JNeurosci
THE JOURNAL OF NEUROSCIENCE

Research Articles: Behavioral/Cognitive

Adaptive History Biases Result from Confidence-weighted Accumulation of Past Choices

Anke Braun[1], Anne E. Urai[1,2] and Tobias H. Donner[1,2,3]

[1]Department of Neurophysiology and Pathophysiology, University Medical Center Hamburg-Eppendorf, 20246 Hamburg, Germany.
[2]Department of Psychology
[3]Amsterdam Brain and Cognition (ABC), University of Amsterdam, 1018 WS Amsterdam, The Netherlands.

Correspondence: a.braun@uke.de, t.donner@uke.de

1 **Adaptive History Biases Result from Confidence-weighted**

2 **Accumulation of Past Choices**

3 Anke Braun[1, *], Anne E. Urai[1,2], Tobias H. Donner[1, 2, 3, *]

4 1 Department of Neurophysiology and Pathophysiology, University Medical Center Hamburg-Eppendorf, 20246

5 Hamburg, Germany. 2 Department of Psychology and 3 Amsterdam Brain and Cognition (ABC), University of

6 Amsterdam, 1018 WS Amsterdam, The Netherlands.

7 *Correspondence: a.braun@uke.de, t.donner@uke.de

8

9 **Abbreviated title:** Confidence-dependent accumulation of decision variables

10 **Number of pages:** 30

11 **Number of figures:** 6

12 **Number of tables:** 0

13 **Number of multimedia:** 0

14 **Number of 3D models:** 0

15 **Number of words for Abstract:** 248

16 **Number of words for Introduction:** 502

17 **Number of words for Discussion:** 1439

18 **Conflict of Interest:** None

19

25

26 **Author contributions:**

27 A.B., A.E.U. and T.H.D. designed research; A.B. performed research, A.B. analyzed data, AEU re-

28 analyzed the data with separate code, A.B., A.E.U. and T.H.D. wrote the paper.

29

30 **Abstract**

31 Perceptual decision-making is biased by previous events, including the history of preceding choices:

32 Observers tend to repeat (or alternate) their judgments of the sensory environment more often than

33 expected by chance. Computational models postulate that these so-called choice history biases result

34 from the accumulation of internal decision signals across trials. Here, we provide psychophysical

35 evidence for such a mechanism and its adaptive utility. Male and female human observers performed

36 different variants of a challenging visual motion discrimination task near psychophysical threshold. In a

37 first experiment, we decoupled categorical perceptual choices and motor responses on a trial-by-trial

38 basis. Choice history bias was explained by previous perceptual choices, not motor responses,

39 highlighting the importance of internal decision signals in action-independent formats. In a second

40 experiment, observers performed the task in stimulus environments containing different levels of auto-

41 correlation and providing no external feedback about choice correctness. Despite performing under

42 overall high levels of uncertainty, observers adjusted both the strength and the sign of their choice

43 history biases to these environments. When stimulus sequences were dominated by either repetitions

44 or alternations, the individual degree of this adjustment of history bias was about as good a predictor

45 of individual performance as individual perceptual sensitivity. The history bias adjustment scaled with

46 two proxies for observers' confidence about their previous choices (accuracy and reaction time).

47 Taken together, our results are consistent with the idea that action-independent, confidence-

48 modulated decision variables are accumulated across choices in a flexible manner that depends on

49 decision-makers' model of their environment.

50

51 **Significance statement:**

52 Decisions based on sensory input are often influenced by the history of one's preceding choices,

53 manifesting as a bias to systematically repeat (or alternate) choices. We here provide support for the

54 idea that such choice history biases arise from the context-dependent accumulation of a quantity

55 referred to as the decision variable: the variable's sign dictates the choice and its magnitude the

56 confidence about choice correctness. We show that choices are accumulated in an action-

57 independent format and a context-dependent manner, weighted by the confidence about their

58 correctness. This confidence-weighted accumulation of choices enables decision-makers to flexibly

59 adjust their behavior to different sensory environments. The bias adjustment can be as important for

60 optimizing performance as one's sensitivity to the momentary sensory input.

61

Introduction

63 It has been known for almost a century that people's judgments of sensory stimuli do not only depend

64 on the current sensory input, but also on their preceding choices (Fernberger, 1920). Several studies

65 have found that humans and other species repeat (or alternate) their perceptual judgments more often

66 than expected by chance (Gold et al., 2008; Busse et al., 2011; de Lange et al., 2013; Akaishi et al.,

67 2014; Fischer and Whitney, 2014; Fründ et al., 2014; Abrahamyan et al., 2016; Pape and Siegel,

68 2016; St. John-Saaltink et al., 2016; Fritsche et al., 2017; Hwang et al., 2017; Urai et al., 2017). Such

69 choice history biases occur also in other domains of decision-making (Leopold et al., 2002; Allefeld et

70 al., 2013; Padoa-Schioppa, 2013).

71   Computational models posit that choice history biases result from the temporal accumulation of

72 signals from past decisions (Yu and Cohen, 2009; Glaze et al., 2015; Bonaiuto et al., 2016). Such a

73 mechanism may serve to continuously update the decision-makers' prior belief about the upcoming

74 stimulus category and adjust their choice behavior to structured environments (Yu and Cohen, 2009;

75 Glaze et al., 2015). In laboratory perceptual tasks, stimulus sequences are typically uncorrelated by

76 design, so that across-trial accumulation degrades performance (Abrahamyan et al., 2016). By

77 contrast, when stimulus sequences exhibit auto-correlations (Goldfarb et al., 2012; Glaze et al., 2015;

78 Abrahamyan et al., 2016; Kim et al., 2017), history biases should improve performance, provided that

79 the accumulation is context-dependent. Specifically, the accumulation should switch sign between

80 environments dominated by either stability or change (Glaze et al, 2015).

81   Perceptual decisions often have to be made under uncertainty due to weak or ambiguous

82 evidence. This uncertainty (or its complement: confidence) might be important for controlling behavior

83 under conditions, in which the decision-maker receives no immediate external feedback. Indeed,

84 fluctuations of confidence play a key role in a normative model, which postulates the accumulation of

85 the internal decision variable over time (Glaze et al, 2015). The decision variable is the basis of both

86 the categorical choice (Bogacz et al., 2006; Gold and Shadlen, 2007) as well as the confidence about

87 its correctness (Kepecs et al, 2008). Correlates of the decision variable are distributed across many

88 brain regions (Gold and Shadlen, 2007; Siegel et al., 2011; Brody and Hanks, 2016) and expressed as

89 motor plans (Gold and Shadlen, 2007; Donner et al., 2009; de Lange et al., 2013) or in action-

90 independent formats (Bennur and Gold, 2011; Hebart et al., 2012; O'Connell et al., 2012; Hebart et al.,

91   2016). These decision-related neural signals also reflect the graded confidence about the choice

92   (Kiani and Shadlen, 2009; Hebart et al., 2016).

93   Our current study addressed three questions. First, do choice history biases originate from

94   signals in motor or action-independent formats? Second, can these signals be accumulated in a

95   sufficiently flexible manner, so as to adjust history biases to repetitive as well as alternating

96   environments? Third, is the strength of such bias adjustment scaled by confidence? We modeled

97   human choice behavior under experimental manipulations tailored to answering these questions.

98

99   **Materials and Methods**

100 **Participants**

101 We analyzed data from 28 participants and two experiments (referred to as Experiment 1 and 2) in

102 total. All participants gave their written informed consent.

103

104 *Experiment 1*

105 Six healthy participants (2 male and 4 female, mean age: 25; range: 22–29 years) took part in the

106 experiment, which was approved by the ethics committee of the Department of Psychology of the

107 University of Amsterdam (reference number 2011-OP-1588).

108

109 *Experiment 2*

110 26 healthy participants (11 male and 15 female, mean age: 26, range: 20 - 36) took part in the

111 experiment, which was approved by the local ethical review board (Ärztekammer Hamburg, reference

112 number PV4714). Four participants were excluded from the data analysis, so that 22 participants

113 remained for the data analysis. Three of the excluded participants did not complete all sessions and

114 one exhibited substantially worse performance than the rest of the group (64 percent correct overall,

115 63 percent correct for the easiest three motion coherence levels).

116

117

118 **Experimental design**

119 The data from both experiments allowed for quantifying choice history biases during a random dot

120 motion discrimination (up vs. down) task. We used large random dot motion patterns in both

121 experiments, so as to minimize stochastic fluctuations in the effective motion energy across trials (Urai

122   et al., 2017).

123

124   *Experiment 1*

125   The following description summarizes the aspects of the experimental design that were most

126   important to the current paper; a comprehensive description can be found in (Tsetsos et al., 2015).

127   Random dot kinematograms (Figure 1A) were composed of 785 (average) white dots on a black

128   screen. The dots were moving within a circular aperture of 9.1° radius. A red fixation cross of 0.4° x

129   0.4° was centered in the middle of the circle. The dot density was 12.07 dots per deg^2. The population

130   of dots was split into "signal dots" and "noise dots". The signal dots moved either upwards or

131   downwards with a velocity of 2.6°/s. The noise dots changed position randomly from frame to frame.

