

---

**Research Article: New Research | Sensory and Motor Systems**

## **Strategic and dynamic temporal weighting for perceptual decisions in humans and macaques**

**Aaron J Levi<sup>1,2,3</sup>, Jacob L. Yates<sup>1,2,3,4,5</sup>, Alexander C. Huk<sup>1,2,3</sup> and Leor N. Katz<sup>1,2,3,6</sup>**

<sup>1</sup>*Department of Neuroscience, The University of Texas at Austin, Austin, Texas, USA*

<sup>2</sup>*Department of Psychology, The University of Texas at Austin, Austin, Texas, USA*

<sup>3</sup>*Center for Perceptual Systems, The University of Texas at Austin, Austin, Texas, USA*

<sup>4</sup>*Center for Visual Science, University of Rochester, Rochester, New York, USA*

<sup>5</sup>*Department of Brain and Cognitive Sciences, University of Rochester, Rochester, New York, USA*

<sup>6</sup>*Laboratory of Sensorimotor Research, National Eye Institute, National Institutes of Health, Bethesda, Maryland, USA*

DOI: 10.1523/ENEURO.0169-18.2018

Received: 30 April 2018

Revised: 8 August 2018

Accepted: 1 September 2018

Published: 11 September 2018

---

**Funding:** Howard Hughes Medical Institute International Student Research Fellowship

**Funding:** <http://doi.org/10.13039/100000053HHS> | NIH | National Eye Institute (NEI)  
R01-EY017366  
T32EY021462

**Funding:** <http://doi.org/10.13039/100000002HHS> | National Institutes of Health (NIH)  
T32DA018296

**Conflict of Interest:** Authors report no conflicts of interest.

AJL, JLY, and LNK performed research and analyzed data. AJL, JLY, ACH, and LNK designed research and wrote the paper.

This research was supported by the Howard Hughes Medical Institute International Student Research Fellowship to L.N.K., the National Eye Institute (R01-EY017366) grant to both A.C.H. and Jonathan Pillow (Princeton University), and the National Institutes of Health under Ruth L. Kirschstein National Research Service Awards T32DA018926 from the National Institute on Drug Abuse and 2T3EY021462 from the National Eye Institute.

**Correspondence should be addressed to** [alevi@utexas.edu](mailto:alevi@utexas.edu) and [leor.katz@nih.gov](mailto:leor.katz@nih.gov)

**Cite as:** eNeuro 2018; 10.1523/ENEURO.0169-18.2018

**Alerts:** Sign up at [eneuro.org/alerts](http://eneuro.org/alerts) to receive customized email alerts when the fully formatted version of this article is published.

Accepted manuscripts are peer-reviewed but have not been through the copyediting, formatting, or proofreading process.

Copyright © 2018 Levi et al.

This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license, which permits unrestricted use, distribution and reproduction in any medium provided that the original work is properly attributed.

1 1. Strategic and dynamic temporal weighting for perceptual decisions in humans and  
2 macaques

3  
4 2. Dynamic temporal weighting for decisions

5  
6 3. Aaron J Levi<sup>1-3</sup>, Jacob L. Yates<sup>1-5</sup>, Alexander C. Huk<sup>1-3</sup>, Leor N. Katz<sup>1-3, 6</sup>

7  
8 <sup>1</sup>Department of Neuroscience, The University of Texas at Austin, Austin, Texas, USA.

9 <sup>2</sup>Department of Psychology, The University of Texas at Austin, Austin, Texas,  
10 USA. <sup>3</sup>Center for Perceptual Systems, The University of Texas at Austin, Austin, Texas,

11 USA. <sup>4</sup>Center for Visual Science, University of Rochester, Rochester, New York,

12 USA. <sup>5</sup>Department of Brain and Cognitive Sciences, University of Rochester, Rochester,

13 New York, USA. <sup>6</sup>Laboratory of Sensorimotor Research, National Eye Institute, National

14 Institutes of Health, Bethesda, Maryland, USA.

15

16 4. AJL, JLY, and LNK performed research and analyzed data. AJL, JLY, ACH, and LNK  
17 designed research and wrote the paper.

18

19 5. Correspondence should be addressed to [alevi@utexas.edu](mailto:alevi@utexas.edu) and [leor.katz@nih.gov](mailto:leor.katz@nih.gov)

20

21 6. 5 figures

22 7. 1 tables

23 8. 0 multimedia

24

25 9. 212 words in Abstract

26 10. 106 words in Significance Statement

27 11. 750 words in Introduction

28 12. 2,482 words in the Discussion

29

30 13. We thank other members of the Huk lab for discussion surrounding the research and  
31 K. Mitchell, C. Carter, and A. Laudano for animal support. This research was supported  
32 by the Howard Hughes Medical Institute International Student Research Fellowship to  
33 L.N.K., the National Eye Institute (R01-EY017366) grant to both J.W.P, and A.C.H., and  
34 the National Institutes of Health under Ruth L. Kirschstein National Research Service  
35 Awards T32DA018926 from the National Institute on Drug Abuse and T32EY021462  
36 from the National Eye Institute.

37

38 14. Authors report no conflicts of interest.

39

40 15. This research was supported by the Howard Hughes Medical Institute International  
41 Student Research Fellowship to L.N.K., the National Eye Institute (R01-EY017366) grant  
42 to both A.C.H. and Jonathan Pillow (Princeton University), and the National Institutes of  
43 Health under Ruth L. Kirschstein National Research Service Awards T32DA018926 from  
44 the National Institute on Drug Abuse and 2T3EY021462 from the National Eye Institute.

45

46 **ABSTRACT**

47

48 Perceptual decision-making is often modeled as the accumulation of sensory evidence  
49 over time. Recent studies using psychophysical reverse correlation have shown that  
50 even though the sensory evidence is stationary over time, subjects may exhibit a time-  
51 varying weighting strategy, weighting some stimulus epochs more heavily than others.  
52 While previous work has explained time-varying weighting as a consequence of static  
53 decision mechanisms (e.g., decision bound or leak), here we show that time-varying  
54 weighting can reflect strategic adaptation to stimulus statistics, and thus can readily take  
55 a number of forms. We characterized the temporal weighting strategies of humans and  
56 macaques performing a motion discrimination task in which the amount of information  
57 carried by motion stimulus was manipulated over time. Both species could adapt their  
58 temporal weighting strategy to match the time-varying statistics of the sensory stimulus.  
59 When early stimulus epochs had higher mean motion strength than late, subjects  
60 adopted a pronounced early weighting strategy, where early information was weighted  
61 more heavily in guiding perceptual decisions. When the mean motion strength was  
62 greater in later stimulus epochs, in contrast, subjects shifted to a marked late-weighting  
63 strategy. These results demonstrate that perceptual decisions involve a temporally  
64 flexible weighting process in both humans and monkeys, and introduce a paradigm with  
65 which to manipulate sensory stimuli in decision-making tasks.

66

67 **SIGNIFICANCE STATEMENT**

68 During decision-making, the weight assigned by subjects to sensory information over  
69 time is not necessarily constant. Such time-varying weighting is often interpreted as a  
70 signature of a particular decision-making model (e.g., higher weighting of early stimulus  
71 information is consistent with a bounded accumulation process). Temporal weighting  
72 may also result, however, from a strategic reweighting of the stimulus evidence itself that  
73 takes place before and/or independent of a decision-making mechanism. Here we use a  
74 psychophysical reverse correlation paradigm to both measure and manipulate temporal  
75 weighting behavior. We demonstrate that both humans and macaques adopt weighting  
76 strategies that are flexible, consistent with dynamic reweighting of the sensory stimulus.

77

78 **INTRODUCTION**

79 Perceptual decisions are typically thought of as resulting from some form of  
80 accumulating samples of a stimulus over time. During this process, a decision variable is  
81 updated as evidence is integrated until a choice is made. In both human and nonhuman  
82 primates, perceptual decision-making has been studied extensively in the context of  
83 motion direction discrimination tasks, where the vast majority of stimuli provide  
84 statistically uniform sensory evidence over time (Gold and Shadlen, 2007). Despite a  
85 stationary level of expected sensory evidence, subjects often assign more weight to  
86 some stimulus epochs over others. In many instances, subjects have exhibited “early  
87 weighting”, where sensory evidence presented in early epochs contributes more to  
88 choices than that in late (Huk and Shadlen, 2005; Kiani et al. 2008, Nienborg and  
89 Cumming, 2009; Yates et al., 2017). In other instances, however, “late weighting” has  
90 been observed, where choices were primarily influenced by sensory evidence presented  
91 in late stimulus epochs (Tsetsos et al., 2012; Cheadle et al., 2014; Bronfman et al.,  
92 2016; Carland et al., 2016). In rodents, a mixture of either early or flat weighting profiles  
93 has been reported (Erich et al., 2015; Scott et al., 2015; Pinto et al., 2017; Licata et al.,  
94 2017).

95         The diverse set of temporal weighting profiles observed across studies and  
96 species may be explained in a number of ways. One approach appeals to mechanistic  
97 models of decision-making. An early weighting strategy, for example, could be explained  
98 as a consequence of bounded accumulation (Huk and Shadlen, 2005; Kiani et al., 2008),  
99 which posits that sensory evidence is accumulated until reaching a bound, whereupon  
100 the decision is made. Because the remainder of the stimulus is ignored once the bound  
101 has been hit, early stimulus epochs contribute more to decisions than late. Late  
102 weighting, in contrast, may be interpreted as a consequence of leaky accumulation

103 (Usher and McClelland, 2001), which stipulates that the representation of sensory  
104 evidence decays over time. In this model, early sensory evidence contributes less to  
105 decisions compared to late.

