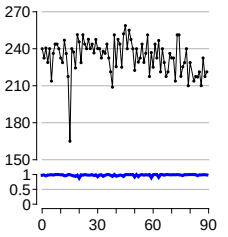
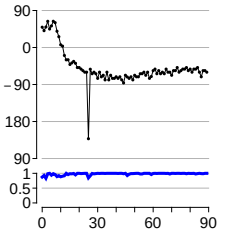
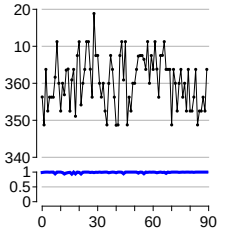
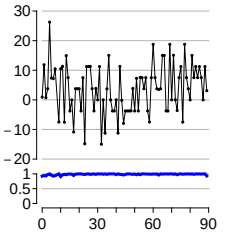
SFM							TFM						
	p	*d*	*q*	a_1	b_1	int.		*p*	*d*	*q*	a_1	b_1	int.
	0	1	1		−0.784			0	0	0			0.931
	1	1	1	−0.227	−0.815			0	0	0			0.148
	0	1	1		−0.887			0	1	1		−0.906	
	0	1	1		−0.586			0	1	1		−0.846	
	0	0	0			−0.365		0	0	0			0.967
	0	0	0			0.887		0	1	1		−0.847	
	1	1	1	−0.278	−0.423			0	0	0			0.986
	1	1	1	−0.300	−0.422			0	1	1		−0.923	
	1	0	0	0.570		0.216		0	1	1		−0.870	
	1	0	0	0.920		0.898		0	1	1		−0.794	
	0	0	1		0.234	−0.233		0	0	0			0.963
	0	0	1		0.203	0.935		0	0	1		0.170	−0.076
	0	1	1		−0.834			1	1	1	0.249	−0.930	
	0	1	1		−0.903			0	0	0			−0.108

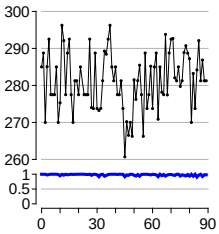
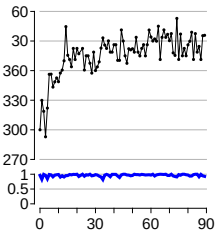
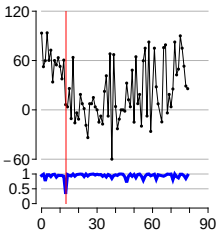
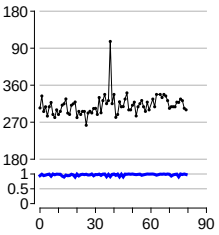
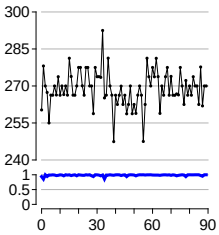
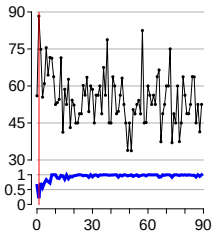
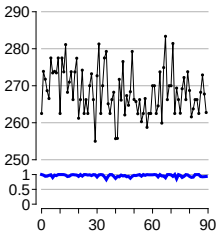
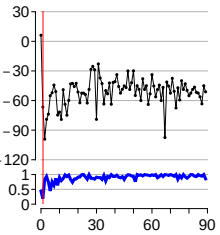
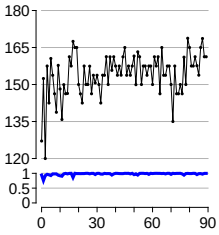
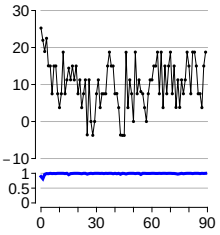
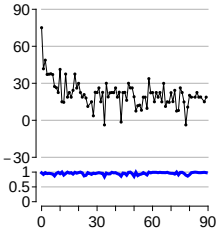
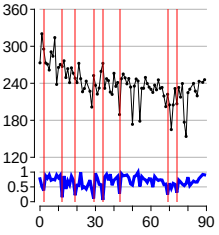
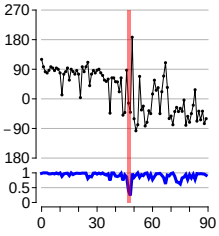
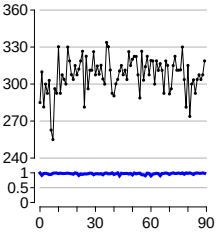
SFM							TFM						
	p	*d*	*q*	a_1	b_1	int.		*p*	*d*	*q*	a_1	b_1	int.
	0	1	1		−0.678			1	1	1	0.396	−0.907	
	1	0	0	0.477		0.936		1	1	1	0.235	−0.876	
	0	1	1		−0.940			0	0	0			0.985
	0	1	1		−0.923			0	0	0			
	1	0	0	0.444		0.840		1	0	0	0.545		0.723
	0	1	1		−0.706			0	1	1		−0.834	
	1	1	1	−0.200	−0.731			1	0	0	0.457		0.674
	0	0	0			0.977		1	0	0	0.453		0.695
	0	1	1		−0.907			0	1	1		−0.855	
	0	0	0			0.992		0	1	1		−0.866	
	0	1	1		−0.898			0	0	0			0.307
	0	0	0			0.904		0	1	1		−0.756	
	0	0	0			−0.205		0	1	1		−0.904	
	1	0	0	0.235		0.642		0	1	1		−0.892	