132   The percentage of signal dots defined the motion coherence, a measure of motion strength. On each

133   trial, three different sequences of dot motion (at the same coherence and direction) were presented in

134   an interleaved fashion, making the effective speed of signal dots 0.87 °/s. One of six different levels of

135   motion coherence, (0.05, 1.26, 3.15, 7.92, 19.91 and 50%) and one of six different viewing durations

136   (150, 300, 600, 1200, 2400, and 4800 ms) were chosen randomly, under the constraint that they

137   occurred equally often within a block of 144 trials. Stimuli were presented on a 22-inch CRT monitor

138   with a resolution of 800 x 600 pixel and a frame rate of 100 Hz at a viewing distance of 68 cm. The

139   participants were instructed to maintain their gaze on the red cross throughout the trial and judge the

140   net motion direction. The motion viewing interval was followed by a variable delay (uniform distribution

141   ranging from 200 to 400 ms), after which the observers had to report their choice by pressing one of

142   two buttons on a computer keyboard, with either the left or the right index finger. Participants received

143   auditory feedback after incorrect responses (a 1000 Hz tone of 100 ms). Perceptual choices ('up' vs.

144   'down' motion direction) were decoupled from motor responses (left vs. right button press) by varying

145   their mapping from trial to trial. This mapping was instructed before motion viewing in one condition

146   ('Pre' condition) and after motion viewing in the other ('Post' condition), by means of a visual cue that

147   presented each direction (as an arrow) on the left or right side (i.e. two possible mappings). This

148   mapping cue was randomly selected on each trial. Conditions alternated across blocks. Observer 1-5

149   participated in both conditions. Observer 6 participated only in the Post condition. The analyses of

150   participants 1-5 presented here were collapsed across both conditions. We obtained the same pattern

151   of results when analyzing the data from both conditions separately (data not shown).

152

153   *Experiment 2*

154   To test for the adaptability of choice history biases, we manipulated the sequential stimulus statistics

155   between experimental sessions, to make people perform the task in 'Repetitive', 'Neutral' (no

156   sequential dependence), or 'Alternating' environments (Figure 1B). Stimuli, task, and procedure for

157   Experiment 2 were identical to Experiment 1, with the following exceptions. The circle within which the

158   dots were moving had an outer radius of 12° and an inner radius of 2°. The density of dots was 1.7

159   dots/deg^2 and each dot had a diameter of 0.2°. The dots moved with a velocity of 11.5°/s. Signal dots

160   had a maximum lifetime of 6 frames. We used the following coherence levels: 0, 5, 10, 20, 40 and

161   60% (equally many trials per coherence level). A red bulls-eye fixation target at the center of the

162   screen as well as randomly moving dots (0% coherence) were presented throughout each block. The

163   first trial of each block started with a baseline interval of 5 s. A beep (duration: 50 ms, sine wave at

164   440 Hz) indicated the onset of the evidence interval with variable coherence levels and directions (see

165   above) after a fixed duration of 0.75 s. A second beep indicated the offset of the evidence interval and

166   prompted the observers' response. Observers reported their perceptual choices by pressing one of

167   two keyboard buttons, with the index finger of the left or right hand. After button press or a response

168   deadline of 3 s, the inter-trial interval started. Inter-trial intervals were uniformly distributed between 1

169   and 5 s. Observers received auditory feedback during the training sessions, but no feedback during

170   the subsequent six sessions of the main experiment. The motion viewing duration of 0.75 s was

171   selected because previous work in monkeys (Kiani et al., 2008) and humans (Tsetsos et al., 2015)

172   found little integration of motion information beyond that duration. We used a fixed mapping between

173   choices and motor responses, whereby the two possible mappings (right-hand button for up, left-hand

174   for down, or vice versa) were counterbalanced across participants. Experiment 2 consisted of seven

175   sessions per participant (one for training and six main sessions), whereby each session was divided

176   into 10 blocks of 60 trials.

177   Critically, the transition probabilities between the two alternative stimulus categories (i.e. up vs.

178   down regardless of coherence) over trials were manipulated across experimental sessions (Figure 1B,

179   right). Specifically, the probability of a repetition was defined as

180    $P(stimulus\ repetition) = 1 - P(stimulus\ alternation) = P(stimulus_n = up \mid stimulus_{n-1} = up) =$

181         $P(stimulus_n = down \mid stimulus_{n-1} = down),$ eq. 1

182   whereby *n* indexes trials. The repetition probability was held constant within each session, but varied

183   across the main experimental sessions between the following values: 0.5 in the 'Neutral' condition, 0.8

184   in the 'Repetitive' condition, and 0.2 in the 'Alternating' condition. The Neutral condition allowed for

185   quantifying observers' intrinsic choice history bias, which we used as a baseline for quantifying their

186   adjustment to the biased sequential statistics of the Repetitive and Alternating conditions.

187   During the training session, the motion direction on each trial was chosen randomly and

188   independently. All participants started with the Neutral condition in session 1 of the main experiment,

189   which was repeated in session 4. Half of the participants then performed the Repetitive condition in

190   sessions 2 and 5 and the Alternating condition in sessions 3 and 6 and conversely for the other half of

191   participants.

192   Observers were instructed to maintain stable fixation and perform the motion discrimination task

193   as accurately as possible. They were informed that the sequential statistics of the stimulus identities

194   would change from session to session, but stay constant within each session. To this end, we told

195   them that the stimulus sequences could be 'as if produced by a coin flip' (Neutral), 'more likely

196   repeating than alternating' (Repetitive), or 'more likely alternating than repeating' (Alternating).

197   Observers were not informed about (i) the order of these conditions, (ii) the exact transition

198   probabilities, (iii) the use of this information for optimizing their behavioral performance.

199

200   **Modeling choice history bias**

201   We used logistic regression to model observers' choice history biases under the different experimental

202   conditions. The basic approach consisted of adding a linear combination of different components of

203   trial history (which depended on the experiment), as a bias term to a logistic function model of the

204   choice probability (Fründ et al., 2014; Urai et al., 2017). We here used a variant that quantified the

205   relative contributions of previous stimuli, choices, and (for Experiment 1) motor responses.

206

207

208   *Basic choice model* using *psychometric function fit*

209   The probability of making one of the two choices $r_t = 1$ ($r_t = 1$ for 'choice up', $r_t = -1$ for 'choice

210   down') on trial t, given the signed stimulus intensity $\widetilde{s_t}$ (i.e., motion coherence times stimulus category)

211   (up or down, coded as 1 and -1) was described by:

212                    $$P(r_t = 1|\widetilde{s_t}) = \gamma + (1 - \gamma - \lambda)\, g(\delta + \alpha\widetilde{s_t}),$$ eq. 2

213   where γ and λ were the lapse rates for the choices $r_t = 1$ and $r_t = -1$, and $g(x) = \frac{1}{1+e^{-x}}$ was the

214   logistic function. The bias term δ, the offset of the psychometric function, described the overall bias for

7

215   one specific choice. α was the slope of the stimulus-dependent part of the psychometric function,

216   quantifying perceptual sensitivity.

217   For visualizing the effect of previous on current choice (Figure 1D), we separated the trials from

218   Neutral into two subsets, conditioned on the choice from the previous trial, and fitted the psychometric

219   function separately to the observed proportion of upward choices in both subsets. Results from three

220   example observers are shown in Figure 1D and discussed in Results.

221

222   *Modeling the contributions of past stimuli, choices, and motor responses to current choice bias*

223   We estimated the contribution of the previous seven stimulus categories and choices by adding a

224   history-dependent bias term to the argument of the logistic function (Fründ et al., 2014):

225    $$P(r_t = 1 | \widetilde{s_t}, \boldsymbol{h_t}) = \gamma + \left(1 - \gamma - \lambda\right) g(\delta(\boldsymbol{h_t}) + \alpha \widetilde{s_t}) \quad \text{eq. 3}$$

226    $$\delta(\boldsymbol{h_t}) = \delta' + \delta_{\text{hist}}(\boldsymbol{h_t}) = \delta' + \sum_{k=1}^{14} \omega_k h_{kt} \quad \text{eq. 4}$$

227   The history bias $\delta_{\text{hist}}(\boldsymbol{h_t}) = \sum_{k=1}^{14} \omega_k h_{kt}$ consisted of the sum of the preceding seven responses r_{t-1} to

228   r_{t-7} and the preceding seven stimulus categories z_{t-1} to z_{t-7}, each multiplied with a weighting factor

229   ω_k. The vector $\boldsymbol{h}_t$ was written as:

230    $$\boldsymbol{h}_t = (r_{t-1}, r_{t-2}, r_{t-3}, r_{t-4}, r_{t-5}, r_{t-6}, r_{t-7}, z_{t-1}, z_{t-2}, z_{t-3}, z_{t-4}, z_{t-5}, z_{t-6}, z_{t-7}).$$

231   All terms in $\boldsymbol{h}_t$ were coded as -1 or 1, with the exception of terms coding for stimuli with zero

232   coherence, which were set to zero. The weighting factors ω_k thus modeled the influence of each of the

233   seven preceding responses and stimulus categories on the current choice. Positive values of ω_k

234   indicated a bias to repeat the choice or stimulus category at the corresponding lag, and negative

235   values of ω_k indicated a tendency to alternate. In this and all subsequent analyses, the parameters of

236   the logistic regression model were fit by maximizing the log-likelihood $L = \sum_t \log P(r_t = 1 | \widetilde{s_t}, \boldsymbol{h_t})$ using

237   an expectation maximization algorithm (Fründ et al., 2014).