106         An alternative approach to explaining the variety in weighing strategies  
107 postulates that the temporal weighting strategy is flexible, and linked to the demands or  
108 structure of the task. This notion is supported by experiments in which weighting  
109 changes systematically with variable trial length and signal timings (Ghose 2006;  
110 Tstetsos et al., 2012; Ossmy et al., 2013; Bronfman et al. 2016), as well as by studies  
111 that explore effects of congruency between serially presented samples (Cheadle et al.,  
112 2014). Irrespective of a stipulated model or mechanism, these studies point to similar  
113 conclusions: subjects may reweigh stimulus information as dictated by the reliability of  
114 the evidence and demands of the task.

115         Without appeal to a specific decision-making mechanism, we set out to  
116 manipulate temporal weighting under the hypothesis that weights should be flexible and  
117 influenced by the dynamic features of the stimulus itself, either independent of or in  
118 addition to constraints imposed by integration mechanisms such as a bound or a leak.  
119 To test this idea, we adopted a motion stimulus designed explicitly for psychophysical  
120 reverse correlation in the presence of experimenter-controlled manipulation of temporal  
121 stimulus statistics (Katz et al. 2016, Yates et al. 2017). The stimulus is similar to classic  
122 motion stimuli used in the study of perceptual decisions (Newsome & Pare, 1988; Britten  
123 et al., 1992), but with two crucial features: (i) the stimulus consists of seven consecutive  
124 motion pulses, each with a predetermined mean motion strength and direction, and thus  
125 can be precisely designed to carry more or less motion evidence at different epochs  
126 (Figure 1); (ii) the stimulus is amenable to psychophysical reverse correlation analysis  
127 such that subject temporal weighting strategy may be computed directly. This motion  
128 discrimination task was performed under three temporal conditions: (1) “Flat-stimulus”, in

129 which the mean motion strength per pulse was constant; (2) “Early-stimulus”, in which  
130 early pulses had high mean motion strength and late pulses had low; and (3) “Late-  
131 stimulus”, in which late pulses had high mean motion strength and early pulses had low  
132 (Figure 2A-C). In all conditions, the task was to report the net motion of the trial.

133         We found that in both time-varied conditions (early-stimulus and late-stimulus),  
134 subjects shifted their temporal weighting strategy, placing highest weight on motion  
135 pulses with the highest mean motion strength. In flat-stimulus sessions, however,  
136 subjects exhibited a large range of temporal weighting strategies despite equal mean  
137 motion strength over time. Overall, these results demonstrate that temporal weighting  
138 strategies in human and monkey observers are flexible, and can be adjusted to suit  
139 temporal stimulus statistics.

140

141 **MATERIALS AND METHODS**

142 **Subjects and apparatus**

143 Data were collected from both monkeys and humans. Monkey data were collected from  
144 two adult rhesus macaques (one female and one male, referred to as M1 and M2  
145 hereafter) aged 10 and 14, weighing 7.7 and 10kg, respectively. All animal procedures  
146 were performed in accordance with the [Author University] animal care committee's  
147 regulations. Both M1 and M2 had standard surgery for implantation of a head-holder.  
148 Some portion of the monkey data was presented previously (Anonymous 2016,  
149 Anonymous 2017). Human data were collected from three subjects (all males, referred  
150 to as H1, H2, H3), aged 23-41 years, all with normal or corrected-to-normal vision.  
151 Experiments were performed with the written consent of each observer and all  
152 procedures were approved by [Author University] review board.

153 For both monkeys and humans, stimuli were presented using the Psychophysics  
154 Toolbox with Matlab (The Mathworks) using a Datapixx I/O box (Vpixx) for precise  
155 temporal registration (Eastman and Huk, 2012). Sample stimulus presentation code is  
156 available upon request. Eye position was tracked using an Eyelink eye tracker (SR  
157 Research), sampled at 1kHz. Monkeys sat in a primate chair (Crist Instruments) and  
158 viewed stimuli on a 55 inch LCD (LG) display (resolution = 1920 x 1080p, refresh rate =  
159 60Hz, background luminance = 26.49 cd/m<sup>2</sup>) that was corrected to have a linear gamma  
160 function. Monkeys viewed the stimulus from a distance of 118cm (such that the screen  
161 width subtended 54 degrees of visual angle, and each pixel subtended 0.0282 degrees  
162 of visual angle). Auditory feedback was played at the end of every trial, and fluid reward  
163 was delivered through a computer-controlled solenoid. Humans viewed stimuli on a  
164 linearized 16.5 inch OLED (LG) display (resolution = 1920 x 1080p, refresh rate = 60Hz,  
165 background luminance = 67.22 cd/m<sup>2</sup>) at a distance of 65.3cm (such that screen width



166 subtended 31 degrees of visual angle, and each pixel subtended 0.0163 degrees of  
167 visual angle).

168

### 169 **Task and stimulus design**

170 Stimulus and task design were identical between monkeys and humans unless  
171 otherwise noted. Subjects were required to discriminate the net direction of a motion  
172 stimulus and communicate their decision with an eye movement to one of two targets,  
173 placed on either side of the motion stimulus. The sequence of task events is presented  
174 in Figure 1. A trial began with the appearance of a fixation point. Once the subject  
175 acquired fixation and held for 400-1200ms (uniform distribution), two targets appeared  
176 and remained visible until the end of the trial. 200-1000ms after target onset, the motion  
177 stimulus was presented at a range of eccentricities from 4-10° for a duration of 1050ms.  
178 The fixation point was extinguished 200-1000ms after motion offset, and the subject was  
179 then required to shift their gaze towards one of the two targets within 600ms (saccade  
180 end points within 3° of the target location were accepted). The timing of each event was  
181 randomly and independently jittered from trial to trial (Figure 1A).

182         The reverse-correlation motion stimulus contained motion towards one direction  
183 or the opposite, with varying motion strength. Spatially, the stimulus consisted of a  
184 hexagonal grid of 19 Gabor elements, 5-7° across, scaled by eccentricity (Figure 1B).  
185 Individual Gabor elements were set to approximate the receptive field (RF) size of a V1  
186 neuron, and the entire motion stimulus approximated the RF size of an MT neuron (Van  
187 Essen et al., 2002). Motion was presented by varying the phase of the sine-wave carrier  
188 of the Gabors. Each Gabor underwent a sinusoidal contrast modulation over time with  
189 independent random phase to prevent perceptual “pop-out” of individual drifting  
190 elements. Gabor spatial frequency (0.8 cycles/°, sigma = 0.1 x eccentricity) and temporal

191 frequency 5-7Hz, yielding velocities of 5.55-7.77 °/s, respectively) were selected to  
192 match the approximate sensitivity of MT neurons (Bair and Movshon, 2004).

193 Each motion stimulus presentation consisted of seven consecutive motion pulses  
194 lasting 150ms each (9 frames), producing a motion sequence of 1050ms in duration in  
195 total. For human subjects S2 and S3, each motion pulse lasted 100ms each (6 frames),  
196 producing a 700ms-long stimulus. On any given pulse, a number of Gabor elements  
197 would have their carrier sine waves drift in unison to produce motion (“signal elements”),  
198 and the remaining would counter-phase flicker (“noise elements”). Signal elements on  
199 any given pulse were assigned at random within the grid and all signal element drifted in  
200 the same direction. Motion strength on pulse  $i$  was defined as the proportion of signal  
201 elements out of the total number of elements, the value of which was drawn from a  
202 Gaussian distribution,  $X_i \sim N(\mu_k, \sigma)$  and rounded to the nearest integer, where  $k$  is the  
203 distribution index for the five trial types (strong left, weak left, noise, weak right, strong  
204 right) and  $\mu_k$  was one of five values: -50%, -10%, 0%, 10%, and 50% (sign indicates  
205 motion in the opposite direction), and  $\sigma$  was set to 15%. Thus, while each pulse within a  
206 sequence could take on any value (and either sign/direction) from distribution  $N(\mu_k, \sigma)$ ,  
207 the expectation of a sequence would be  $\mu_k$  (Figure 1). The subjects were rewarded for  
208 selecting the target consistent with the sign of the motion pulse sequence sum (i.e., the  
209 net direction), independent of the distribution  $\mu_k$  from which the pulses were drawn.

210 The distributions  $N(\mu_k, \sigma)$  were most commonly set to the values listed above but  
211 were occasionally varied to better maintain individual subject performance around  
212 threshold. Overall, humans performed sessions with  $\mu_{strong}$  ranging from 35-50% and  
213  $\mu_{weak}$  ranging from 10-20%, with  $\sigma$  ranging from 10-24% coherence. Macaques  
214 performed sessions with  $\mu_{strong}$  ranging from 50-70% and  $\mu_{weak}$  ranging from 10-20%, with  
215  $\sigma$  ranging from 8-24% coherence.