SFM

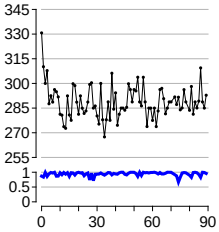
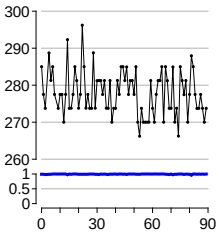
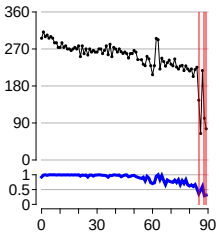
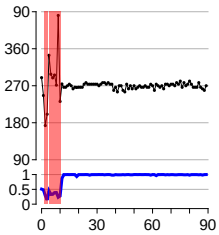
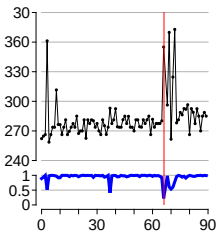
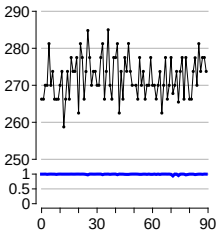
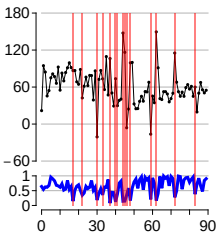
	p	d	q	a_1	b_1	int.
	1	0	0	0.253		−0.918
	0	1	1		−0.717	
	0	1	1		−0.821	
	0	1	1		−0.351	
	1	0	1	0.832	−0.589	−0.302
	1	1	0	−0.320		
	0	0	0			−0.161
	1	0	1	0.723	−0.395	0.927
	0	0	0			0.103
	0	0	0			0.977
	0	1	1		−0.901	
	0	0	0			0.985
	0	0	0			
	1	1	1	−0.230	−0.855	

TFM

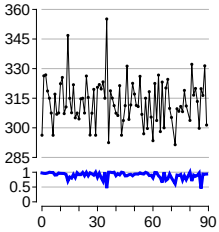
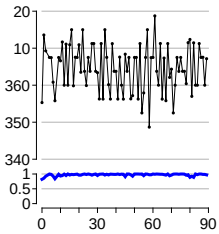
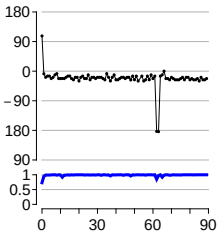
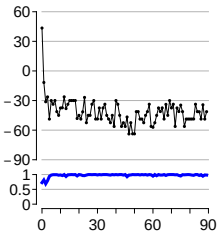
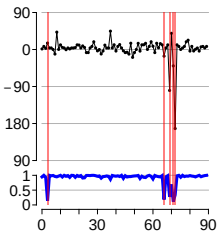
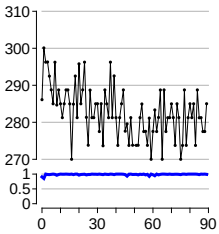
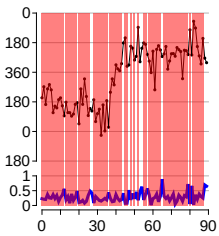
	p	d	q	a_1	b_1	int.
	0	0	0			−0.845
	1	1	1	−0.365	−0.458	
	0	1	1		−0.868	
	1	1	1	0.957	−0.742	
	1	0	0	0.213		
	0	1	1		−0.625	
	0	1	1		−0.864	
	0	0	0			−0.766
	0	1	1		−0.690	
	0	1	1		−0.401	
	0	1	1		−0.935	
	1	0	0	0.249		
	0	1	1		−0.820	
	0	0	0			0.062