238   In Experiment 1, perceptual choices and motor responses were further decoupled through a

239   mapping that varied from trial to trial. Thus, we could independently estimate the relative contribution

240   of previous choices and motor responses to the current choice bias. We added the last seven choices

241    $c_{t-1}, c_{t-2}, c_{t-3}, c_{t-4}, c_{t-5}, c_{t-6}, c_{t-7}$, each one multiplied with a separate set of history weights ω'_k, to the

242   history bias term $\delta_{\text{hist}}(\boldsymbol{h_t}, \boldsymbol{c_t})$.

243    $$\delta(\boldsymbol{h_t}, \boldsymbol{c_t}) = \delta' + \delta_{\text{hist}}(\boldsymbol{h_t}, \boldsymbol{c_t}) = \delta' + \sum_{k=1}^{14} \omega_k h_{kt} + \sum_{k=1}^{7} \omega'_k c_{kt} \quad \text{eq. 5}$$

244   Experiment 1 contained not only trial-to-trial variations in motion direction and coherence, but also in

245   the duration of the dot motion stimulus. To assess the effect of this manipulation, we first fitted the

246   psychometric functions separately for each of the six different motion-viewing durations (on the current

247   trial) and compared the resulting weights within each observer. The viewing duration had only

248   negligible impact on the history weights (data not shown), indicating that the history contributions were

249   invariant across viewing durations. Consequently, we fitted the psychometric functions to the data

250   from all trials, and analyzed the history weights, irrespective of viewing duration.

251   Experiment 1 also contained two conditions (Pre and Post), in which observers were instructed

252   about the required mapping between choice and response either before or after the presentation of

253   the sensory evidence. The analyses presented in Results collapsed across both conditions, but we

254   also verified that there were no differences between these conditions when analyzing them separately

255   (data not shown).

256

257   *Modeling the contributions of past correct (and incorrect) choices to current bias*

258   The weights for previous correct and incorrect choices were estimated by re-combining the weights for

259   previous stimuli and choices estimated by means of equations 3 and 4. Specifically, the weights for

260   correct choices were computed as the sum of choice and stimulus weights and the weights for

261   incorrect choices were computed as the difference between choice and stimulus weights (Fründ et al.,

262   2014). Please note that this is equivalent to fitting a regression model with predictors encoding correct

263   or incorrect choices, along with the chosen category (Busse et al, 2011; Abrahamyan et al, 2016).

264

265   *Modeling the contribution of past decision confidence to current bias*

266   We used a model of statistical decision confidence based on signal detection theory (Kepecs et al.

267   2008; Sanders et al. 2016; Urai et al. 2017) in order to define two behavioral proxies of confidence that

268   could be used in the present study. The model assumes that choices are made based on an internal

269   decision variable (*dv*), which is computed as a transformation of sensory input, corrupted by noise. A

270   choice is made by comparing *dv* to a criterion *c*. Confidence is a function of the distance between *dv*

271   and *c*. When *dv* is far from *c*, the choice is likely to be correct; the probability of the choice being

272   correct approaches chance as *dv* approaches *c*. Specifically, $confidence_i = f(|dv_i - c|)$ where

273    $f(x) = \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{x}{\sigma\sqrt{2}}\right)\right]$ is a monotonic sigmoid function that maps the distance metric x on the

274   probability of making a correct choice (Kepecs et al, 2008; Lak et al. 2014). The model predicts that

275   confidence (i) is larger on correct than on error trials, and (ii) scales oppositely as a function of

276   stimulus strength for correct and error trials. We used two behavioral proxies of the so-defined

9

277   confidence to investigate its impact on the adjustment of choice history biases: (i) Choice accuracy,

278   with correct choices being associated with larger confidence than incorrect choices for all levels of

279   evidence strength in the model described above; and (ii) reaction times (RT), which has been found to

280   reflect decision confidence as defined above in empirical work (Sanders et al., 2016; Urai et al., 2017).

281   When assessing the confidence-dependence of the bias adjustment (i.e. changes in history

282   weights), we restricted the model to the immediately preceding trial (lag 1), at which the bias

283   adjustment was expected to be strongest, but we now estimated the weights separately for each of the

284   different levels of previous motion coherence. This enabled us to control for the trial-to-trial variations

285   of stimulus strength, thus isolating the impact of internal trial-to-trial fluctuations of confidence.

286   In our analysis of the impact of choice accuracy, separate predictors coded for the choice or

287   stimulus categories for each level of (non-zero) previous motion coherence. Because choice accuracy

288   was undefined at 0% coherence, we estimated a single choice weight for previous trials where no

289   decision-relevant sensory evidence was presented. Specifically, we included six regressors in the

290   model that each coded for the previous choice at a given coherence level (zero elsewhere) and we

291   included five regressors that each coded for the previous stimulus category at a given non-zero

292   coherence level (zero elsewhere). To assess the impact of choice accuracy, the stimulus and choice

293   weights were transformed into weights for correct and incorrect choices by re-combining the stimulus

294   and choice weights as described in the section *Modeling the contributions of past correct (and*

295   *incorrect) choices to current bias* above.

296   To assess the effect of RTs, we first normalized RT to make it scale positively with confidence,

297   because of its negative scaling with decision confidence (the shortest RTs correspond to the most

298   confident trials (Sanders et al., 2016; Urai et al., 2017): For each observer and condition, we

299   transformed single-trial RTs as follows:

300    $$RTconf = z(max(log(RT)) - log(RT)) \qquad \text{eq. 6}$$

301   where z denoted z-scoring per individual and condition. This transformation was only applied to

302   simplify the interpretation of the corresponding history terms in terms of confidence-weighting. Without

303   this transformation, the resulting weights were qualitatively identical but sign-flipped, thus reflecting the

304   complement to confidence, decision uncertainty (data not shown).

305   We added a modulation by the above-defined $RTconf$ variable to the logistic regression model,

306   as introduced in (Urai et al., 2017). To this end, we added a term describing the interaction between

307   choice and stimulus category at lag 1 with the previous trial's $RTconf$ separately for each previous

308   coherence level: $\sum_{k=1}^{11}\omega'_k h_t RTconf_{kt}$. Specifically, the interaction terms in this model were six

309   regressors for previous choice multiplied by previous $RTconf$ (one for each coherence level) and five

310   regressors for previous stimulus category multiplied by previous $RTconf$ (one for each non-zero

311   coherence level), and a nuisance covariate $\sum_{k=1}^{6}\omega''_k RTconf_{kt}$ for the main effect of $RTconf$. The full

312   bias term in this model was as follows:

313    $\delta(\boldsymbol{h_t}, \boldsymbol{RTconf_t}) = \delta' + \delta_{\text{hist}}(\boldsymbol{h_t}, \boldsymbol{RTconf_t}) = \delta' + \sum_{k=1}^{11}\omega_k h_{kt} + \omega'_k h_{kt} RTconf_{kt} + \sum_{k=1}^{6}\omega''_k RTconf_{kt}$

314   eq. 7

315

316   *Modeling history contributions in synthetic, non-adjusting observers*

317   We performed two sets of simulations to ensure that the context-dependent shifts in the history

318   weights exhibited by participants in Experiment 2 were not just passively 'inherited from' the correlated

319   stimulus sequences in the Repetitive and Alternating conditions. The rationale of these simulations

320   was to fit the behavior of synthetic observers. These were matched to the behavior of each of our

321   participants in all parameters, except for the history weights displayed in the biased conditions.

322   In the first set of simulations, we constructed observers, who based their decisions only on the

323   current stimulus. For each participant, we estimated the parameters of the psychometric function

324   described by eqs. 3 and 4 from the data of the respective biased conditions. This 'memory-less

325   observer model' was the set of fitted parameters, but with history bias $\delta_{\text{hist}}(\boldsymbol{h_t})$ set to zero; it allowed

326   us to compute the probability of making an 'up'-choice ($r_t = 1$) for any given stimulus intensity in the

327   absence of any influence of past events. To simulate the model's performance in the two biased

328   experimental conditions (Alternating and Repetitive), we used the original sequences of stimuli (motion

329   coherences times directions) seen by each observer in these two conditions, and computed the choice

330   probability for each trial by putting the model parameters and stimulus categories in eq.2. Based on

331   these choice probabilities, we then drew the choices on each trial by a weighted coin-flip, resulting in a

332   sequence of choices generated by the model. We then fitted this choice and stimulus sequence, again

333   using the model specified in eqs. 3 and 4 allowing us to estimate the synthetic observers' stimulus and

334   choice weights. The resulting values served as a reference for the history biases expected as a result

335   of discriminating a biased sequence of stimuli without memory.

336   The second set of simulations was as the first set of simulations, with the exception that the

337   synthetic observers had the same (non-zero) history weights estimated for the real participants in the

338   Neutral condition, but did not adjust these biases to the biased environments. Again, this enabled us

339   to simulate choice patterns of the synthetic observers exposed to the stimulus sequences used in our

340   actual experiment, and to use these choice patterns to estimate the simulated observers' choice and

341   stimulus weights, as described for the first set.

342   In both sets of simulations, we presented the same stimulus sequence 50 times, to average out

343   the effect of binomial noise that was needed to generate choices from the logistic function. This

344   yielded more precise estimates of the model parameters than was possible in the human observers.

345

346   **Statistical tests**

347   We used parametric 2-tailed t-tests for all statistical comparisons of regression weights reported in this

348   paper. The rationale was that we could then also provide Bayes factors (Bf), in order to quantify the

349   posterior belief in the null hypothesis given the evidence (Rouder et al., 2009). $Bf_{10} < 1/3$ indicates

350   evidence in favor of the null hypothesis, $Bf_{10} > 3$ indicates evidence for the alternative hypothesis, and

351    $Bf_{10} = 1$ indicates inconclusive evidence. When performing multiple t-tests of regression weights (e.g.,

352   across seven lags or coherence levels), false discovery rate correction (Benjamini and Hochberg,

353   1995) was applied to correct for multiple comparisons.