216

217

218 **Temporal manipulation of stimulus**

219 In the standard stimulus design described above, the mean of the motion strength  
220 distribution  $N(\mu_k, \sigma)$  would be held constant throughout a stimulus presentation. In other  
221 words, the mean of the distribution from which  $X_i$  was drawn was fixed at  $\mu_k$ , for pulses 1  
222 to 7 (Figure 2A). We refer to this as the “flat-stimulus” condition and treat it as a  
223 baseline, because it is similar to most variants of the classic moving dot stimuli used in  
224 the past (Newsome & Pare, 1988; Britten et al., 1992; Britten et al., 1996; Gold and  
225 Shadlen 2007). In the time-varying stimulus conditions (the early-stimulus or late-  
226 stimulus),  $\mu_k$  was varied over pulses 1 to 7. Figure 2B depicts a stimulus condition in  
227 which motion strength is reduced substantially in early pulses (relative to baseline  
228 levels), but not late. In this “late-stimulus” condition,  $\mu_k$  is set to 0 for the first pulse ( $i=1$ ),  
229 and reaches its expected value ( $\mu_k$ ) by pulse 7. The transition from 0 at pulse 1 to  $\mu_k$  at  
230 pulse 7 is governed by a logistic function with parameters chosen to result in a smooth  
231 transition between the first 3 and last 3 pulses (midpoint=4, slope=0.3). Although  $\mu_k$  is  
232 near zero for the early pulses,  $\sigma$  is unchanged such that even though the expectation for  
233 motion on pulse one is zero, the motion strength and direction will vary from trial to trial  
234 (see example trials in Figure 2B). In other words, random draws of  $X_i$  from distribution  
235  $N(\mu_k, \sigma)$  where  $\mu_k=0$  still carry motion information, albeit less correlated with the net  
236 motion outcome of the trial as a whole. The opposite is done for the “early-stimulus”  
237 condition (Figure 2C), in which the first pulses maintain mean motion strength equal to  
238  $\mu_k$ , and later pulses have a mean near zero. This stimulus design ensures that pulse  
239 sequences drawn from the  $\mu_k = 0$  Gaussian (i.e. “zero-mean trials”) maintain a 0 mean  
240 throughout all 7 pulses, regardless to whether the stimulus condition is flat, early, or late.  
241 These trials were difficult because the motion strength and direction of each pulse is  
242 small and independent of the sequence, and the net motion summed to a small

243 directional outcome. About one quarter of macaque sessions also contained frozen seed  
244 trials, in which an identical stimulus was displayed for 5-10% of trials. These trials  
245 summed to exactly zero and the subject was rewarded at random.

246 All subjects began the experiments with the flat-stimulus condition. After multiple  
247 sessions of stable psychophysical performance within a condition, the stimulus was  
248 changed to either the late- or early-stimulus conditions. Finally, after multiple sessions of  
249 stable psychophysical performance under the second condition, they began the third and  
250 final condition. Subjects were only exposed to one stimulus condition per session and  
251 were not informed of which stimulus condition they were viewing before or during any  
252 given session.

253

#### 254 **Data analysis**

255 Sessions with a minimum of 250 successfully completed trials were included in data  
256 analysis. Sessions were excluded from analysis if subject accuracy was lower than 85%  
257 for the strongest motion values (17 / 235 sessions for macaques, 0 / 52 for humans).  
258 Additionally, 30 macaque sessions were excluded from analysis for having  
259 psychophysical thresholds greater than 2 median absolute deviations about the median.  
260 Overall, 188 and 52 sessions were included for macaques and humans, respectively,  
261 with median session lengths of 295 and 632 successfully completed trials, netting a total  
262 of 129,922 and 15,275 trials overall.

263 All analyses were performed in Matlab (The Mathworks). Subject choices in the  
264 direction-discrimination task were analyzed with a maximum likelihood fit of a three-  
265 parameter logistic function (Wichmann and Hill 2001) assuming a Bernoulli distribution of  
266 binary choices, in which the probability of a rightward choice is  $P$  and leftward choice is  
267  $1-P$ , where  $P$  is given by:

268 (1) 
$$p = \gamma + (1 - 2\gamma)\left(\frac{1}{1 + e^{-\beta(x-\alpha)}}\right)$$

269 where  $x$  is the net motion strength value (z-scored over all sessions for each subject  
 270 separately),  $\alpha$  is the bias parameter (reflecting the midpoint of the function in units of  
 271 motion strength),  $\beta$  is the slope (i.e., sensitivity, in units of log-odds per motion strength),  
 272 and  $\gamma$  captures the lapse rate as the offset from the 0 and 1 bounds. Error estimates on  
 273 the parameters were obtained from the square root of the diagonal of the inverse  
 274 Hessian (2<sup>nd</sup> derivative matrix) of the negative log-likelihood.

275 The temporal weighting kernel (which we also refer to as “temporal weighting  
 276 strategy” or “temporal weighting profile”) was computed using ridge regression via  
 277 maximum likelihood. The log posterior of the psychophysical weights is given by:

278

279 (2) 
$$\mathcal{L}(\mathbf{w}) = \sum_{i=1}^N [Y(i)\mathbf{w}^T X(i) - \log(1 + \exp(\mathbf{w}^T X(i)))] + \lambda \|\mathbf{w}\|^2$$

280

281 where  $\mathbf{Y} \in \{0, 1\}$  is a vector of choice on every trial and  $\mathbf{X}$  is a matrix of the seven pulses  
 282 on each trial, augmented by a column of ones (to capture bias).  $\lambda$  was estimated using  
 283 evidence optimization (Sahani and Linden, 2003). Psychophysical weights are  
 284 normalized by the Euclidean norm of the vector of weights. The seven temporal weights  
 285 assigned to the seven motion pulses,  $w$ , were computed by using all trials within a  
 286 session. These include trials where  $\mu_k$  was set to zero (i.e. “zero-mean trials”, where  
 287 motion on a given pulse is temporally independent of all other pulses in the sequence)  
 288 and trials where  $\mu_k$  was set to a non-zero value (“signal trials”, where motion is  
 289 correlated over pulses). Psychophysical reverse correlation is traditionally performed on  
 290 noise trials exclusively, but logistic regression effectively whitens the stimulus  
 291 covariance, such that we could include all trials and increase our statistical power,

292 regardless to whether they have correlated temporal structure. We verified the whitening  
 293 step by comparing the psychophysical kernel computed on all trials to the kernel  
 294 computed on only zero-mean trials and calculating the Pearson correlation between the  
 295 pair of kernels (i.e. between the 7 weights of the 'all-trials-kernel' and the 7 weights of  
 296 the 'zero-mean-kernel') for each combination of subject and stimulus condition. This  
 297 yielded 14 Pearson correlation values with a median of 0.886 ([0.819 0.952], 1 SEM)  
 298 demonstrating a strong agreement between results of the two methods of reverse  
 299 correlation for the subject-averaged data per condition. We also verified the whitening  
 300 step at the level of individual sessions, using the same approach. This yielded 240  
 301 Pearson correlation values (one for each session) with a median of 0.846 ([0.829 0.864],  
 302 1 SEM), indicating a strong agreement between reverse correlation methods, even on  
 303 single sessions.

304 The vector of weights,  $w$ , describes the temporal weighting adopted by the  
 305 subject for a given set of trials. If the individual weights have a similar value, then that  
 306 implies that the subject had weighted all pulses equally on average. If some weights are  
 307 larger than others, that implies uneven weighting over time. We summarized temporal  
 308 weighting by performing linear regression on the 7 weights and using the slope of the fit  
 309 as a metric of temporal structure, where negative slopes indicate early psychophysical  
 310 weighting and positive slopes indicate late. Comparisons of temporal weighting profiles  
 311 across experimental conditions and species were assessed using the slope of the linear  
 312 fit plus or minus 95% confidence intervals. Wilcoxon sign tests were used to evaluate  
 313 whether slopes differed significantly from zero. ANOVA was used to assess differences  
 314 in mean slope across experimental conditions. Bartlett's test was used to evaluate  
 315 differences in variance between distributions of slopes across experimental conditions.

Test use	Test	Data structure	Power
----------	------	----------------	-------

Psychophysical weighting calculated on all trials vs. only zero-mean trials	Pearson correlation	Linear	Subjects: Median $r = 0.886$ [0.819 0.952], 1 SEM  Single session: Median $r = 0.846$ , [0.829 0.864], 1SEM
Differences in slope of linear fit to temporal weights between flat-, late-, and early-stimulus conditions in humans and macaques (Figure 3).	Confidence intervals	Linear	Slope of linear fit, [95% confidence interval]  Macaques Flat: -0.050, [-0.069-0.031]; Late: 0.051, [0.004 0.098]; Early: -0.094, [-0.111 -0.077]  Humans: Flat: -0.013, [-0.032 0.006]; Late: 0.053, [0.006 0.100]; Early: -0.083, [-0.119 -0.048];
Comparison of slope of linear fit to temporal weights during early-stimulus condition between humans and macaque subjects (Figure 3)	Confidence intervals	Linear	Slope of linear fit, [95% confidence interval]  Humans: -0.013, [-0.032 0.006] Macaques -0.050, [-0.069 -0.031]
Comparison of psychometric functions across conditions (Figure 3).	Confidence intervals	Linear	Slope of psychometric function, [95% confidence interval]  Macaques Early: 3.39 [3.22 3.56], Flat: 2.16 [2.13 2.18], Late: 2.9339, [2.83 3.03]  Humans Early: 2.77 [2.56 2.99], Flat: 2.14 [2.00 2.28], Late: 2.60 [2.43 2.77]
Average slope of temporal weights for flat-, early-, and late-stimulus conditions compared to 0 (Figure 5).	Wilcoxon sign test	Non-Gaussian	$p < 0.0001$ , all conditions
Comparing group means for slopes of temporal weights for flat-, early-, and late-stimulus conditions. (Figure 5).	ANOVA	Gaussian	$p < 0.0001$

Comparing variance of slopes during flat stimulus condition vs. late- and early-stimuli (Figure 5).	Bartlett's test	Non-Gaussian	Flat-to-early $p < 0.0001$  Flat-to-late, $p < 0.0001$
Evaluating linear relationship between psychophysical threshold and slope of temporal weights (Figure 5).	Pearson correlation	Linear	Flat: $r = -0.29, p < 0.001$  Early: $r = 0.46, p = 0.038$  Late: $r = 0.05, p = 0.75$
Evaluating linear relationship between psychophysical threshold and energy of temporal weights (Figure 5).	Pearson correlation	Linear	Flat: $r = 0.40, p < 0.0001$  Early: $r = -0.004, p = 0.99$  Late: $r = 0.31, p = 0.048$