SFM							TFM						
	p	*d*	*q*	a_1	b_1	int.		*p*	*d*	*q*	a_1	b_1	int.
	0	0	1		0.233	−0.189		1	0	1	0.802	−0.545	0.857
	0	1	1		−0.930			0	1	1		−0.645	
	1	0	1	0.898	−0.696	−0.662		0	1	1		−0.891	
	1	0	1	0.949	−0.786	−0.421		1	0	1	0.726	−0.530	−0.706
	1	0	0	0.161				0	1	1		−0.781	
	0	0	0			0.977		1	0	0	0.201		0.758
	0	1	1		−0.922			0	1	1		−0.821	
	0	0	1		0.187	0.958		1	1	1	0.250	−0.950	
	0	1	1		−0.692			1	0	1	0.933	−0.396	0.950
	0	1	1		−0.831			1	0	1	0.747	−0.464	0.180
	0	1	1		−0.600			0	1	1		−0.847	
	0	1	1		−0.752			1	0	1	0.856	−0.726	−0.543
	0	1	1		−0.836			1	0	1	0.808	−0.624	0.567
	0	1	1		−0.745			0	0	0			−0.748

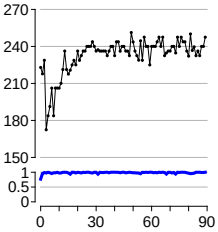
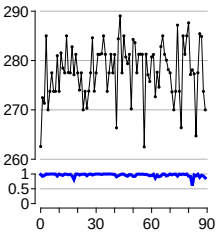
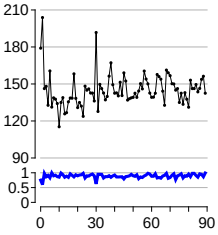
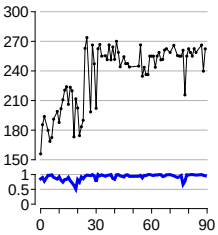
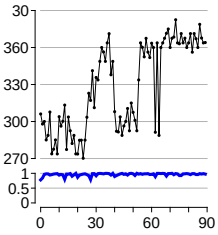
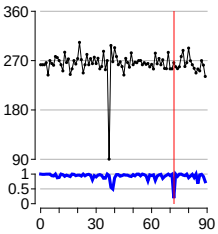
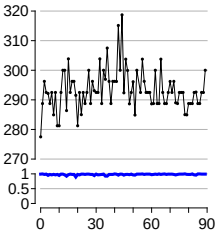
SFM

	p	d	q	a_1	b_1	int.
	1	0	0	0.289		−0.300
	1	0	0	0.319		0.878
	0	0	1		0.188	−0.135
	0	0	0			0.979
	0	1	1		−0.651	
	0	1	1		−0.576	
	1	0	0	0.246		
	1	1	1	−0.690	0.909	
	0	1	1		−0.901	
	1	0	1	0.824	−0.582	0.890
	0	0	0			−0.036
	0	0	0			0.990
	0	1	1		−0.882	
	1	0	1	0.892	−0.739	−0.508

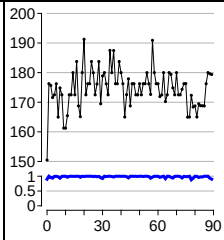
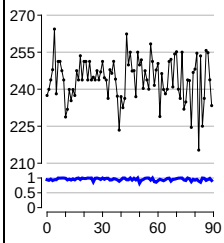
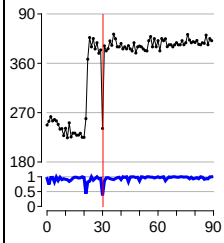
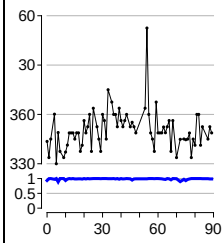
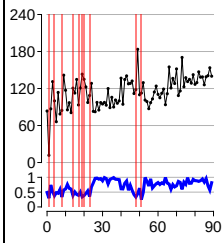
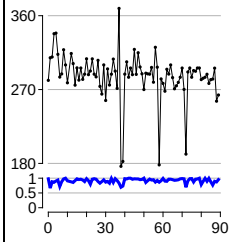
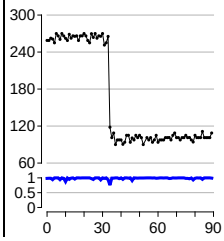
TFM