354   When testing correlation coefficients computed for individual participants (the so-called 'adaptivity

355   indices' defined in Results) against zero, we first Fisher z-transformed the Pearson correlation

356   coefficients and then submitted them to simple t-tests. We used the parametric 2-tailed Steiger's test

357   (Steiger, 1980) for comparing across-subjects correlations between individual adaptivitiy indices

358   (correlation coefficients) and their proportion of correct choices with the corresponding correlations

359   between individual perceptual sensitivity and the proportion correct choices.

360   Finally, we used circular statistics, specifically Rayleigh's test, to assess the clustering of

361   orientations of the lines connecting the weights from Neutral with those from either Repetitive or

362   Alternating conditions, respectively. A Hotelling test (van den Brink et al., 2014) was used to assess

363   the difference in mean directions of adjustment between these two conditions.

364   The results from all the all regression weights and individual adaptivity indices (correlation

365   coefficients) in Experiment 2 were analogous when replacing the parametric tests with non-parametric

366   permutation tests (Efron and Tibshirani, 1998) with N = 10.000 permutations.

367

368   **Results**

369   We here report results from two experiments (referred to as Experiment 1 and 2) quantifying choice

370   history biases during the random dot motion discrimination task that is widely used in

371   neurophysiological studies of perceptual decision-making (Gold and Shadlen, 2007; Siegel et al.,

372   2011; Kelly and O'Connell, 2015). The two experiments aimed to manipulate different aspects of

373   choice behavior. Analyses of behavior from Experiment 1 were previously published (Tsetsos et al.,

374   2015), but those analyses did not assess sequential effects. Here, we re-analyzed these data to

375   quantify the dependence of choice on previous stimuli, choices, and motor responses.

376   Figure 1C and D illustrates behavioral patterns generated by choice history biases in example

377   observers from the Neutral condition of Experiment 2 (i.e. no correlations among successive stimuli).

378   Figure 1C shows, for one observer, a 'streak' of eight repeats of the same choice, followed by five

379   repeats of the other choice. These streaks occur in the face of trial-to-trial variations of the direction of

380   the random stimuli. Critically, such apparent biases towards one or the other choice emerge only

381   locally in time. Choice history biases are therefore distinct from the 'global' biases towards one

382   particular choice that result from uneven probabilities of the two stimulus categories or uneven payoffs

383   for the two options (Bogacz et al., 2006; Mulder et al., 2012; de Lange et al., 2013). One way to isolate

384   choice history biases is to fit, for each observer, two separate psychometric functions (relating signed

385   stimulus strength to choice probability), each conditioned on the choice the observer made on the

386   previous trial. Choice history biases are then evident as horizontal shifts between these two functions.

387   Figure 1D displays the resulting functions of three example observers (Neutral condition from

388   Experiment 2) with an intrinsic bias to repeat (left panel, same observer as in Figure 1C) or to

389   alternate choices (right), or no bias (middle). A more comprehensive approach is to explicitly model

390   the relative contribution of previous choices, or other experimental variables from previous trials, to

391   current choice bias (Busse et al., 2011; Fründ et al., 2014) (see Materials and Methods, section

392   *Modeling choice history bias*). We used this statistical modeling approach throughout this paper.

393   Our analyses pursued two main objectives. First, we aimed to disentangle and compare the

394   contribution of decisional and motor processing stages to the history biases. Second, we aimed to

395   quantify the adjustment of choice history biases to the environment, as a function of varying levels of

396   decision confidence, in the absence of external feedback.

397

398   **Experiment 1: Disentangling the impact of previous stimuli, choices, and motor responses**

399   In laboratory tasks, perceptual choice and motor response used for reporting the choice are typically

400   coupled, but can be decoupled with little effect on performance on the current trial (Tsetsos et al.,

401   2015). While there is evidence for either decisional or motor origin of history biases (Akaishi et al.,

402   2014; Pape and Siegel, 2016; St. John-Saaltink et al., 2016), their relative contributions have not yet

403   been systematically compared across several trials in the past. To do so, we reanalyzed data from a

404   previously published study (Tsetsos et al., 2015), in which observers performed a random dot motion

405   task under trial-to-trial variations in the mapping between choice and motor response. The direction of

406   motion was chosen randomly and independently on each trial, so that maximizing performance

407   required basing choices solely on the current stimulus and not on its history (i.e., previous stimuli,

408   choices, or motor responses).

409   Observers showed a significant tendency to repeat their previous choices (indicated by positive

410   choice weights), but not their motor responses (Figure 2A). The effect of the previous choice on

411   current choice was positive and stronger than the effect of the previous motor response (Figure 2A).

412   The response weights did not differ significantly from zero for any lag ($Bf_{10} < 0.45$ for all lags).

413   Preceding stimulus categories (up/down) exhibited negative, albeit not statistically significant, weights

414   at longer lags (Figure 2B), possibly reflecting the impact of long-lasting, repulsive effects of direction-

415   selective sensory adaptation mechanisms (Kohn, 2007) on choice behavior.

416   These results indicate that the commonly observed choice repetition biases are specifically due to

417   previous choices and not the motor responses used to the report them, which has implications for their

418   neural bases (see Discussion). We next investigated the adjustment of choice history biases (under

419   fixed mapping) to varying environmental statistics in order to gain deeper insights into their functional

420   origin and adaptive utility.

421

422   **Experiment 2: Confidence-dependent adjustment of choice history biases to the environment**

423   In laboratory tasks used to study perceptual choice, it is common to generate random sequences of

424   the two alternative stimulus categories. But the states of natural environments, and hence the sensory

425   signals generated by them, often exhibit significant auto-correlations across time, so that it might be

426   beneficial for decision-makers to adjust their choice history biases to this correlation structure (Yu and

427   Cohen, 2009). In Experiment 2, we tested for such adjustments, by systematically manipulating the

428   repetition probabilities between the two possible motion directions across three conditions blocked by

429   experimental session: Repetitive, Alternating, and Neutral (two sessions per condition; see Figure 1B

430   and Materials and Methods). Importantly, observers received no external feedback about the

431   correctness of their choices. This enabled us to study the impact of their decision confidence on the

432   adjustment of their choice history biases to environmental statistics.

433

434   *Adjustment of choice history biases to environmental statistics*

435   The manipulation of the environmental statistics had robust effects on observers' history biases. We

436   visualized those in two complementary ways focusing on different aspects of the data. Both our

437   approaches were guided by the statistical structure of the Repetitive and Alternating conditions, which

438   yielded characteristic profiles of the probability of stimulus repetitions as a function of lag: For both

439   Repetitive and Alternating conditions, repetition probability was most strongly biased (i.e., different

440   from 0.5) at lag 1, and progressively approaching 0.5 for larger lags (Figure 3A). Thus, the strongest

441   effects were expected for events from the preceding trial (i.e., lag 1).

442   Our first approach, therefore, focused on the weights for lag 1. When plotting the choice weights

443   against stimulus weights, data points located in the upper-right triangular part indicated a tendency to

444   repeat the previous choice or stimulus categories (up/down), whereas data points in the lower-left

445   triangular part indicated a tendency to alternate (Figure 3B, dashed diagonal line). If observers

446   adjusted their choice patterns to the Repetitive and Alternating environments, their history weights

447   should have shifted in the corresponding directions. This is what we observed (Figure 3B; compare

448   dots of different colors). The weights were close to zero in the Neutral condition (group average, red

449   'x'); weights shifted towards repetition in the Repetitive condition (group average, green arrow in

450   Figure 3B), and alternation in the Alternating condition (group average, blue arrow in Figure 3B),

451   respectively. The vector angles of the shift from Neutral were significantly different from uniform ($z =$

452   7.69, $p = 0.0003$ in Repetitive and $z = 8.64$, $p < 0.0001$ in the Alternating condition, Rayleigh's test),

453   and the shift angles were significantly different between Repetitive and Alternating ($F(2, 20) = 60.28$, p

454   < 0.0001, Hotelling test). The adjustment of choice history bias was also evident when fitting the

455   psychometric function conditioned on the previous choice (as in Figure 1D). Both conditions were

456   characterized by a history-dependent shift, in opposite directions (Figure 3C; difference in shift

457   between Repetitive and Alternating: $t(21) = 3.21$, $p = 0.0042$). By contrast, previous choice had no

458   effect on the slope of the psychometric function (difference in history-dependent change in the slope

459   between Repetitive and Alternating: $t(21) = -0.60$, $p = 0.5532$, $Bf_{10} = 0.2627$).

460   It is noteworthy that the direction of the shift of history weights between the different

461   environmental statistics was largely along the positive diagonal (Figure 3B) that corresponds to equal

15

462   weights for previous stimuli and choices. Thus, the bias adjustment was largely driven by correct

463   choices (where previous choices and stimuli were identical). We thus used the weight of previous

464   correct choices (i.e., the sum of choice and stimulus weights, see Materials and Methods) in all

465   subsequent analyses as a single metric of the bias adjustment. The importance of previous correct

466   choice for the bias adjustment was indicative of the role of decision confidence, an aspect that we

467   elaborate on in the section *Modulation of choice history bias adjustment by decision confidence* below.