**Table 1. Statistical table.** Descriptions and results of statistical tests performed on results.

316



317 **RESULTS**

318 Overall, subjects performed over 145,000 trials of a one-interval motion direction  
319 discrimination task. After viewing a sequence of motion pulses, they indicated the net  
320 perceived direction by moving their eyes to one of two targets (Figure 1). In addition to  
321 the usual practice of varying the net strength and direction of motion across trials, the  
322 temporal statistics of the motion stimulus were manipulated within trials (in different  
323 series of sessions). Thus, sessions varied in whether the motion stimulus offered an  
324 equal amount of motion information over time (flat-stimulus condition), or whether some  
325 epochs contained more motion information than others (early-stimulus and late-stimulus  
326 conditions) (Figure 2A-C). This design is amenable to psychophysical reverse correlation  
327 such that in addition to computing standard subject performance as a function of  
328 stimulus strength, we calculated the psychophysical weights assigned by the subject to  
329 the motion stimulus over each epoch. We refer to the resulting weights as the temporal  
330 weighting strategy or temporal weighting profile. We found that both human and monkey  
331 observers shifted their temporal weighting profile in response to the differential temporal  
332 structure of motion statistics across the three stimulus conditions. We first present our  
333 subject-averaged results, followed by an examination of the differences between species  
334 and individual subjects.

335

336 **Temporal weighting strategies shift in response to stimulus statistics**

337 Changes in temporal stimulus statistics led to clear shifts in the psychophysical  
338 weighting strategy in all subjects. We consider the flat-stimulus condition as a baseline,  
339 both because of the stationary statistics of the stimulus over time, and because the vast  
340 majority of stimuli used in the study of perceptual decision-making have temporally  
341 stationary statistics. In the flat-stimulus condition, subjects exhibited an inclination

342 towards early weighting, with the highest weight on the first three pulses and then a  
343 steady decrease as time went on (Figure 2D). The temporal weighting measurements  
344 were complimented by a standard analysis of subject psychometric performance. These  
345 indicate that observers were well engaged in the task and based their choices on the net  
346 strength and direction of the motion stimulus (Figure 2G).

347         During late-stimulus sessions, subjects shifted their strategy to place higher  
348 weight on the later pulses, which more often carried high motion information and were  
349 therefore more reliably correlated with the final trial outcome. Temporal weights in the  
350 late-stimulus condition started low, increasing to a peak at the fifth or sixth motion pulse,  
351 followed by a decreased weight on the seventh (final) pulse (Figure 2E). Even though  
352 the late-stimulus condition had less motion information in early pulses, and  
353 consequently, less motion information overall compared to the flat-stimulus condition,  
354 subjects still exhibited standard psychometric performance, basing their choices on the  
355 net motion strength and direction (Figure 2H).

356         In sharp contrast to the late-stimulus sessions, during early-stimulus sessions,  
357 subjects showed steep early weighting, where the first three pulses were weighted the  
358 highest followed by a large decrease (Figure 2F). As with the late-stimulus condition,  
359 even though the temporal weighting profile shifted markedly, both species exhibited  
360 standard psychometric performance (Figure 2I).

361         The differences in temporal weighting strategies as a function of stimulus  
362 condition were robust, and consistent across species (Figure 3). Temporal weighting in  
363 the late-stimulus condition was significantly different than the weighting in the baseline  
364 flat-stimulus condition in macaques (Figure 3A, Flat: -0.050, [-0.069 -0.031]. Late: 0.051,  
365 [0.004 0.098]; slope of linear fit, [95% confidence intervals]) and in humans (Figure 3B,  
366 Flat: -0.013, [-0.032 0.006]; Late: 0.053, [0.006 0.100]; slope of linear fit, [95%  
367 confidence intervals]). Temporal weighting in the early-stimulus condition was

368 also significantly different than the weighting in the flat-stimulus condition for humans  
369 (Figure 3B, Flat: -0.013, [-0.032 0.006]. Early: -0.083, [-0.119 -0.048]; slope of linear fit  
370 [95% confidence intervals]), and in the monkey who performed the early-stimulus  
371 condition, M1 (Figure 3A, Flat -0.050, [-0.069-0.031]. Early: -0.094, [-0.111 -0.077], slope  
372 of linear fit [95% confidence intervals]), even though M1's weighting strategy for the flat-  
373 stimulus condition was very early to begin with. Such early weighting for a flat-stimulus  
374 condition has been observed in various forms in previous reports (Huk and Shadlen,  
375 2005; Kiani et al., 2008; Nienborg and Cumming, 2009; Katz et al., 2016; Yates et al.,  
376 2017; Odoemene et al., 2017). The difference in temporal weighting between the early-  
377 and late-stimulus conditions was highly significant in both species (Humans, Early:-  
378 0.083, [-0.119 -0.048]. Late: 0.053, [0.006 0.100]; slope of linear [95% confidence  
379 intervals]. Macaques, Early: -0.094 [-0.111 -0.077]. Late: 0.051, [0.004 0.098]; slope of  
380 linear fit [95% confidence intervals]).

381 In addition, no differences in temporal weighting strategy were observed between  
382 species within either the early- or late-stimulus conditions. In the flat-stimulus condition,  
383 in contrast, macaques exhibited an early weighting that was substantially steeper than  
384 that exhibited by the human observers (Figure 3A and B, blue curve. Humans: -0.013, [-  
385 0.032 0.006]. Macaques -0.050, [-0.069 -0.031] slope of linear fit [95% confidence  
386 intervals])

387 Lastly, the species-averaged psychometric functions exhibit a standard sigmoidal  
388 relationship between motion strength and choices in all stimulus conditions,  
389 demonstrating that subjects were properly engaged in the task. In the flat-stimulus  
390 condition, however, psychophysical performance was slightly decreased relative to  
391 performance in the early- and late-stimulus conditions, in both macaques (Figure 3C.  
392 Early: 3.39 [3.22 3.56], Flat: 2.16 [2.13 2.18], Late: 2.93, [2.83 3.03]; slope of  
393 psychometric function and [95% confidence intervals]) and humans (Figure 3D. Early:

394 2.77 [2.56 2.99], Flat: 2.14 [2.00 2.28], Late: 2.60 [2.43 2.77]; slope of psychometric  
395 function and [95% confidence intervals]).

396 In summary, observers performing perceptual decisions shifted their temporal  
397 weighting strategy dynamically, and placed the most value on pulses with the highest  
398 motion expectation, whenever they were located in time.

399

#### 400 **Ruling out extrema detection as a behavioral strategy**

401 In all experiments, every trial was rewarded based on the true net direction of  
402 motion presented across the seven pulses, regardless of the underlying, generating  
403 distribution. Thus, integration of the motion information over all pulses would be ideal to  
404 maximize accuracy and reward. However, the possibility exists that subjects were not  
405 performing conventional temporal integration. For example, subjects could base their  
406 decisions on the strongest motion pulse within a trial as opposed to incorporating  
407 information from all pulses. Our stimulus design enabled us to perform a post-hoc  
408 analysis to test whether subjects were performing this strategy of extrema detection  
409 (Figure 3E).

410 We selected trials in which the direction of the strongest motion pulse (i.e. the  
411 pulse with the largest number of signal-carrying Gabor elements) was in conflict with the  
412 net direction of motion of the full trial (termed “inconsistent trials”). Most choices in these  
413 trials were in favor of the net direction of motion, as opposed to the direction of the  
414 extreme single pulse, in both human and macaque subjects (Figure 3E). We then  
415 compared these inconsistent trials to trials that were matched for difficulty, but in which  
416 the direction of the strongest pulse was in the same direction as the trial’s net direction  
417 (termed “consistent trials”). If subjects were performing extrema detection, then  
418 performance should be worse on inconsistent trials (where the strongest pulse was in  
419 the opposite direction of the net) compared to consistent trials. In contrast to this idea,

420 no subject performed significantly worse on inconsistent trials, demonstrating that  
421 extreme pulse strengths did not influence subject choices non-linearly in their favor,  
422 ruling the extrema detection strategy as unlikely in this task.

423

#### 424 **Variability in temporal weighting strategy depends on stimulus condition**

425 When averaged across sessions and subjects, temporal weighting profiles tell a fairly  
426 straightforward story: subjects adopt a late weighting strategy for the late-stimulus, an  
427 early weighting strategy for the early-stimulus, and a flat-to-early weighting strategy for  
428 the flat-stimulus. Here we sought to quantify the weighing strategy at a higher resolution  
429 by looking at performance for individual subjects and sessions.

430         When each subject is considered individually, results were largely consistent with  
431 the average weighting profiles reported above. In the late-stimulus condition, human and  
432 macaque subjects' weighting was extremely similar (Figure 4A). All observers exhibited  
433 a single-humped psychophysical weighting profile in which peak weight was at pulse five  
434 or six, before a drop off on pulse seven. Even the unexpected drop in weighting of the  
435 last pulse was shared. In the early stimulus-condition (Figure 4B), subject M1 and  
436 subject H2 exhibited fairly linear early weighting patterns, and the remaining two human  
437 subjects showed slightly higher weights on the second pulses rather than the first,  
438 though still globally consistent with early weighting. Individual performance in the flat-  
439 stimulus condition (Figure 4C), however, was more variable than in the late and early  
440 conditions. In monkey subjects, M1 showed very strong early weighting, while M2  
441 exhibited u-shaped weights. Human subjects deployed generally flat weights on  
442 average, but did so in idiosyncratic ways compared to the very stereotyped strategies of  
443 the early and late conditions. On average, each subject changed their temporal  
444 weighting as dictated by early- and late-stimulus conditions compared to the flat-stimulus

445 condition (Figure 4D). Overall, temporal weighing strategies adopted in the flat-stimulus  
446 condition were more variable than those adopted in the early- or late-stimulus conditions  
447 at the level of individual subjects.

