Cherry-picking information: humans actively sample evidence to support prior beliefs

Paula Kaanders[1,2,*], **Pradyumna Sepulveda**[3], **Tomas Folke**[4], **Pietro Ortoleva**[5], **and Benedetto De Martino**[3, 6, *]

[1]Department of Experimental Psychology, University of Oxford, Oxford, United Kingdom; [2]Wellcome Centre for Integrative Neuroimaging, University of Oxford, Oxford, United Kingdom; [3]Institute of Cognitive Neuroscience, University College London, London, United Kingdom; [4]Centre for Business Research, Cambridge Judge Business School, University of Cambridge, Cambridge, United Kingdom; [5]Department of Economics, Columbia University, New York, United States of America; [6]Wellcome Centre for Human Neuroimaging, University College London, London, United Kingdom

* Correspondence: paula.kaanders@psy.ox.ac.uk, benedettodemartino@gmail.com

Abstract

No one likes to be wrong. Previous research has shown that participants may underweight information incompatible with previous choices, a phenomenon called confirmation bias. In this paper we argue that a similar bias exists in the way information is actively sought. We investigate how choice influences information gathering using a perceptual choice task and find that participants sample more information from a previously chosen alternative. Furthermore, the higher the confidence in the initial choice, the more biased information sampling becomes. As a consequence, when faced with the possibility of revising an earlier decision, participants are more likely to stick with their original choice, even when incorrect. Critically, we show that agency controls this phenomenon. The effect disappears in a fixed sampling condition where presentation of evidence is controlled by the experimenter, suggesting that the way in which confirmatory evidence is acquired critically impacts the decision process. These results suggest active information acquisition plays a critical role in the propagation of strongly held beliefs over time.

40 We are constantly deciding what information to sample to guide future decisions. This is
41 no easy feat, as there is a vast amount of information available in the world. In theory, an
42 agent should look for information to maximally reduce their uncertainty (Schulz &
43 Gershman, 2019; Schwartenbeck et al., 2019; Yang et al., 2016). There is some evidence
44 that this is indeed what humans do (Kobayashi & Hsu, 2019; Meyer & Shi, 1995; Steyvers
45 et al., 2009; Wilson et al., 2014), but a range of other drivers of information search have
46 been identified that make us deviate from the sampling behaviour of optimal agents
47 (Eliaz & Schotter, 2007; Gesiarz et al., 2019; Hunt et al., 2016; Kobayashi et al., 2019;
48 Rodriguez Cabrero et al., 2019; Sharot, 2011; van Lieshout et al., 2018; Wang & Hayden,
49 2019).
50
51 Confirmation bias is defined as the tendency of agents to seek out or overweight evidence
52 that aligns with their beliefs while avoiding or underweighting evidence that contradicts
53 them (Hart et al., 2009; Lord et al., 1979; Nickerson, 1998; Stroud, 2008; Wason, 1960,
54 1968). This effect has often been described in the context of what media people choose
55 to consume (Bakshy et al., 2015; Bennett & Iyengar, 2008; Pariser, 2012). Recently,
56 cognitive scientists have realised confirmation bias may not be restricted to this domain
57 and may reveal a fundamental property of the way in which the brain drives information
58 search. As such, researchers have started to investigate confirmation bias in simple
59 perceptual and value-based choice tasks. For example, sensitivity to new information has
60 been shown to decrease after choice (Bronfman et al., 2015), while Talluri, Urai et al.
61 (2018) have shown that sensitivity to a stimulus is increased when it is consistent with
62 the stimulus presented before. Furthermore, in Palminteri et al. (2017) learning rates
63 were found to be higher for feedback indicating positive obtained outcomes as well as for
64 negative forgone outcomes. This effect has been replicated in recent studies that suggest
65 confirmation bias during learning might be adaptive in some contexts (Tarantola, Folke
66 et al., 2021; Lefebvre et al., 2020; Salem-Garcia et al., 2021).
67
68 Most of the existing research on perceptual decision-making has focused on the
69 weighting of incoming information during evidence accumulation or belief updating.
70 Other findings suggest that agents also actively search for confirmatory evidence (Hunt
71 et al., 2016, 2018; Tolcott et al., 1989; Wason, 1968), but none of these experiments
72 investigated the effect of biased information sampling in simple, perceptual choice. As
73 most paradigms currently used to study confirmation bias do not explicitly focus on
74 active information sampling performed by the agent, therefore it is unclear whether
75 confirmation bias arises from evidence underweighting, biased information sampling, or
76 both.
77
78 Another important aspect of confirmation bias is how it modulates and is modulated by
79 decision confidence. Confidence is known to play an important role in guiding decisions
80 (Bogacz et al., 2010; Boldt et al., 2019; Folke et al., 2017; Rabbitt, 1966; Yeung &
81 Summerfield, 2012). It is therefore likely it may also guide the information search
82 preceding choice (Desender et al., 2019). In fact, Rollwage et al. (2020) recently showed
83 that confidence indeed affects a neural signal of confirmatory evidence integration. By
84 definition, the higher confidence is in a choice, the stronger the decision-maker's belief is
85 in the correctness of this choice (Fleming & Lau, 2014; Pouget et al., 2016). Accordingly,
86 we predict that confirmation bias in information sampling will be stronger after choices
87 made with high confidence.
88