468   Our second approach focused on an assessment of the full time courses of the history weights.

469   The temporal profiles of the stimulus repetition probabilities in the two biased conditions exhibited

470   markedly different patterns: In the Repetitive condition the temporal profile exhibited a monotonic

471   decay towards 0.5, whereas it exhibited a damped oscillation around 0.5 in the Alternating condition

472   (Figure 3A). The correlation between both time courses for Repetitive and Alternating conditions was

473   negative. In what follows, we refer to these time courses as 'history templates', to indicate that these

474   characterize the statistical structure of the environment.

475   Indeed, participants' history weights exhibited profiles that were similar to those of the history

476   templates (Figure 3D, compare with Figure 3A). We use the term 'history kernel' to refer to the

477   individual courses of the weights for correct choices as a function of lag. We quantified their similarity

478   with the corresponding history templates by means of temporal correlation (Figure 3E). These

479   correlations were significant in both conditions (Repetitive: $t(21) = 2.57$, $p = 0.0179$, Alternating: $t(21) =$

480   4.83, $p < 0.0001$). Thus, participants adjusted their history biases to the statistical structure of their

481   environments with a time course matched to the full environmental statistics. In what follows, we refer

482   to this similarity metric as 'adaptivity index'.

483   One concern might be that even an observer who only discriminates the current sensory

484   evidence, without any active accumulation of past experimental events, might exhibit similar shifts in

485   the history weights between the Alternating and Repetitive conditions, by virtue of the stimulus

486   statistics propagating into the history weights without any active adjustment of the observer. To

487   address this concern, we simulated the performance of two types of synthetic observers, which were

488   constructed individually for each of our participants. These had the same perceptual sensitivity as

489   each participant, but without any adjustment of stimulus and choice weights to the different

490   environmental conditions (see Materials and Methods). The first set of synthetic observers had

491   stimulus and choice weights of zero. The second set of synthetic observers had the same history

492   biases (i.e., non-zero choice and stimulus weights) as our participants in the Neutral condition, but did

493   not adjust these to the changing environmental statistics. The choice and stimulus weights obtained

494   for both these models were not significantly different from zero (Figure 4A and B) (all p-values > 0.567

495   and Bf_{10} ranging from 0.22 to 0.62).

496   These simulation results indicate that the effect of the correlations between stimuli on choice

497   patterns was reliably soaked up by the stimulus-dependent part of our statistical model (i.e. the slope

498   of the psychometric function). In other words, the systematic deviations of the history weights between

499   Repetitive and Alternating conditions evident in the real observers were not just passively 'inherited

500   from' the correlations evident in the stimulus sequences, but rather due to an active adjustment of their

501   biases.

502

503

504

505   *History bias adjustment predicts performance in biased environments*

506   While the bias adjustment was highly consistent across participants, individuals differed in the extent

507   to which they shifted their history biases between conditions (i.e., the magnitude of their adaptivity

508   indices, Figure 3E). We correlated the individual adaptivity indices with the proportion of correct

509   choices to assess their predictive value for overall task performance (Figure 5A). The more strongly

510   observers adjusted, the more successful they were in both, the Repetitive (Figure 5A, left panel) and

511   Alternating (middle) condition. We found no evidence for such an effect in the Neutral condition

512   (Figure 5A, right, Bf_{10} = 0.4276). As expected, perceptual sensitivity (i.e., the slope of the

513   psychometric function) was also strongly predictive of individual performance in all three conditions

514   (Figure 5B). In both biased environments the adaptivity index was similarly predictive of performance

515   as perceptual sensitivity (Repetitive: Steiger's test z = 0.28, p = 0.7791; Alternating: z = 0.4, p =

516   0.6893), while sensitivity was a better predictor in the Neutral environment (z = 2.70, p = 0.0068). In

517   other words, the adjustment to the environmental statistic can be about as important for maximizing

518   reward rate as sensitivity to the momentary sensory evidence.

519   Taken together, our results reported so far supported the idea that participants accumulated

520   internal signals from their previous correct decisions into biases for their current choice – a process

521   that adjusted their behavior to the statistics of their environment and improved performance. Some

522   previous accounts of sequential effects have postulated the accumulation of external variables, such

523   as stimulus repetitions (Yu and Cohen, 2009; Meyniel et al., 2016), performance feedback

524   (Abrahamyan et al., 2016), or reward (Sugrue, 2004). Our experimental conditions precluded any of

525   the above: Observers performed under generally high uncertainty about the veridical stimulus

526   identities, and they did not receive external feedback about choice outcomes. We reasoned that,

527   under these conditions, observers may have accumulated the internal decision variables, on which

528   they based their choices in a context-dependent manner (i.e., with opposite sign for Repetitive and

529   Alternating environments). This interpretation is in line with a normative model of sequential effects

530   (Glaze et al, 2015). In statistical decision theory, as well as in neural signals observed in the brain, the

531   decision variable not only encodes the categorical choice, but also the graded confidence about that

532   choice (Kepecs et al., 2008; Kiani and Shadlen, 2009; Hebart et al., 2016). Consequently, we

533   reasoned that the impact of previous choices on current bias should depend on the confidence

534   associated with the previous choices. Our final set of analyses tested this hypothesis.

535   *Modulation of choice history bias adjustment by confidence*

536   We here use the term 'decision confidence' in a statistical sense, to refer to the posterior probability

537   that a choice is correct, given the evidence (Kepecs et al., 2008; Pouget et al., 2016; Sanders et al.,

538   2016; Urai et al., 2017). The key features of a model formalizing this construct are reproduced in

539   Figure 6A (Kepecs et al., 2008; Sanders et al., 2016; Urai et al., 2017; see Methods). This definition of

540   confidence is agnostic about the link to the subjective sense of confidence, or the ability to report this

541   sense of confidence (but see Sanders et al, 2016).

542   We used two experimental variables consistent with this definition of confidence: accuracy and

543   RT. Correct choices are overall associated with higher confidence for all non-zero evidence strengths

544   (i.e., coherence levels, Figure 6A, top). The scaling of RT with motion coherence exhibited the same

545   characteristic signature as uncertainty (i.e., the complement of confidence) as reported in previous

546   studies (Sanders et al., 2016; Urai et al., 2017): RT decreased with coherence for correct choices, but

547   increased for incorrect choices (Figure 6D, compare to Figure 6A). Linear regression revealed an

548   opposite-signed relationship between motion coherence and RT, separately for correct (β = -0.150,

549   s.e.m. = 0.027, p = 0.005) and error trials (β = 0.628, s.e.m. = 0.025, p = 0.025).

550   As predicted, the leverage of the preceding choice on bias adjustment (i.e., difference between

551   Repetitive and Alternating) was larger when the previous choice was correct than incorrect, even

552   when controlling for the level of previous motion coherence (Figure 6B, C). There was a significant

553   effect of previous correct choice at all of the previous coherence levels, while there was no such effect

554   for previous incorrect choices at any of the previous coherence levels (Figure 6B; Bf_{10}: 0.23, 0.69,

555   0.23, 0.34, and 0.28 for previous coherence levels: 0.05, 0.1, 0.2, 0.4, and 0.6, respectively).

556   When pooled across previous coherence levels, the weights for previous correct choices deviated

557   significantly from zero in both conditions (Figure 6C; Repetitive: $t(21) = 2.58$, $p = 0.0174$; Alternating:

558    $t(21) = -4.42$, $p = 0.0002$). Again there was no such effect for previous errors (Figure 6C; Repetitive:

559    $t(21) = -0.13$, $p = 0.8985$, $Bf_{10} = 0.22$; Alternating: $t(21) = 0.35$, $p = 0.7309$, $Bf_{10} = 0.24$). The weights

560   were significantly larger for correct than incorrect previous choices in Repetitive ($t(21) = 3.06$, $p =$

561   0.0060, and the other way around in the Alternating ($t(21) = -5.19$, $p < 0.0001$Figure 6C). Please note

562   that while the weights were averaged across previous coherence levels for visualization in Figure 6C,

563   they were first estimated separately for each previous coherence in order to factor out effects of trial-

564   to-trial fluctuations in sensory evidence strength (see Materials and Methods). These results were

565   qualitatively identical after random subsampling of the correct trials, so as to match the smaller

566   number of incorrect trials for each previous coherence level (data not shown), ruling out the concern

567   that the stronger bias adjustment after correct choices may have been due to the larger number of

568   correct than error trials. In sum, these results were consistent with the idea that the weight of choices

569   in the across-trial accumulation process depended on internal (i.e., stimulus-independent) fluctuations

570   in decision confidence.

571   To assess the modulatory effect of the second confidence proxy, RT, on the bias adjustment, we

572   built on an extension of the statistical model by multiplicative interaction terms. This quantified the

573   degree to which the impact of previous correct choices on current bias was modulated by previous RT

574   (see Materials and Methods for details). In these model fits, we transformed RT to scale positively with

575   decision confidence, a variable we refer to as $RTconf$ (Materials and Methods). Again, we split trials

576   by their motion coherence to assess the modulatory effect of $RTconf$ on the impact of correct choices

577   (i.e. the weights for the interaction term $RTconf \; x \; correct$), over and above variations in evidence

578   strength.