448         When each session is considered individually, variability in temporal weighing  
449 strategy is evident both between and within each of three stimulus conditions. To  
450 quantify the degree of early vs. late single-session weighing, we fit a line to the seven  
451 temporal weights of the observer for each session, and used the slope of this fit to  
452 summarize the temporal weighing profile: A positive slope indicates late weighting, a  
453 negative slope indicates early weighting, and a slope around zero indicates flat (or  
454 equal) weighting over time. The distribution of weighting slopes for all experimental  
455 sessions in the early-stimulus condition had an average of -0.079 (significantly less than  
456 zero, Wilcoxon sign test,  $p < 0.0001$ ), with no single individual sessions having a slope  
457 greater than zero (Figure 5A). The average slope for all late-stimulus sessions was  
458 0.051 (significantly greater than zero, Wilcoxon sign test,  $p < 0.0001$ ), with only 2 of 42  
459 sessions having a slope less than zero. These distributions of weighting slopes reveal  
460 distinct populations across conditions (ANOVA,  $p < 0.0001$ ), indicating that even at the  
461 resolution of single sessions, distinct strategies were adopted during the early- and late-  
462 stimulus conditions. The distribution of weighting slopes from the flat-stimulus condition  
463 had a mean of -0.0356, denoting slight early weighting (significantly less than zero,  
464 Wilcoxon sign test,  $p < 0.0001$ ), but also differed in that it had a considerably larger  
465 range of results. The standard deviation of flat-stimulus weighting slopes was more than  
466 double that of the early- or late-stimulus weighting slope distributions (Bartlett's test, flat-  
467 to-early,  $p < 0.0001$ ; flat-to-late,  $p < 0.0001$ ), indicating that subjects adopted a larger  
468 variety of temporal weighing strategies in this condition. It is worth noting that some of  
469 the variance in all three of the distributions comes from noise inherent to fitting a two-  
470 parameter linear model to the seven weights that constitute the temporal weighing

471 strategy; nevertheless, the difference in distribution widths is substantial and therefore  
472 likely meaningful.

473

#### 474 **Relationship between temporal weighting and psychometric performance**

475 We next sought to examine the relationship between temporal weighting  
476 strategies and psychometric performance in the direction discrimination task. We  
477 compared the slope of the temporal weights to psychophysical threshold (i.e. the motion  
478 strength at which subject performed at 75% correct) for each stimulus condition (Figure  
479 5B). During the flat-stimulus condition, a negative correlation was present  
480 ( $r = -0.29$ ,  $p < 0.001$ ), indicating that adopting an early weighting strategy is detrimental  
481 to psychophysical performance. The early-stimulus sessions exhibited a positive  
482 correlation between temporal weighting slope and psychophysical threshold ( $r = 0.46$ ,  $p =$   
483  $0.038$ ), indicating that in the early-stimulus condition, an early weighting strategy is  
484 preferable. Little to no correlation was observed in the late-stimulus sessions ( $r = 0.05$ ,  $p$   
485  $= 0.75$ ).

486 Perhaps more compelling was the relationship between psychophysical threshold  
487 and the energy of the temporal weights, where energy was measured as the sum of the  
488 squared residuals of each weight from the mean of the seven weights (Figure 5C). This  
489 measurement gives us an estimation of variation or deviation from a consistent, flat  
490 weighting scheme. Here, flat-stimulus sessions showed a strong positive relationship  
491 between threshold and weighting energy ( $r = 0.40$ ,  $p < 0.0001$ ), demonstrating that  
492 during flat-stimulus sessions, employing weights that are highly variable from temporal  
493 uniformity (i.e. have high energy) is detrimental to psychophysical performance. Late-  
494 stimulus sessions showed a moderate positive correlation ( $r = 0.31$ ,  $p = 0.048$ ) and  
495 early-stimulus sessions showed no obvious linear relationship ( $r = -0.004$ ,  $p = 0.99$ ).

496           Taken together, larger variability in weighting and higher energy appear to be  
497 detrimental towards psychometric performance. These were most pronounced in the flat-  
498 stimulus condition, offering a potential explanation for the slight and unexpected  
499 decrease in psychophysical behavior during the flat-stimulus relative to early- and late-  
500 stimulus conditions (Figure 3C, 3D).  
501



502 **DISCUSSION**

503           We used psychophysical reverse correlation in the context of manipulations of  
504 temporal stimulus statistics to examine observers' ability to update their temporal  
505 weighting strategy to match the time course of available evidence in a dynamic motion  
506 discrimination task. First, we found that when motion strength was systematically varied  
507 over time within a stimulus presentation, subjects changed their temporal weighting  
508 strategy to weight the periods of strong motion more heavily than those with weak  
509 motion. Second, weighting strategies were rather consistent across species and  
510 subjects, with the exception of the flat-stimulus condition. Third, session-to-session  
511 variability in strategy was greater in the flat-stimulus condition than in the late- and early-  
512 stimulus conditions. Each of these findings is discussed in more detail below.

513

514 **Temporal weighting likely reflects a combination of dynamic sensory**  
515 **reweighting and decision-making mechanisms**

516           The observation of early sensory evidence exerting a larger effect on decisions  
517 than late evidence (i.e., "early weighting") has been identified in prior work and has been  
518 interpreted within the context of a drift diffusion decision-making model. Early weighting  
519 is often interpreted as a straightforward consequence of accumulation to a decision  
520 bound—sensory data arriving after the bound has been hit does not impact the  
521 accumulator (Huk and Shadlen, 2005; Kiani et al., 2008; Okazawa et al., 2018;  
522 Kawaguchi et al., 2018). Just as past work has taken such early weighting as a signature  
523 of bounded accumulation, late weighting has been posited to reflect leaky integration.  
524 However, such models have been increasingly updated to accommodate either sort of  
525 behavioral signature (Usher and McClelland, 2001; Tsetsos et al., 2012; Bronfman et al.,  
526 2016). Thus, while time varying weighting has been identified before, it is almost always

527 discussed as diagnostic about the structure of a decision-making mechanism, i.e.,  
528 perfect or leaky integration to a bound (fixed or collapsing).

529         The shifts we identified in temporal weighting strategies show that time-varying  
530 weighting of a stimulus is a flexible strategy that adapts to the statistical structure of the  
531 stimulus. This flexibility highlights the possibility of a more direct re-weighting of the  
532 sensory signal itself, regardless of downstream impacts, such as a bound or a leak in the  
533 sensory integration system. Temporal weighting strategies need not be solely the result  
534 of static decision-making mechanisms, but rather could reflect a dynamic strategy for  
535 directly weighting incoming stimulus. Another group made a similar observation  
536 (Cheadle et al., 2014), but in contrast to our findings, their results highlighted sequential  
537 dependencies within single trials and were interpreted via an appeal to normalization.  
538 Such normalization of evidence could be a part of many decision mechanisms, while the  
539 strategic shifts we identified here point to the possibility of a more general and flexible  
540 mechanism of dynamic re-weighting of sensory evidence. By demonstrating an adaptive  
541 weighting strategy that easily shifts towards the most reliable motion information, we  
542 suggest that temporal weighting strategies could be interpreted as a gain on the  
543 incoming stimulus, rather than byproducts of mechanisms beyond the sensory stage of  
544 processing. Indeed, even when presenting a temporally uniform (flat) stimulus, the  
545 neural representation of that stimulus will impose its own time-varying signal to noise  
546 properties on whatever downstream circuits may receive that information for integration  
547 or other such computations (Yates et al., 2017; Churchland et al., 2010; Osborne et al.  
548 2004). It is therefore possible that changes in temporal weighting strategy in the  
549 presence of temporally dynamic stimuli are due to direct reweighting of the time-varied  
550 responses in sensory circuits.

551         It remains to be seen whether or not the observed time-varying weighting in  
552 sensory brain areas can be changed in response to temporal manipulations of the

553 stimulus of the sort we employed, but the well documented effects of temporal attention  
554 in multiple visual cortical areas (Ghose and Maunsell, 2002) lend credence to this  
555 hypothesis. Likewise, changes in spike-count correlation structure with task instruction  
556 have been shown to reflect feedback in early sensory areas (Bondy et al., 2018)  
557 suggesting a possible source for context-dependent reweighting in the current  
558 experiments as well. Notably, our data do not rule out the impacts of decision  
559 mechanisms. The existence of a bound at later stages of decision formation could still  
560 interact with stimulus re-weighting. This could be further sculpted by urgency signals or  
561 time-varying bounds (Ditterich, 2006; Bogacz et al. 2006; Churchland et al., 2008; Cisek  
562 et al., 2009; Hanks et al., 2014; Okazawa et al. 2018). In fact, a potential example of  
563 such an interaction between stimulus reweighting and a bounded decision mechanism  
564 might be present in the late weighting behavior we observed, which often manifested  
565 with a seemingly idiosyncratic, low weight on the final pulse. Although subjects clearly  
566 down-weighted the first few pulses, and up-weighted pulses 5 and 6, the low weight on  
567 the final pulse could be explained as a byproduct of achieving the bound before the  
568 end of the stimulus, even in the late-stimulus condition.