89   We used two perceptual binary forced choice tasks with two choice phases separated by
90   a free sampling phase to test our hypotheses that confirmation bias arises from biased
91   information sampling, that this effect influences future choice and that confidence
92   mediates this behavioural tendency.
93
94
95   **Results**
96
97   Here we present the results of two independent experiments we conducted to test our
98   hypotheses.
99
100   *Experiment 1*
101
102   Participants performed a perceptual two-alternative forced choice (2AFC) task (Figure
103   1A), in which they were briefly presented with patches containing multiple dots.
104   Participants were asked to judge which patch contained the most dots. After their initial
105   choice (phase 1), they rated their confidence that they were correct. Subsequently, in the
106   second phase of the trial (phase 2) participants could freely sample (i.e., see) each dot
107   patch using key presses to switch between them as frequently as they liked for 4000ms.
108   Participants could only view one patch at a time. They were then prompted to make a
109   second choice (phase 3), in which they could either confirm their initial choice or change
110   their mind. After this second decision, they again reported their degree of confidence.
111
112   As expected, participants' performance was highly sensitive to the amount of perceptual
113   evidence presented in a given trial (i.e. the difference in the amount of dots in the left and
114   right patches; Fig 1B). Participants were also more accurate in the second choice
115   compared to the first choice, showing that they used the additional sampling time
116   between choice phases to accumulate additional perceptual evidence that allowed them
117   to increase accuracy on the second choice (t_{27}=8.74, p<0.001).
118
119   Our main hypothesis was that participants would prefer to gather perceptual information
120   that was likely to confirm their initial decision. Conversely, we expected them to be less
121   likely to acquire evidence that would disconfirm it. To perform well on this task, an agent
122   would have to attend equally to the two patches, as the goal requires computing the
123   difference in dots between the two patches. Therefore, any imbalance in sampling time
124   would not necessarily be beneficial to completing the task. However, we expected that in
125   the free sampling phase (phase 2) of the trial participants would spend more time
126   viewing their chosen patch. In line with this hypothesis, during the sampling phase,
127   participants spent more time viewing the patch they chose in the first decision phase
128   relative to the unchosen alternative (Figure 1C; t_{27}=7.28, p<0.001).
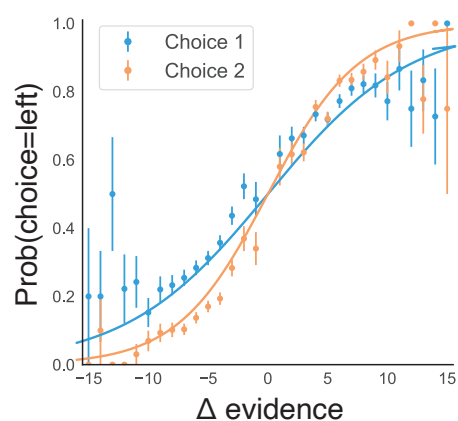
129
130