579   Larger values of $RTconf$ were associated with a stronger impact of the previous (correct) choice

580   on the current bias (Figure 6E, F), an effect that was robust even when we evaluated each previous

581   coherence level separately (Figure 6E). When pooled across previous coherence levels, the

582   interaction weight was significantly larger than zero in the Repetitive condition (Figure 6F; $t(21) = 3.84$,

583    $p = 0.0009$), indicating a confidence-dependent enhancement of the tendency to repeat correct

584   choices in that condition. Conversely, the interaction weight was significantly smaller than zero in the

585   Alternating condition (Figure 6F; t(21) = -3.71, p = 0.0013), indicating a confidence-dependent

586   enhancement of the tendency to alternate correct choices. Thus, even within the correct choices,

587   evidence-independent fluctuations in the associated confidence (indexed by $RTconf$) boosted their

588   impact on future choice bias.

589   Taken together, two independent proxies of decision confidence, choice accuracy and reaction

590   time, both supported the conclusion that decision confidence boosted the adjustment of choice history

591   biases to the structure of the environment.

592

593

594

595   **Discussion**

596   Choice history biases are a pervasive phenomenon in perceptual decision-making (Fernberger, 1920;

597   Fründ et al., 2014). Here, we showed that these biases were largely dominated by categorical choices

598   rather than motor responses and could be flexibly adjusted to environmental statistics even in the

599   absence of feedback about choice outcome. In line with recent normative accounts, the strength of

600   this adjustment was modulated by previous decision confidence. In environments with strong

601   sequential structure it governed individual performance to a similar extent as perceptual sensitivity.

602   Taken together, our results yield new insights into the functional origins and adaptive utility of choice

603   history biases, with direct implications for their neural bases.

604   An important novel contribution of our study is the discovery of a confidence-weighted adjustment

605   of choice history biases to changing environments. We propose that this was due to a context-

606   dependent accumulation of decision variables across trials. A similar accumulation process has been

607   proposed to explain sequential effects under strong and unambiguous evidence (Yu and Cohen, 2009;

608   Meyniel et al., 2016). Consequently, those latter models describe the accumulation of external

609   observables rather than internal decision variables. The latter are often dissociated from external

610   observables when the decision-maker is uncertain about the environmental state due to degraded

611   evidence. While temporal accumulation is a widely established mechanism in perceptual choice

612   (Bogacz et al., 2006; Gold and Shadlen, 2007; Ratcliff and McKoon, 2008; Wang, 2008; Ossmy et al.,

613   2013), previous models focus on the within-trial accumulation of the momentary sensory evidence.

614   Across-trial accumulation of information is long established in the theory of reinforcement learning, but

615   there it pertains to the accumulation of rewards (i.e., external signals about choice outcome) and

616   spans substantially longer timescales than sensory evidence accumulation (Sutton and Barto, 1998;

617   Sugrue, 2004; Glimcher, 2011). Our current results are indicative of an accumulation mechanism that

618   operates (i) with a timescale situated in between those established for sensory evidence accumulation

619   and action value learning, and (ii) on internal decision variables, which themselves result from the

620   faster (within-trial) accumulation of sensory evidence.

621   Such a context-dependent across-trial accumulation of decision variables has been postulated by

622   a recent normative account (Glaze et al., 2015), and shown to account for history biases in simple

623   saccadic choice (Kim et al., 2017). Little is currently known about the neural basis of this process. Our

624   current work sets the stage for probing into its neural basis, by experimentally establishing key

625   behavioral hallmarks of this accumulation process within the most widely used task in the

626   neurophysiology of perceptual decision-making (Gold and Shadlen, 2007; Siegel et al., 2011; Kelly

627   and O'Connell, 2015). It will now be important to explore the underlying mechanisms through direct

628   recordings of neural activity under conditions as used here.

629   One recent study provided similar evidence for an effective adjustment of human observers to

630   changing environmental statistics (Abrahamyan et al., 2016). Our current results and those from

631   Abrahamyan et al (2016) complement each other in quantifying the adaptability of human choice

632   history biases. In Abrahamyan et al. the nature of the change in the environment was different from

633   the one we have used here: In their study, observers received unambiguous feedback about the

634   outcome of each choice, and the manipulation of the stimulus sequence depended on the participants'

635   success or failure. Consequently, the process adjusting history biases likely depended on the

636   combination of choices and their outcome. By contrast, in our study, the environments differed in their

637   statistical structure independent of observers' choices. Furthermore, participants remained uncertain

638   about their choice outcomes, and therefore could only base their history biases on internal signals.

639   This was likely the key aspect that mediated the confidence-weighting of the impact of previous

640   choices on current bias in our present study. Thus, the adjustment effects in our study and the one

641   from Abrahamyan and colleagues (2016) likely resulted from distinct computational mechanisms.

642   The interpretation provided above as well as the normative framework by Glaze et al (2015) offer

643   a natural interpretation of the modulatory effect of decision confidence on the adjustment of choice

644   behavior. In statistical decision theory, as well as in neural signals observed in the brain, the decision

645   variable does not only encode the categorical choice, but also the graded confidence about that

646   choice (Kepecs et al., 2008; Kiani and Shadlen, 2009; Hebart et al., 2016). This quantity is the best

647   proxy for the true state of the environment available to the decision-maker in the absence of external

648   feedback. A decision variable large in magnitude implies large confidence and predicts accurate as

649   well as fast decisions (Sanders et al., 2016). Thus, across-trial accumulation of decision variables into

650   choice history bias predicts that correct or fast choices have a stronger impact on the history bias

651   adjustment to environmental statistics – just as we observed in our second experiment. The same idea

652   can account for the observation (Urai et al., 2017) that 'intrinsic' history biases emerging under

653   random stimulus sequences, regardless of their direction (i.e., towards alternation or repetition), are

654   weaker following low-confidence decisions (i.e., long reaction times). In such contexts, corresponding

655   to the Neutral condition in our Experiment 2, observers' biases might result from biased internal

656   representations of the environmental structure (i.e., biased 'subjective hazard rates' in the model by

657   Glaze et al, 2015)). Taken together, confidence-weighting of the impact of previous choices on current

658   bias may be a diagnostic feature of the accumulation of graded decision variables across trials that we

659   propose as a mechanism underlying the history bias adjustment in our experiment.

660   In our account the strength of history bias adjustment depends on the magnitude of the decision

661   variable, which is also the sole source of variations in confidence in the confidence model from Figure

662   6A (Kepecs et al, 2008). According to this model the difference in confidence between correct and

663   error trials increases as a function of stimulus strength. Thus, one would expect the impact of

664   correctness on bias adjustment to also increase as a function of stimulus strength. Such an increase

665   was not evident in our data (Figure 6B). A possible explanation is that choice accuracy was less

666   closely coupled to the decision variable than postulated by the confidence model from Figure 6A. For

667   example, some errors will be caused by noise downstream from the decision variable and

668   consequently not affect the bias adjustment. Thus, motor errors will counteract the dependence of

669   history bias adjustment on correctness. Because misperception of the true stimulus category becomes

670   less likely with stronger evidence, the relative contribution of motor errors to incorrect choices will

671   increases as a function of evidence strength. This might explain why the effect of previous correctness

672   on history bias adjustment did not increase as a function of previous evidence strength in our data.

673   Similar considerations hold for our second confidence proxy, reaction time.

674   Our analyses revealed that the contributions of previous stimuli, perceptual choices, and motor

675   responses were dissociable in terms of their strength, sign, and time course. Importantly, the dominant

676   and consistent bias in standard conditions with random stimulus sequences, was to repeat preceding

677   choices, rather than motor responses. Two recent studies similarly decoupled perceptual choice and

678   motor response (Akaishi et al., 2014; Pape and Siegel, 2016). One of them (Pape and Siegel, 2016)

679   showed that a bias to alternate response hands from trial to trial systematically contributed to

680   sequential effects, due to activity dynamics within motor cortex. This motor response alternation bias

681   was superimposed onto a choice repetition bias in their study, but Pape and Siegel (2016) did not

682   compare the magnitude and time course of these two effects directly. When performing such a direct

683   comparison, we here found the contribution of previous choices to be significantly stronger, and more

684   prolonged in time. The predominance of choices over motor responses is consistent with the results

685   (focusing on the preceding decision only) from Akaishi et al (2014). Taken together, the data by

686   Akaishi et al (2014) and our present study indicate that history biases in perceptual decision-making

687   are governed by decision variables encoded in an abstract, action-independent format. Such

688   representations of the decision variable exist in associative brain regions, such as posterior parietal or

689   prefrontal cortex (Bennur and Gold, 2011; Hebart et al., 2012, 2016), which also exhibit the short-term

690   memory dynamics necessary for the persistence of biases in the decision-making machinery (Wang,

691   2002; Bonaiuto et al., 2016; Morcos and Harvey, 2016).

692   We conclude that human observers accumulate action-independent, graded decision variables

693   across trials towards biases for upcoming choices in a context-dependent manner. This process

694   enables observers to adjust their choice behavior to environmental statistics in the absence of

695   unambiguous information about choice outcome. Our findings are in line with normative theory and

696   constrain the candidate neural sources of choice history biases.

697

698   **References**

699   Abrahamyan A, Silva LL, Dakin SC, Carandini M, Gardner JL (2016) Adaptable history biases in human

700   perceptual decisions. Proc Natl Acad Sci 113:E3548–E3557.

701   Akaishi R, Umeda K, Nagase A, Sakai K (2014) Autonomous Mechanism of Internal Choice Estimate Underlies

702   Decision Inertia. Neuron 81:195–206.

703   Allefeld C, Soon CS, Bogler C, Heinzle J, Haynes J-D (2013) Sequential dependencies between trials in free

704   choice tasks. arxiv.org.