569

570 **Increased variability during the flat-stimulus condition provides insights**  
571 **into previous variability in the literature**

572 Variability in temporal weighting strategy during the flat-stimulus condition was  
573 far larger than in either the early- or late-stimulus conditions. This substantial variability  
574 is of general relevance to the study of evidence accumulation because it is typically  
575 performed using stimuli that are similar to our flat-stimulus condition, in that their  
576 expectation is stationary over time. Although the average weighting strategies for both  
577 humans and macaques in the flat-stimulus condition trend towards early weighting,

578 session-by-session analysis of weighting slopes revealed robust variability (Figure 5).  
579 Few if any prior studies have characterized individual session strategies, likely owing to  
580 low statistical power of alternate designs that rely on post hoc characterization or  
581 infrequent probe trials. Our results suggest that even individual subject averages may  
582 gloss over strategic variability within the observer that occurs over sessions. Likewise,  
583 even the relatively high-resolution session-averages we present here may gloss over  
584 strategic variability over single trials, variability that current trial-based psychophysical  
585 methods lack the resolution to resolve. Consequently, all temporal weighting strategies  
586 presented here (and elsewhere, as far as we know) are computed as an average over  
587 multiple trials, each with a potentially unique temporal weighting strategy.

588         The large session-by-session variability in weighing strategies observed here  
589 may serve to reconcile those presented elsewhere. In the flat-stimulus condition, all time  
590 points (i.e., pulses) are equally informative of the trial outcome, and thus the flat-stimulus  
591 condition is more forgiving of different temporally biased weighting strategies compared  
592 to the early and late conditions, for which only about 1/2 of the stimulus contained  
593 informative evidence on average. Thus, increased variability in weighting strategies  
594 during the flat-stimulus condition compared to early- and late-stimulus conditions is likely  
595 a consequence of temporally uniform stimulus statistics—a feature of most reverse  
596 correlation studies.

597         The consistency of temporal weighting across species displayed in the late and  
598 early stimulus conditions also suggests that, at least for humans and macaques,  
599 interspecies differences need not be a major player in variability of weighting. This is of  
600 possible broader interest, for example, in linking to rodent work (Erlich et al 2015, Scott  
601 et al. 2015, Morcos and Harvey 2016, Pinto et al. 2017, Odoemene et al. 2017, Licata et  
602 al. 2017).

603           One discrepancy across species was present in the flat-stimulus condition, in  
604 which macaque subjects (on average, but most pronounced in M1) displayed an early-  
605 weighting strategy (despite flat stimulus expectation) compared to the flat-weighting  
606 strategy displayed by humans. This could be for a number of reasons. Macaques  
607 performed many more trials and sessions than human subjects, raising the possibility  
608 that extensive training may result in faster decisions, based on early epochs of the  
609 stimulus. This may be further accentuated by a desire to perform more trials and obtain  
610 more liquid reward (a factor not included in experiments with human subjects). While  
611 such a strategy does not in fact change the trial duration or, in turn, the speed-accuracy  
612 tradeoff, it might factor into macaques' behavior. It is noteworthy that the species  
613 difference is only present in the flat-stimulus condition, and not the time-variant  
614 conditions. We believe this is because the flat-expectation and fixed-duration design is  
615 lenient with respect to temporal weighting, granting subjects the liberty to adopt any  
616 number of temporal weighting strategies (Figure 5). This is very different than the time-  
617 varying conditions, which place clear constraints on the temporal weighting strategies  
618 that would benefit the subject. These considerations may serve to reconcile past  
619 conflicting results in different task designs and species, and inform new work going  
620 forward.

621

### 622 **Difficulties in interpreting temporal weighting strategies in light of stimulus** 623 **and task design**

624           Stimulus and task design must be considered to properly interpret the shape of a  
625 temporal weighting strategy. Given that single trials are always rewarded based on the  
626 true net motion presented, regardless of their underlying distribution, all motion pulses  
627 are always informative. Therefore, it is intuitive that highest overall accuracy would be

628 realized via a strategy that assigns equal weighting across all pulses. However, this was  
629 not uniformly present in our dataset, indicating that subjects did not perform the task  
630 optimally. Importantly, the assumption of equal weighting is only a part of an optimality  
631 argument, as equal but low weighting of incoming sensory data would of course be  
632 suboptimal too. Complete optimality of the decision mechanism is a difficult standard to  
633 assess without a detailed characterization of signal and noise in both the stimulus and  
634 the sensory neural representation (Geisler, 1989). Given that most relevant experimental  
635 paradigms do not avail themselves straightforwardly to a formal and complete ideal  
636 observer model, the shape of the temporal weighting provides only partial insight into  
637 decision formation, without a gold standard for the overall level of accuracy.

638         A similar difficulty is present in evaluating the optimality of temporal integration in  
639 fixed duration tasks. Classically, tests of optimal temporal integration appeal to the  
640 relation between viewing duration and accuracy (Kiani et al. 2008, Katz et al. 2015).  
641 However, two issues we have discussed with respect to temporal weighting also speak  
642 to limitations in evaluating optimality in temporal integration via the relation between  
643 viewing duration and accuracy. First, underweighting the sensory evidence before  
644 accumulating is suboptimal, but is not captured by such an analysis, which would lump  
645 such an effect in with sensory noise. Second, although a sensory stimulus may have  
646 certain temporal properties, the neural representation of the sensory stimulus is likely to  
647 have time-varying signal to noise properties (Osborne et al., 2004; Churchland et al.,  
648 2010; Yates et al., 2017). Standard viewing-duration analyses do not distinguish  
649 between the stimulus and the neural signals that are actually used. These two issues  
650 likely interact, with the potential for dynamic strategic weighting to either mirror or  
651 compensate for the dynamics of the incoming sensory stream— making canonical  
652 functional forms of the relations between accuracy and duration rather imperfect tests of  
653 a unique posited mechanism (Huk et al. 2017).

654           Other aspects of experimental design may complexify inferences drawn from the  
655 assessment of temporal weighting as well. For example, although early weighting may  
656 be a general default state (potentially driven by extensive training and/or the default  
657 structure of decision mechanisms), variable duration paradigms may fortify an early  
658 weighting strategy. Variable duration paradigms can be thought of as loosely analogous  
659 to our early-stimulus condition, in that as time progresses, the expected stimulus  
660 strength falls off (owing to the end of the variable duration stimulus). Reaction time tasks  
661 can also facilitate an early weighting strategy as the subject is typically incentivized to  
662 respond as fast as possible, placing more weight on early samples within a stream  
663 (Okazawa et al., 2018). Lastly, time-varying confidence may play a role in shaping  
664 temporal weighting strategies too (Kiani et al., 2009; Kawaguchi et al., 2018). Taken  
665 together, the patterns of selective temporal weighting we have discussed imply that it will  
666 be fruitful to characterize evidence accumulation at a fine grain and to allow for the  
667 potential interplay of both flexible and fixed mechanisms in sculpting the resulting  
668 dynamics.

669           Our characterizations of temporal weighting are of course inherently limited by  
670 the assumptions of logistic regression. While it is clear that subjects weigh temporal  
671 sections of the stimulus in proportion to their expected motion signal, it seems unlikely  
672 that the way the brain performs this task is completely described by logistic regression.  
673 There are likely a cascade of nonlinearities between stimulus and response that cannot  
674 be fully described by a set of linear weights passed through a sigmoid, which implies that  
675 the exact pattern and magnitudes of an individual temporal kernel is an incomplete  
676 description of the decision process. However, given the close correspondence between  
677 kernels computed using only flat-expectation, zero-mean trials and kernels computed  
678 using all trials (where there is often temporal correlation in the stimulus), any nonlinearity  
679 in mapping from stimulus to sensory evidence appears to have a minimal impact on our

680 core result: differences between temporal stimulus statistics can exert systematic and  
681 interpretable effects on temporal weighting strategies.

682 More generally, our results provide an opportunity to reconnect perceptual  
683 decision-making models with other frameworks for information integration. For example,  
684 the dynamic temporal weighting we observed has a direct connection to classical  
685 Bayesian integration (Kording and Wolpert, 2006; Angelaki et al. 2009; Hillis et al. 2004;  
686 Knill 2007; Fetsch et al., 2009). Over repeated exposure to a given stimulus condition,  
687 subjects learn to weigh stimulus cues according to reliability. In our experiment, time  
688 epochs (motion pulses) can be thought of as akin to cues: each motion pulse is a cue  
689 towards the trial's net direction, but during early- and late-stimulus conditions subjects  
690 must learn to down-weight noisy epochs, and up-weight reliable ones. Cue combination  
691 with reliability-based weighting has been commonly observed both within and across  
692 sensory domains (Hillis et al., 2004; Morgan et al., 2008; Angelaki et al., 2009; Fetsch et  
693 al., 2009; Fetsch et al., 2011). While Bayesian integration has been discussed  
694 specifically with respect to bounded accumulation (Beck et al., 2008), it also lends itself  
695 to a reliability-based readout of a temporally dynamic sensory representation. Time  
696 points in the sensory response with a higher signal-to-noise ratio may be more strongly  
697 weighted towards choice. For example, as discussed above, a tendency towards early  
698 weighting in the flat stimulus condition could be reflective of temporal variation during  
699 sensory encoding rather than an effect of downstream mechanisms such as a bound.

700 We are encouraged by this mapping to a Bayesian framework, and the implication that  
701 further manipulations of reliability of evidence in time can continue to build tighter links  
702 (or reveal contrasts) between cue integration and temporal integration (Katz et al., 2015;  
703 Hanks et al., 2011).