- Experiment 1 -

A

B

C

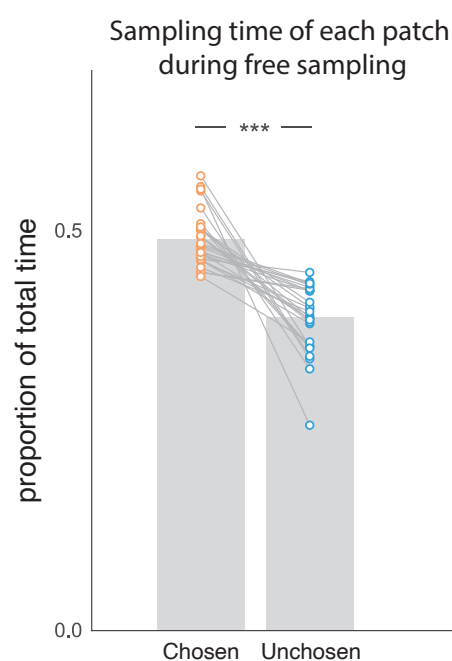Sampling time of each patch during free sampling
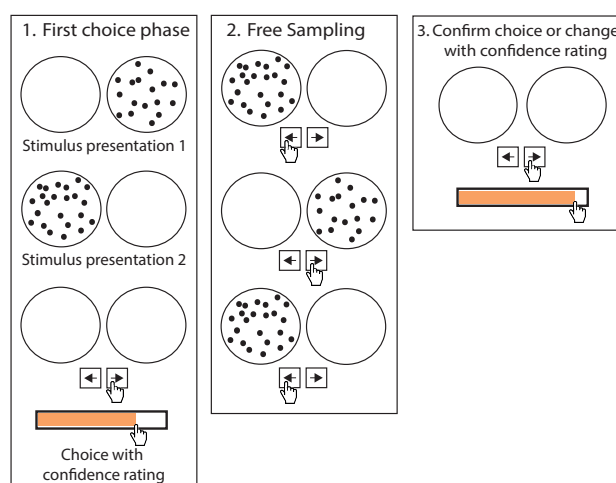
131 **Figure 1.** Task design and participant behaviour for experiment 1. (A) Task structure. Participants had to
132 choose which of two dot patches contained the most dots (phase 1) and rate their confidence in this choice.
133 Then participants were given 4000ms to view the dots which they could allocate between the two patches
134 in whichever way they liked (phase 2) by pressing the left and right arrow keys. Finally, participants made
135 a second choice about the same set of stimuli and rated their confidence again (phase 3). (B) Participants
136 effectively used the stimuli to make correct choices and improved upon their performance on the second
137 choice. This psychometric curve is plotting the probability of choosing the left option as a function of the
138 evidence difference between the two stimuli for each of the two choice phases. (C) In the free sampling
139 phase (phase 2) participants spent more time viewing the stimulus they chose on the preceding choice
140 phase than the unchosen option. Data points represent individual participants.