	p	d	q	a_1	b_1	int.
	0	1	1		−0.949	
	1	0	0	−0.196		−0.650
	1	1	1	0.265	−0.872	
	0	0	1		−0.260	0.079
	0	0	1		0.761	0.863
	0	0	1		0.437	−0.320
	1	0	1	0.787	−0.561	0.707
	0	1	1		−0.415	
	1	0	1	0.754	−0.507	0.896
	0	0	0			0.078
	0	1	1		−0.865	
	1	1	1	−0.380	−0.490	
	0	0	0			−0.073
	0	1	1		−0.929	

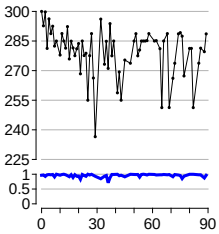
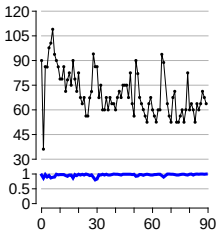
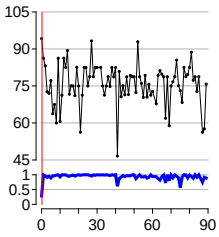
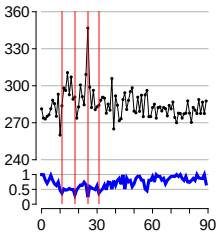
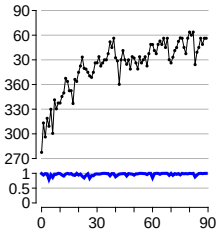
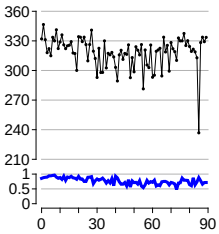
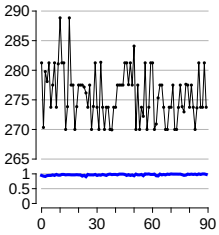
SFM

	p	d	q	a_1	b_1	int.
	0	1	1		−0.596	
	0	1	1		−0.554	
	0	0	0			−0.132
	0	1	1		−0.911	
	0	1	1		−0.897	
	1	0	0	0.235		−0.506
	0	1	1		−0.735	
	1	1	1	0.326	−0.811	
	1	1	0	−0.521		
	1	1	0	−0.521		
	0	0	0			0.064
	0	0	0			0.887
	1	0	1	0.888	−0.743	−0.381
	1	0	1	0.887	−0.724	0.896

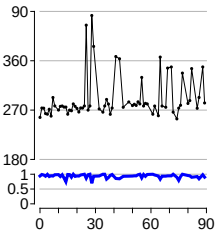
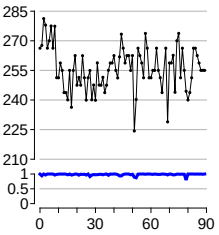
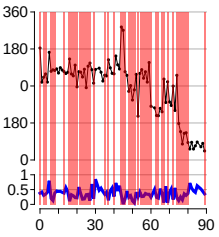
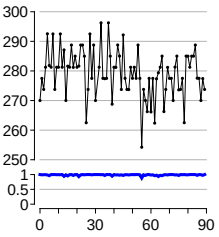
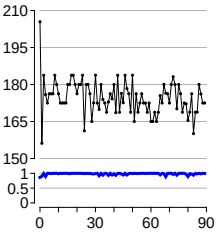
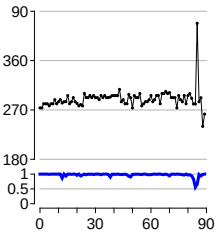
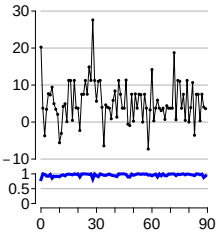
TFM

	p	d	q	a_1	b_1	int.
	1	0	1	−0.582	0.822	−0.971
	1	0	1	0.841	−0.650	0.095
	0	0	0			−0.412
	0	0	0			−0.851
	0	1	1		−0.309	
	1	1	0	−0.397		
	0	0	0			0.955
	1	0	1	0.873	−0.666	−0.166
	0	1	1		−0.830	
	1	0	1	0.834	−0.461	0.579
	0	1	1		−0.935	
	1	0	0	0.319		−0.790
	0	1	1		−0.908	
	0	1	1		0.181	