704 In summary, past work has used reverse correlation and time-varied stimuli to  
705 probe temporal integration. In the present study we used a reverse correlation task in the



706 context of tractable manipulations of stimulus statistics, allowing for direct control over a  
707 subject's temporally weighting strategy. Although the neural correlates of such changes  
708 remain uncertain, the ability to both manipulate and characterize temporal weighting  
709 strategies should provide a powerful tool for neurophysiological experiments to come.  
710

711 **References**

712

713 Angelaki, D. E., Gu, Y., &amp; Deangelis, G. C. (2009). Multisensory integration:

714       psychophysics, neurophysiology, and computation. *Current Opinion in*715       *Neurobiology*, 19(4), 452–458. <http://doi.org/10.1016/j.conb.2009.06.008>

716 Bair, W., &amp; Movshon, J. A. (2004). Adaptive temporal integration of motion in

717       direction-selective neurons in macaque visual cortex. *The Journal of*718       *Neuroscience*, 24(33), 7305–7323.

719 Beck, J. M., Ma, W. J., Kiani, R., Hanks, T., Churchland, A. K., Roitman, J., et al. (2008).

720       Probabilistic population codes for Bayesian decision making. *Neuron*, 60(6),721       1142–1152. <http://doi.org/10.1016/j.neuron.2008.09.021>722       <http://doi.org/10.1523/JNEUROSCI.0554-04.2004>

723 Bogacz, R., Brown, E., Moehlis, J., Holmes, P., &amp; Cohen, J. D. (2006). The

724       physics of optimal decision making: a formal analysis of models of

725       performance in two-alternative forced-choice tasks. *Psychological Review*,726       113(4), 700–765. <http://doi.org/10.1037/0033-295X.113.4.700>

727 Bondy, A. G., Haefner, R. M., &amp; Cumming, B. G. (2018). Feedback determines

728       the structure of correlated variability in primary visual cortex. *Nature*729       *Neuroscience*, 21(4), 598–606. <http://doi.org/10.1038/s41593-018-0089-1>

730 Britten, K. H., Newsome, W. T., Shadlen, M. N., Celebrini, S., &amp; Movshon, J. A.

731       (1996). A relationship between behavioral choice and the visual responses

732       of neurons in macaque MT. *Visual Neuroscience*, 13(1), 87–100.

733 Britten, K. H., Shadlen, M. N., Newsome, W. T., &amp; Movshon, J. A. (1992). The

734       analysis of visual motion: a comparison of neuronal and psychophysical

735       performance. *The Journal of Neuroscience*, 12(12), 4745–4765.

- 736 Bronfman, Z. Z., Brezis, N., & Usher, M. (2016). Non-monotonic Temporal  
737 Weighting Indicates a Dynamically Modulated Evidence-Integration  
738 Mechanism. *PLoS Computational Biology*, *12*(2), e1004667.  
739 <http://doi.org/10.1371/journal.pcbi.1004667>
- 740 Carland, M. A., Marcos, E., Thura, D., & Cisek, P. (2016). Evidence against  
741 perfect integration of sensory information during perceptual decision  
742 making. *Journal of Neurophysiology*, *115*(2), 915–930.  
743 <http://doi.org/10.1152/jn.00264.2015>
- 744 Cheadle, S., Wyart, V., Tsetsos, K., Myers, N., de Gardelle, V., Hecce Castañón,  
745 S., & Summerfield, C. (2014). Adaptive gain control during human  
746 perceptual choice. *Neuron*, *81*(6), 1429–1441.  
747 <http://doi.org/10.1016/j.neuron.2014.01.020>
- 748 Churchland, A. K., Kiani, R., & Shadlen, M. N. (2008). Decision-making with  
749 multiple alternatives. *Nature Neuroscience*, *11*(6), 693–702.  
750 <http://doi.org/10.1038/nn.2123>
- 751 Churchland, M. M., Yu, B. M., Cunningham, J. P., Sugrue, L. P., Cohen, M. R.,  
752 Corrado, G. S., et al. (2010). Stimulus onset quenches neural variability: a  
753 widespread cortical phenomenon. *Nature Neuroscience*, *13*(3), 369–378.  
754 <http://doi.org/10.1038/nn.2501>
- 755 Cisek, P., Puskas, G. A., & El-Murr, S. (2009). Decisions in changing conditions:  
756 the urgency-gating model. *The Journal of Neuroscience*, *29*(37), 11560  
757 11571. <http://doi.org/10.1523/JNEUROSCI.1844-09.2009>
- 758 Ditterich, J. (2006). Evidence for time-variant decision making. *The European Journal*  
759 *of Neuroscience*, *24*(12), 3628–3641.  
760 <http://doi.org/10.1111/j.1460-9568.2006.05221.x>
- 761 Erlich, J. C., Brunton, B. W., Duan, C. A., Hanks, T. D., & Brody, C. D. (2015).

- 762 Distinct effects of prefrontal and parietal cortex inactivations on an  
763 accumulation of evidence task in the rat. *eLife*, 4, 8166.  
764 <http://doi.org/10.7554/eLife.05457>
- 765 Eastman, K. M., & Huk, A. C. (2012). PLDAPS: A Hardware Architecture and  
766 Software Toolbox for Neurophysiology Requiring Complex Visual Stimuli  
767 and Online Behavioral Control. *Frontiers in Neuroinformatics*, 6, 1.  
768 <http://doi.org/10.3389/fninf.2012.00001>
- 769 Fetsch, C. R., Pouget, A., Deangelis, G. C., & Angelaki, D. E. (2011). Neural  
770 correlates of reliability-based cue weighting during multisensory  
771 integration. *Nature Neuroscience*, 15(1), 146–154.  
772 <http://doi.org/10.1038/nn.2983>
- 773 Fetsch, C. R., Turner, A. H., Deangelis, G. C., & Angelaki, D. E. (2009). Dynamic  
774 reweighting of visual and vestibular cues during self-motion perception.  
775 *The Journal of Neuroscience*, 29(49), 15601–15612.  
776 <http://doi.org/10.1523/JNEUROSCI.2574-09.2009>
- 777 Geisler, W. S. (1989). Sequential ideal-observer analysis of visual  
778 discriminations. *Psychological Review*, 96(2), 267–314.
- 779 Ghose, G. M. (2006). Strategies optimize the detection of motion transients.  
780 *Journal of Vision*, 6(4), 429–440. <http://doi.org/10.1167/6.4.10>
- 781 Ghose, G. M., & Maunsell, J. H. R. (2002). Attentional modulation in visual cortex  
782 depends on task timing. *Nature*, 419(6907), 616–620.  
783 <http://doi.org/10.1038/nature01057>
- 784 Gold, J. I., & Shadlen, M. N. (2007). The neural basis of decision making. *Annual*  
785 *Review of Neuroscience*, 30(1), 535–574.  
786 <http://doi.org/10.1146/annurev.neuro.29.051605.113038>
- 787 Hanks, T. D., Mazurek, M. E., Kiani, R., Hopp, E., & Shadlen, M. N. (2011).

- 788 Elapsed decision time affects the weighting of prior probability in a  
789 perceptual decision task. *The Journal of Neuroscience*, 31(17), 6339  
790 6352. <http://doi.org/10.1523/JNEUROSCI.5613-10.2011>
- 791 Hanks, T., Kiani, R., & Shadlen, M. N. (2014). A neural mechanism of speed  
792 accuracy tradeoff in macaque area LIP. *eLife*, 3, 433.  
793 <http://doi.org/10.7554/eLife.02260>
- 794 Hillis, J. M., Watt, S. J., Landy, M. S., & Banks, M. S. (2004). Slant from texture  
795 and disparity cues: optimal cue combination. *Journal of Vision*, 4(12),  
796 967–992. <http://doi.org/10.1167/4.12.1>
- 797 Huk, A. C., & Shadlen, M. N. (2005). Neural activity in macaque parietal cortex  
798 reflects temporal integration of visual motion signals during perceptual  
799 decision making. *The Journal of Neuroscience*, 25(45), 10420–10436.  
800 <http://doi.org/10.1523/JNEUROSCI.4684-04.2005>
- 801 Huk, A.C., Katz, L. N., Yates, J. L. (2017). The Role of the Lateral Intraparietal Area in  
802 (the Study of) Decision Making. *Annual Review of Neuroscience*, 40(1), 349  
803 372.
- 804 Katz, L. N., Hennig, J. A., Cormack, L. K., & Huk, A. C. (2015). A Distinct  
805 Mechanism of Temporal Integration for Motion through Depth. *The Journal*  
806 *of Neuroscience*, 35(28), 10212–10216.  
807 <http://doi.org/10.1523/JNEUROSCI.0032-15.2015>
- 808 Katz, L. N., Yates, J. L., Pillow, J. W., & Huk, A. C. (2016). Dissociated functional  
809 significance of decision-related activity in the primate dorsal stream.  
810 *Nature*, 535(7611), 285–288. <http://doi.org/10.1038/nature18617>
- 811 Kawaguchi, K., Clery, S., Pourriahi, P., Seillier, L., Haefner, R., & Nienborg, H.  
812 (2018). Using confidence inferred from pupil-size to dissect perceptual  
813 task-strategy: support for a bounded decision-formation process. *bioRxiv*,