705   Benjamini Y, Hochberg Y (1995) Controlling the False Discovery Rate: A Practical and Powerful Approach to

706   Multiple Testing. J R Stat Soc Ser B Methodol 57:289–300.

707   Bennur S, Gold JI (2011) Distinct Representations of a Perceptual Decision and the Associated Oculomotor

708   Plan in the Monkey Lateral Intraparietal Area. J Neurosci 31:913–921.

709   Bogacz R, Brown E, Moehlis J, Holmes P, Cohen JD (2006) The physics of optimal decision making: A formal

710   analysis of models of performance in two-alternative forced-choice tasks. Psychol Rev 113:700–765.

711   Bonaiuto JJ, Berker A de, Bestmann S (2016) Response repetition biases in human perceptual decisions are

712   explained by activity decay in competitive attractor models. eLife 5 Available at:

713   http://elifesciences.org/lookup/doi/10.7554/eLife.20047 [Accessed July 7, 2017].

714   Brody CD, Hanks TD (2016) Neural underpinnings of the evidence accumulator. Curr Opin Neurobiol 37:149–

715   157.

716   Busse L, Ayaz A, Dhruv NT, Katzner S, Saleem AB, Scholvinck ML, Zaharia AD, Carandini M (2011) The

717   Detection of Visual Contrast in the Behaving Mouse. J Neurosci 31:11351–11361.

718   de Lange FP, Rahnev DA, Donner TH, Lau H (2013) Prestimulus Oscillatory Activity over Motor Cortex Reflects

719   Perceptual Expectations. J Neurosci 33:1400–1410.

720   Donner TH, Siegel M, Fries P, Engel AK (2009) Buildup of Choice-Predictive Activity in Human Motor Cortex

721   during Perceptual Decision Making. Curr Biol 19:1581–1585.

722   Efron B, Tibshirani R (1998) An introduction to the bootstrap. Boca Raton: Chapman & Hall/CRC.

723   Fischer J, Whitney D (2014) Serial dependence in visual perception. Nat Neurosci 17:738–743.

724   Fritsche M, Mostert P, de Lange FP (2017) Opposite Effects of Recent History on Perception and Decision. Curr

725   Biol 27:590–595.

726   Fründ I, Wichmann FA, Macke JH (2014) Quantifying the effect of intertrial dependence on perceptual

727   decisions. J Vis 14:9–9.

728   Glaze CM, Kable JW, Gold JI (2015) Normative evidence accumulation in unpredictable environments. eLife 4

729   Available at: http://elifesciences.org/lookup/doi/10.7554/eLife.08825 [Accessed March 13, 2017].

730   Glimcher PW (2011) Understanding dopamine and reinforcement learning: The dopamine reward prediction

731   error hypothesis. Proc Natl Acad Sci 108:15647–15654.

732   Gold JI, Law C-T, Connolly P, Bennur S (2008) The Relative Influences of Priors and Sensory Evidence on an

733   Oculomotor Decision Variable During Perceptual Learning. J Neurophysiol 100:2653–2668.

734   Gold JI, Shadlen MN (2007) The Neural Basis of Decision Making. Annu Rev Neurosci 30:535–574.

735   Goldfarb S, Wong-Lin K, Schwemmer M, Leonard NE, Holmes P (2012) Can Post-Error Dynamics Explain

736   Sequential Reaction Time Patterns? Front Psychol 3 Available at:

737   http://journal.frontiersin.org/article/10.3389/fpsyg.2012.00213/abstract [Accessed March 13, 2017].

738   Hebart MN, Donner TH, Haynes J-D (2012) Human visual and parietal cortex encode visual choices

739   independent of motor plans. NeuroImage 63:1393–1403.

740   Hebart MN, Schriever Y, Donner TH, Haynes J-D (2016) The Relationship between Perceptual Decision

741   Variables and Confidence in the Human Brain. Cereb Cortex 26:118–130.

742   Hwang EJ, Dahlen JE, Mukundan M, Komiyama T (2017) History-based action selection bias in posterior

743   parietal cortex. Nat Commun 8 Available at: http://www.nature.com/articles/s41467-017-01356-z

744   [Accessed November 29, 2017].

745   Kelly SP, O'Connell RG (2015) The neural processes underlying perceptual decision making in humans: Recent

746   progress and future directions. J Physiol-Paris 109:27–37.

747   Kepecs A, Uchida N, Zariwala HA, Mainen ZF (2008) Neural correlates, computation and behavioural impact of

748   decision confidence. Nature 455:227–231.

749   Kiani R, Hanks TD, Shadlen MN (2008) Bounded Integration in Parietal Cortex Underlies Decisions Even When

750   Viewing Duration Is Dictated by the Environment. J Neurosci 28:3017–3029.

751   Kiani R, Shadlen MN (2009) Representation of Confidence Associated with a Decision by Neurons in the

752   Parietal Cortex. Science 324:759–764.

753   Kim TD, Kabir M, Gold JI (2017) Coupled Decision Processes Update and Maintain Saccadic Priors in a Dynamic

754   Environment. J Neurosci 37:3632–3645.

755   Kohn A (2007) Visual Adaptation: Physiology, Mechanisms, and Functional Benefits. J Neurophysiol 97:3155–

756   3164.

757   Leopold DA, Wilke M, Maier A, Logothetis NK (2002) Stable perception of visually ambiguous patterns. Nat

758   Neurosci 5:605–609.

759   Meyniel F, Maheu M, Dehaene S (2016) Human Inferences about Sequences: A Minimal Transition Probability

760   Model Gershman SJ, ed. PLOS Comput Biol 12:e1005260.

761   Morcos AS, Harvey CD (2016) History-dependent variability in population dynamics during evidence

762   accumulation in cortex. Nat Neurosci 19:1672–1681.

763   Mulder MJ, Wagenmakers E-J, Ratcliff R, Boekel W, Forstmann BU (2012) Bias in the Brain: A Diffusion Model

764   Analysis of Prior Probability and Potential Payoff. J Neurosci 32:2335–2343.

765   O'Connell RG, Dockree PM, Kelly SP (2012) A supramodal accumulation-to-bound signal that determines

766   perceptual decisions in humans. Nat Neurosci 15:1729–1735.

767   Ossmy O, Moran R, Pfeffer T, Tsetsos K, Usher M, Donner TH (2013) The Timescale of Perceptual Evidence

768   Integration Can Be Adapted to the Environment. Curr Biol 23:981–986.

769   Padoa-Schioppa C (2013) Neuronal Origins of Choice Variability in Economic Decisions. Neuron 80:1322–1336.

770   Pape A-A, Siegel M (2016) Motor cortex activity predicts response alternation during sensorimotor decisions.

771   Nat Commun 7:13098.

772   Pouget A, Drugowitsch J, Kepecs A (2016) Confidence and certainty: distinct probabilistic quantities for

773   different goals. Nat Neurosci 19:366–374.

774   Ratcliff R, McKoon G (2008) The Diffusion Decision Model: Theory and Data for Two-Choice Decision Tasks.

775   Neural Comput 20:873–922.

776   Rouder JN, Speckman PL, Sun D, Morey RD, Iverson G (2009) Bayesian t tests for accepting and rejecting the

777   null hypothesis. Psychon Bull Rev 16:225–237.

778   Sanders JI, Hangya B, Kepecs A (2016) Signatures of a Statistical Computation in the Human Sense of

779   Confidence. Neuron 90:499–506.

780   Siegel M, Engel AK, Donner TH (2011) Cortical Network Dynamics of Perceptual Decision-Making in the

781   Human Brain. Front Hum Neurosci 5 Available at:

782   http://journal.frontiersin.org/article/10.3389/fnhum.2011.00021/abstract [Accessed April 8, 2017].

783   St. John-Saaltink E, Kok P, Lau HC, de Lange FP (2016) Serial Dependence in Perceptual Decisions Is Reflected

784   in Activity Patterns in Primary Visual Cortex. J Neurosci 36:6186–6192.

785   Steiger JH (1980) Tests for comparing elements of a correlation matrix. Psychol Bull 87:245–251.

786   Sugrue LP (2004) Matching Behavior and the Representation of Value in the Parietal Cortex. Science

787   304:1782–1787.

788   Sutton RS, Barto AG (1998) Reinforcement Learning: An Introduction (Adaptive Computation and Machine

789   Learning). The MIT Press.

790   Tsetsos K, Pfeffer T, Jentgens P, Donner TH (2015) Action Planning and the Timescale of Evidence

791   Accumulation Marshall JAR, ed. PLOS ONE 10:e0129473.

792   Urai AE, Braun A, Donner TH (2017) Pupil-linked arousal is driven by decision uncertainty and alters serial

793   choice bias. Nat Commun 8:14637.

794   van den Brink RL, Wynn SC, Nieuwenhuis S (2014) Post-Error Slowing as a Consequence of Disturbed Low-

795   Frequency Oscillatory Phase Entrainment. J Neurosci 34:11096–11105.

796   Wang X-J (2002) Probabilistic decision making by slow reverberation in cortical circuits. Neuron 36:955–968.

797   Wang X-J (2008) Decision Making in Recurrent Neuronal Circuits. Neuron 60:215–234.

798   Yu A, Cohen JD (2009) Sequential effects: Superstition or rational behavior. Adv Neural Inf Process Syst

799   21:1873–1880.

800

801   **Legends**

802   **Figure 1. Quantifying choice history bias and behavioral task.**

803   (**A**), (**B**) Behavioral tasks. Observers judged the net direction (up vs. down) of a dynamic random dot

804   pattern of variable direction and coherence. (**A**) Experiment 1, decoupling choice and motor response.