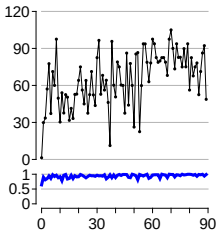
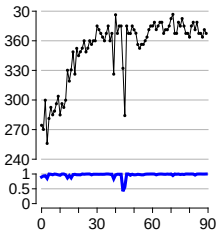
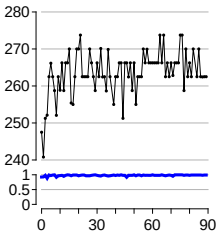
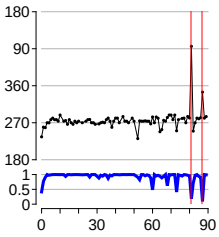
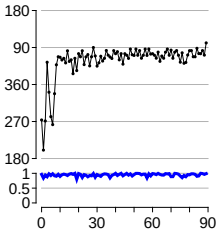
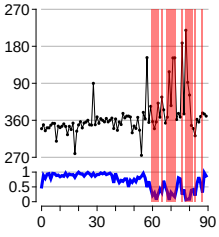
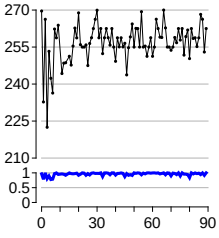
SFM

	p	*d*	*q*	a_1	b_1	int.
	1	0	0	0.438		−0.147
	1	0	0	0.467		0.924
	1	1	1	0.412	−0.910	
	0	0	0			−0.883
	1	0	0	0.166		−0.223
	0	0	0			−0.901
	0	0	0			−0.164
	0	1	1		−0.636	
	1	1	1	−0.917	0.569	
	0	1	1		−0.453	
	0	1	1		−0.858	
	1	0	0	0.216		0.484
	0	1	1		−0.918	
	0	1	1		−0.834	

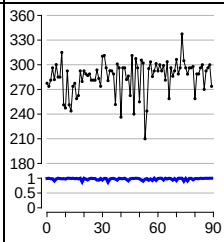
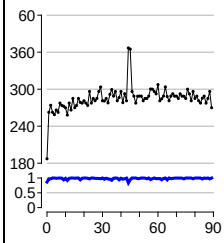
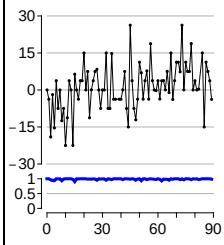
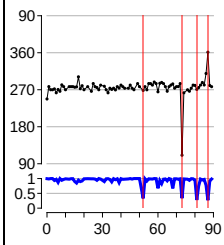
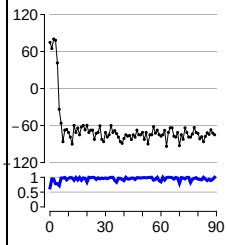
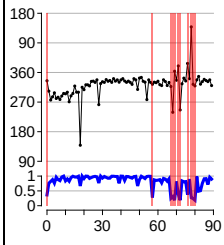
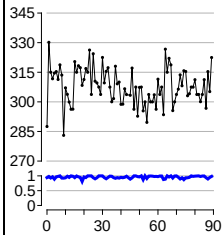
TFM

	p	*d*	*q*	a_1	b_1	int.
	1	1	1	0.438	−0.949	
	0	0	1		0.409	−0.742
	1	0	1	0.779	−0.498	−0.227
	0	0	1		0.226	−0.937
	0	0	0			0.064
	1	1	1	−0.291	−0.439	
	0	1	1		−0.916	
	1	0	1	−0.983	0.923	−0.966
	1	0	1	0.816	−0.575	−0.968
	0	1	1		−0.930	
	1	0	1	0.910	−0.613	0.260
	0	1	1		−0.908	
	0	1	1		−0.937	
	0	0	0			0.090