- 814 269159. <http://doi.org/10.1101/269159>
- 815 Kiani, R., & Shadlen, M. N. (2009). Representation of confidence associated with  
816 a decision by neurons in the parietal cortex. *Science (New York, N.Y.)*,  
817 324(5928), 759–764. <http://doi.org/10.1126/science.1169405>
- 818 Kiani, R., Hanks, T. D., & Shadlen, M. N. (2008). Bounded integration in parietal  
819 cortex underlies decisions even when viewing duration is dictated by the  
820 environment. *The Journal of Neuroscience*, 28(12), 3017–3029.  
821 <http://doi.org/10.1523/JNEUROSCI.4761-07.2008>
- 822 Knill, D. C. (2007). Robust cue integration: a Bayesian model and evidence from  
823 cue-conflict studies with stereoscopic and figure cues to slant. *Journal of*  
824 *Vision*, 7(7), 5.1–24. <http://doi.org/10.1167/7.7.5>
- 825 Körding, K. P., & Wolpert, D. M. (2006). Bayesian decision theory in sensorimotor  
826 control. *Trends in Cognitive Sciences*, 10(7), 319–326.  
827 <http://doi.org/10.1016/j.tics.2006.05.003>
- 828 Licata, A. M., Kaufman, M. T., Raposo, D., Ryan, M. B., Sheppard, J. P., &  
829 Churchland, A. K. (2017). Posterior Parietal Cortex Guides Visual  
830 Decisions in Rats. *The Journal of Neuroscience*, 37(19), 4954–4966.  
831 <http://doi.org/10.1523/JNEUROSCI.0105-17.2017>
- 832 Morcos, A. S., & Harvey, C. D. (2016). History-dependent variability in population  
833 dynamics during evidence accumulation in cortex. *Nature Neuroscience*,  
834 19(12), 1672–1681. <http://doi.org/10.1038/nn.4403>
- 835 Morgan, M. L., Deangelis, G. C., & Angelaki, D. E. (2008). Multisensory  
836 integration in macaque visual cortex depends on cue reliability. *Neuron*,  
837 59(4), 662–673. <http://doi.org/10.1016/j.neuron.2008.06.024>
- 838 Newsome, W. T., & Paré, E. B. (1988). A selective impairment of motion  
839 perception following lesions of the middle temporal visual area (MT). *The*

- 840 *Journal of Neuroscience*, 8(6), 2201–2211.
- 841 Nienborg, H., & Cumming, B. G. (2009). Decision-related activity in sensory  
842 neurons reflects more than a neuron's causal effect. *Nature*, 459(7243)  
843 89–92. <http://doi.org/10.1038/nature07821>
- 844 Odoemene, O., Pisupati, S., Nguyen, H., & Churchland, A. K. (2017). Visual  
845 evidence accumulation guides decision-making in unrestrained mice. *bioRxiv*,  
846 <http://doi.org/10.1101/195792>
- 847 Okazawa, G., Sha, L., Purcell, B. A., & Kiani, R. (2018). Psychophysical reverse  
848 correlation reflects both sensory and decision-making processes. *bioRxiv*,  
849 <http://doi.org/10.1101/273680>
- 850 Osborne, L. C., Bialek, W., & Lisberger, S. G. (2004). Time course of information  
851 about motion direction in visual area MT of macaque monkeys. *The*  
852 *Journal of Neuroscience*, 24(13), 3210–3222.  
853 <http://doi.org/10.1523/JNEUROSCI.5305-03.2004>
- 854 Ossmy, O., Moran, R., Pfeffer, T., Tsetsos, K., Usher, M., & Donner, T. H. (2013).  
855 The timescale of perceptual evidence integration can be adapted to the  
856 environment. *Current Biology : CB*, 23(11), 981–986.  
857 <http://doi.org/10.1016/j.cub.2013.04.039>
- 858 Pinto, L., Koay, S. A., Engelhard, B., Yoon, A. M., Deverett, B., Thiberge, S. Y., et  
859 al. (2017). An accumulation-of-evidence task using visual pulses for mice  
860 navigating in virtual reality. <http://doi.org/10.1101/232702>
- 861 Sahani, M., & Linden, J. F. (2003). Evidence optimization techniques for estimating  
862 stimulus-response functions. In *Advances in neural information processing*  
863 *systems* (pp. 317-324).
- 864 Scott, B. B., Constantinople, C. M., Erlich, J. C., Tank, D. W., & Brody, C. D.

- 865 (2015). Sources of noise during accumulation of evidence in unrestrained  
866 and voluntarily head-restrained rats. *eLife*, 4, e11308.  
867 <http://doi.org/10.7554/eLife.11308>
- 868 Tsetsos, K., Gao, J., McClelland, J. L., & Usher, M. (2012). Using Time-Varying  
869 Evidence to Test Models of Decision Dynamics: Bounded Diffusion vs. the  
870 Leaky Competing Accumulator Model. *Frontiers in Neuroscience*, 6, 79.  
871 <http://doi.org/10.3389/fnins.2012.00079>
- 872 Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: the  
873 leaky, competing accumulator model. *Psychological Review*, 108(3), 550–  
874 592.
- 875 Wichmann FA, Hill NJ. The psychometric function: I. Fitting, sampling, and  
876 goodness of fit *Perception & Psychophysics*. 2001;63:1293–1313.
- 877 Yates, J. L., Park, I. M., Katz, L. N., Pillow, J. W., & Huk, A. C. (2017). Functional  
878 dissection of signal and noise in MT and LIP during decision-making.  
879 *Nature Neuroscience*, 20(9), 1285–1292. <http://doi.org/10.1038/nn.4611>  
880



881 **Figure captions**

882 **Figure 1. Sequence of trial events.** **A.** Subjects fixated on a central point through the  
883 appearance of targets and motion stimulus until the disappearance of the fixation point  
884 (“go”). Choices were made with saccades to the target corresponding to the perceived  
885 net direction of motion. Initial fixation time, target-on duration, and time until fixation point  
886 disappearance were randomly varied. **B.** An example frame of the Gabor motion pulse  
887 stimulus. The stimulus is composed of nineteen Gabor patches, where motion strength  
888 is denoted by the proportion of coherently drifting Gabors out of the total number  
889 elements in the stimulus. **C.** Motion pulse values are generated from Gaussian  
890 distributions spanning a large range of possible motion strengths in either direction. A  
891 single trial consists of seven motion pulses, each randomly drawn from one of the  
892 Gaussians. Example trials for pulses drawn from each Gaussian (strong left/right, weak  
893 left/right, and zero) are presented in cartoon form where the number of arrows  
894 represents the number of coherently drifting Gabor elements.

895

896 **Figure 2. Temporal weighting profiles and psychometric functions for humans and**  
897 **macaques across flat-, late-, and early-stimulus conditions.** **A-C.** Top: schematic of  
898 the Gaussian distributions that generate the motion pulses. In the flat-stimulus (**A**),  
899 Gaussians remain stationary over time. In the late- (**B**) and early-(**C**) stimulus conditions,  
900 the distribution means for signal trials are varied over time. Bottom: Examples sessions  
901 for each stimulus condition. Motion pulse values are drawn from their color-matched  
902 Gaussians on each pulse such that the mean of many trials (bold line) reflects the  
903 temporal structure of the mean of the Gaussians. Motion pulse values in single trials  
904 (semi-transparent traces) vary considerably, in accordance with the variance of color-  
905 matched Gaussians. **D-F.** Temporal weighting profiles averaged across all subjects

906 (human and macaque), and sessions within the flat- (**D**), late- (**E**), and early-stimulus (**F**)  
907 conditions, showing the mean weight assigned to each of the seven motion pulses. Error  
908 bars represent  $\pm 1$  SEM. **G-I**. Psychometric performance averaged over all session for  
909 flat- (**G**), late- (**H**), and early-stimulus (**I**) conditions, fit by a logistic function capturing the  
910 dependence of choice on stimulus strength. Error bars represent  $\pm 1$  SEM (often  
911 occluded by points).

912

913 **Figure 3. Comparison of temporal weighting and psychometric functions within**  
914 **species across stimulus conditions. A, B.** Temporal weighting profiles for macaques  
915 (**A**) and humans (**B**) averaged over all sessions in the early-, flat-, and late-stimulus  
916 conditions, fit by a linear model (semi-transparent lines) to capture the overall trend of  
917 the weights. Error bars represent  $\pm 1$  SEM.. **C, D.** Psychometric behavior of macaques  
918 (**C**) and humans (**D**) averaged over all sessions in the early-, flat-, and late-stimulus  
919 conditions, fit by a logistic function to capture the dependence of choice on stimulus  
920 strength. Error bars represent  $\pm 1$  SEM. **E.** Each subject's proportion correct for  
921 inconsistent trials (where the strongest pulse is in the opposite direction of the full-trial,  
922 net direction) and difficulty-matched consistent trials (where the strongest pulse is in the  
923 same direction as the full-trial, net direction). Error bars represent 95% binomial  
924 confidence intervals.

925

926 **Figure 4. Temporal weighting strategies for individual subjects across stimulus**  
927 **conditions. A–C.** Average temporal weighting strategies for individual human and  
928 macaque subjects (columns) during the late- (**A**), early- (**B**) and flat-stimulus (**C**)  
929 condition. Error bars represent  $\pm 1$  SEM. **D.** A within-subject comparison of the shift in  
930 temporal weighting strategies from flat-stimulus to early (top) and flat-stimulus to late

931 (bottom), represented as the slope of the linear model fit to subject temporal weights.

932 Error bars represent  $\pm 1$  SEM.

933

934 **Figure 5. Variability in temporal weighting profiles and psychometric performance.**

935 **A.** Distribution of temporal weighting profiles over sessions and subjects across the  
936 early- flat- and late-stimulus conditions, represented as the slope of the linear model fit  
937 to the temporal weights of each session. Negative slope values indicate an early  
938 weighting strategy, positive values indicate late. Triangles denote the median. **B.** The  
939 relationship between psychometric performance (75% psychophysical threshold) and  
940 temporal weighting (slope of linear fit to temporal weights), over all sessions across the  
941 three stimulus conditions. **C.** The relationship between psychometric performance (75%  
942 psychophysical threshold) and temporal weighting energy (sum of squared errors of  
943 temporal weight values from their mean), over all sessions across the three stimulus  
944 conditions.  
945