805   After a blank fixation interval, a choice-response mapping cue was shown before (Pre) or after the

806   presentation (Post) of the motion stimulus, which also varied in duration. Observers responded after

807   dot motion offset in the Pre-condition and after mapping cue offset in the Post-condition. Auditory

808   feedback was provided after incorrect responses. (**B**) Experiment 2, manipulating stimulus repetition

809   probabilities. Left: Random dot motion and fixation cross were shown throughout the trial. A beep

810   indicated the onset of the evidence interval, which contained some level of coherent motion (0% on

811   some trials). A second beep indicated the evidence offset and start of the response interval (deadline:

812   3 s). Right: Three repetition probabilities between motion directions across trials yielded three

813   environmental conditions: Neutral (repetition probability of 0.5), Repetitive (repetition probability of 0.8)

814   and Alternating (repetition probability of 0.2). (**C**) Signed motion coherence levels (cyan) and

815   categorical choices (purple) from a sequence of 15 trials recorded in Neutral in Experiment 2. Positive

816   values of stimulus intensity correspond to upward motion and negative ones to downward motion. (**D**)

817   Psychometric functions conditioned on previous choice in Neutral exhibit history biases in three

818   example participants. See main text for details.

819

820

821

822   **Figure 2. Stronger impact of previous choice than of previous motor response on current bias.**

823   **(A)** Impact of previous choices and motor responses as a function of lag. **(B)** As (A), but for impact of

824   previous stimulus categories. Shaded areas, s.e.m.; dots, p < 0.05 (FDR-corrected t-test) across

825   participants.

826

827   **Figure 3. Adjustment of choice history biases to environmental statistics.**

828   **(A)** Stimulus repetition probabilities for Repetitive, Alternating and Neutral. In Repetitive, repetition of

829   the motion direction from two trials back could occur due to a sequence of two repetitions or two

830   alternations (probability: 0.8 x 0.8 + 0.2 x 0.2 = 0.68). In Alternating, the probability of repetition of the

831   same direction oscillated around 0.5 as a function of lags, with decreasing deviation from 0.5. **(B)**

832   Impact of previous stimuli and choices on current choice for lag 1. Dots, single observers, arrows

833   changes of group mean weights from Neutral (red cross) during Repetitive and Alternating,

834   respectively. **(C)** Psychometric functions conditioned on previous choice (group average). Left:

835   Repetitive, leftward shift from dashed to dotted line corresponding to a bias to repeat the previous

836   choice. Right: Alternating, leftward shift from dotted to dashed line, indicating a bias to alternate the

837   previous choice. **(D)** Correct weights as functions of lags in Repetitive and Alternating. **(E)** Adaptivity

838   indices (correlation coefficient with the history template) computed from correct kernels from the

839   Repetitive and Alternating conditions. Dotted line, correlation between history templates for Repetitive

840   and Alternating. Shaded areas, s.e.m.; dots, $p < 0.05$ (FDR-corrected t-test) across participants; *, $p <$

841   0.05; **** $p < 0.0001$.

842

843   **Figure 4. No differences in history weights for synthetic observers without bias adjustment.**

844   Results of two simulations of synthetic observers without bias adjustment analyzed as real observer

845   data for Figure 3B, D. (**A**) Synthetic observers with all parameters taken from the real observers'

846   estimates for Repetitive and Alternating, but history weights set to 0. (**B**) Synthetic observers with all

847   parameters taken from the real observers' estimates for Neutral. Note the difference in the y-axis scale

848   between these simulated observers and the real observer data in Figure 3. Shaded areas, s.e.m.

849   across synthetic observers.

850

851

852   **Figure 5. Behavioral performance depends on bias adjustment and perceptual sensitivity.**

853   **(A)** Correlation between adaptivity index and proportion of correct choices. Left: Repetitive. Middle:

854   Alternating. Right: Neutral. Insets: Correlations for the simulated observers without bias adjustment

855   (compare Figure 4A). **(B)** Correlation between sensitivity (i.e. slope of the psychometric function) and

856   the proportion of correct choices. Left: Repetitive. Middle: Alternating. Right: Neutral.

28

857

858   **Figure 6. Modulation of bias adjustment by proxies of confidence in previous choice.**

859   **(A**) Scaling of model-based confidence and uncertainty with evidence strength on correct and error

860   trials. Adapted from (Urai et al., 2017) under a CC-BY 4.0 license. **(B)** Difference between previous

861   choice weights from Repetitive and Alternating, sorted by previous choice correctness and coherence.

862   **(C)** Comparison between previous correct and incorrect weights, for Repetitive and Alternating.

863   Weights were first calculated separately for each previous coherence level and then pooled across

864   coherence. **(D)** Reaction time as function of motion coherence sorted by correctness (pooled across

865   Repetitive and Alternating). **(E)** Difference between previous $RT conf\ x\ correct$ modulation weights

866   from Repetitive and Alternating, sorted by previous coherence. See main text for details of the

867   multiplicative modulation model. **(F)** $RT conf\ x\ correct$ weights for Repetitive and Alternating.

868   Modulation weights were first calculated separately for each previous coherence level and then pooled

869   across coherence. Shaded areas, s.e.m.; dots, p < 0.05 (FDR-corrected t-test) across participants; *, p

870   < 0.05; **, p < 0.01; ***, p < 0.001, ****, p < 0.0001.

871

A

Pre

Fixation 0.4 - 0.5s | mapping cue 1s | Delay 0.2 - 0.4s | Evidence 0.15 - 4.8s | Response

Post

Time → Fixation 0.4 - 0.5s | Evidence 0.15 - 4.8s | Delay 0.2- 0.4s | mapping cue | Response

B

up / down

Time → Fixation 1 - 5 s | Evidence 0.75 s | Response max. 3 s

Neutral 0.5 0.5 0.5 0.5

Repetitive 0.2 0.8 0.8 0.2

Alternating 0.8 0.2 0.2 0.8

C

Choice history bias, S20

Upward motion coherence | up / down Choice

Trial number

D

Repetition bias, S20 | No history bias, S09 | Alternation bias, S11

Probability up choices

Upward motion coherence

All choices
Previous choice: up
Previous choice: down

A

B

A

B

C

D

E

A

Simulated observers with individual parameters
from biased conditions and history weights set to 0
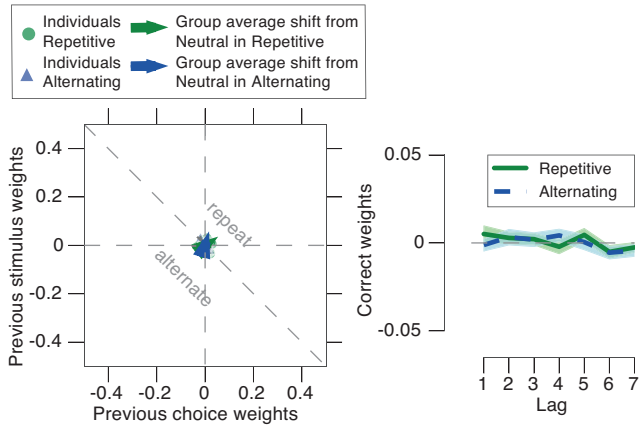
B

Simulated observers with individual parameters from Neutral

A

Repetitive

Simulated observers
without bias adjustment

r = -0.0623
p = 0.7831

0.95
0.85
0.75
0.65

-1   0   1

Alternating

Simulated observers
without bias adjustment

r = 0.2355
p = 0.2913

-1   0   1

Neutral

Simulated observers
without bias adjustment

r = 0.183
p = 0.4149

-1   0   1

Proportion correct

0.95
0.9
0.85
0.8
0.75
0.7
0.65

r = 0.6057
p = 0.0028

-1 -0.5   0   0.5   1

r = 0.6492
p = 0.0011

-1 -0.5   0   0.5   1

r = 0.3074
p = 0.1641

-1 -0.5   0   0.5   1

Adaptivity index (correlation between history kernel and template)

B

Proportion correct

0.95
0.9
0.85
0.8
0.75
0.7
0.65

r = 0.6602
p = 0.0008

7   9   11   13   15

0.95
0.9
0.85
0.8
0.75
0.7
0.65

r = 0.7177
p = 0.0002

7   9   11   13   15

0.95
0.9
0.85
0.8
0.75
0.7
0.65

r = 0.8141
p < 0.0001

7   9   11   13   15

Perceptual sensitivity α

A

Confidence

1

0.5

— Error
— Correct

Uncertainty

0.5

0

weak strong
Evidence

Uncertainty = 1 - confidence

B

Repetitive - Alternating

Difference in choice weight

1.0

0.5

0.0

−0.5

— Previous correct
— Previous error

0.05   0.1   0.2   0.4 0.6
Previous coherence

p < 0.05, corrected
● Previous correct
● Previous correct vs. previous error

C

Repetitive   Alternating

*   **   n.s.   *** **** n.s.

Weight

2
1.5
1
0.5
0
−0.5
−1
−1.5
−2

Prev.   Prev.   Prev.   Prev.
correct error correct error

D

Reaction time (s)

0.55

0.50

0.45

0.40

0.35

— Correct
— Error

0.0   0.2   0.4   0.6
Motion coherence

E

Repetitive - Alternating

Difference in RTconf x correct weight

0.5

0.4

0.3

0.2

0.1

0.0

0.05   0.1   0.2   0.4 0.6
Previous coherence

● p < 0.05, corrected

F

Repetitive Alternating

***   **

RTconf x correct weight

0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
−0.8