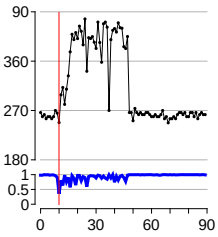
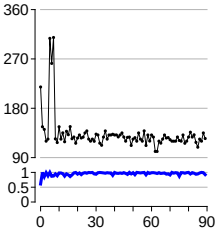
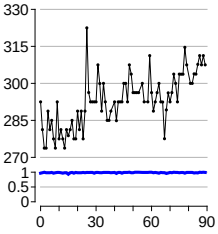
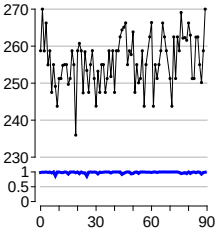
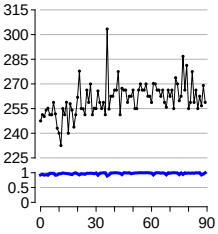
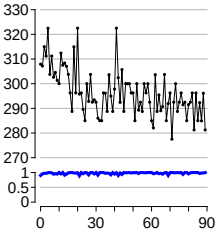
SFM

	p	*d*	*q*	a_1	b_1	int.
	0	1	1		−0.865	
	0	1	1		−0.836	
	0	1	1		−0.539	
	0	1	1		−0.568	
	0	1	1		−0.768	
	0	1	1		−0.702	
	0	0	1		0.410	−0.058
	0	1	1		−0.922	
	0	0	1		0.174	−0.311
	0	1	1		−0.465	
	0	1	1		−0.795	
	0	1	1		−0.845	
	0	1	1		−0.930	
	1	1	1	−0.403	−0.631	

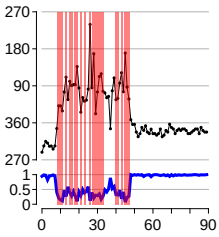
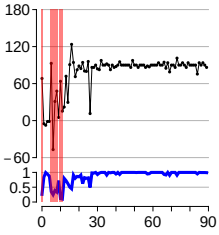
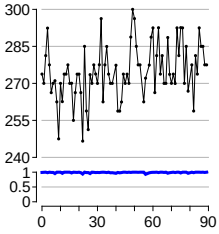
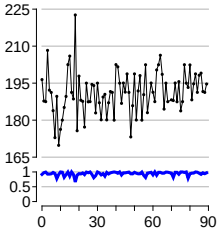
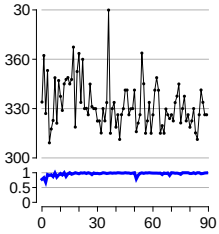
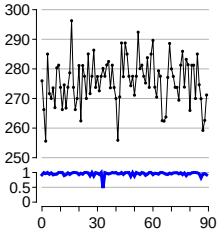
TFM

	p	*d*	*q*	a_1	b_1	int.
	0	1	1		−0.943	
	1	0	0	0.160		−0.884
	0	1	1		−0.747	
	0	0	1		0.656	−0.902
	0	0	0			0.973
	0	1	1		−0.903	
	1	0	0	0.270		0.098
	0	1	1		−0.937	
	1	0	0	0.153		0.269
	1	1	0	0.211		
	0	1	1		−0.799	
	0	1	1		−0.782	
	0	0	0			0.581
	1	0	1	0.684	−0.478	−0.746

SFM

	p	d	q	a_1	b_1	int.
	0	1	1		−0.483	
	0	1	1		−0.434	
	1	0	1	0.592	−0.311	0.529
	0	1	0			
	0	1	1		−0.680	
	0	1	1		−0.758	
	1	1	1	0.191	−0.941	
	0	1	1		−0.946	
	0	1	1		−0.874	
	0	1	1		−0.848	
	0	1	1		−0.834	
	0	1	1		−0.810	

TFM

	p	d	q	a_1	b_1	int.
	0	1	1		−0.326	
	0	1	1		−0.440	
	0	1	1		−0.780	
	0	1	1		−0.599	
	1	1	1	0.233	−0.940	
	0	0	0			−0.969
	0	0	0			−0.917
	0	1	1		−0.948	
	0	0	0			0.828
	0	1	1		−0.919	
	0	0	1		0.214	0.096
	0	0	1		−0.178	−0.945