141

142

143 Confidence in a choice reflects the strength of the participant's belief that their choice
144 was correct. Consequently, we hypothesised that participants' preference for gathering
145 confirmatory evidence for a given choice would be modulated by their confidence in
146 that choice. As such, we expected choices made with higher confidence would lead to a
147 stronger sampling bias favouring the chosen patch over the unchosen patch. In a
148 hierarchical regression predicting sampling time difference between the two patches,
149 we found a significant interaction between choice and confidence, such that the higher
150 the degree of confidence was in the first choice, the more sampling was biased in favour
151 of that choice (Figure 2; $t_{26.96}=5.26$, $p<0.001$; see supplemental materials S1). We also
152 saw a significant main positive effect of choice on sampling time difference, such that a
153 chosen stimulus was likely to be sampled for longer during the sampling phase (Figure
154 2B; $t_{26.96}=9.64$, $p<0.001$) as shown in the previous section. There was also a significant
155 main positive effect of evidence difference on sampling time difference, whereby the
156 correct stimulus (containing the most dots) was sampled for longer (Figures 2B;
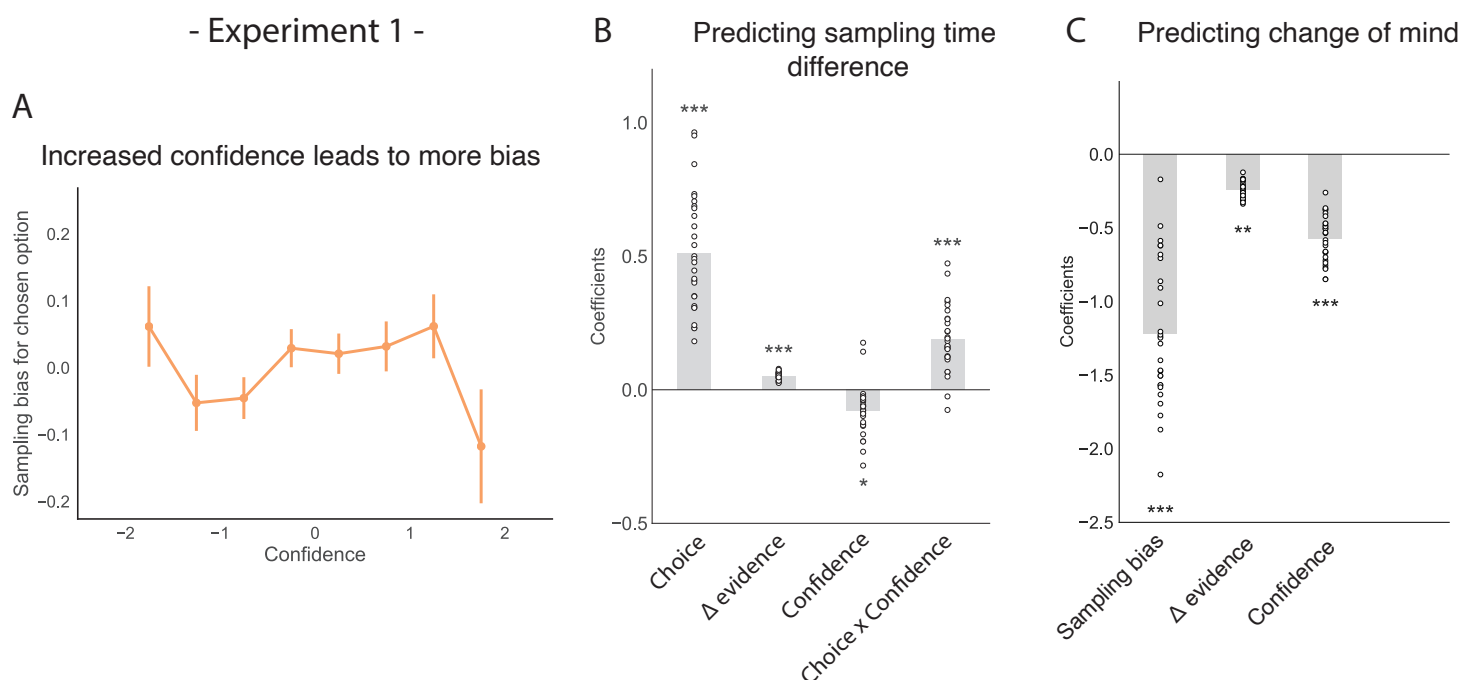157 $t_{26.96}=13.02$, $p<0.001$).

158

159

160 **Figure 2.** The effect of choice on sampling behaviour is mediated by confidence in experiment 1.
161 Participants were less likely to change their mind if they showed a strong sampling bias for their initially
162 chosen option in the sampling phase. (A) Sampling bias in favour of the chosen option increases as a

163 function of confidence in the initial choice. Confidence and sampling bias are both normalised in this plot.
164 (B) There is a significant main effect of choice on sampling time difference, such that an option is sampled
165 for longer if it was chosen, and a significant interaction effect of Choice x Confidence, such that options
166 chosen with high confidence are sampled for even longer. (C) There is a main negative effect of sampling
167 bias on change of mind, such that participants were less likely to change their mind in the second decision
168 phase (phase 3) the more they sampled their initially chosen option in the free sampling phase (phase 2).
169 (B-C) Plotted are fixed-effect coefficients from hierarchical regression models predicting the sampling time
170 (how long each patch was viewed in the sampling phase) difference between the left and right stimuli. Data
171 points represent regression coefficients for each individual participant.

174 Our choices are determined by the evidence we have gathered before making a decision.
175 Therefore, we expected that biased sampling in the free sampling phase (phase 2) would
176 impact decisions in the second decision phase (phase 3). Specifically, we hypothesised
177 that the more strongly participants preferred sampling their chosen patch, the more
178 likely they were to choose it again. Using a multinomial hierarchical logistic regression,
179 we found that bias in sampling time towards the previously chosen option was indeed
180 predictive of future choice. In other words, the more participants sampled a previously
181 chosen option, the more likely they were to choose it again (Figure 2C; z=-11.0, p<0.001;
182 see S2). Furthermore, evidence difference and confidence were significantly negatively
183 related to subsequent changes of mind, whereby participants were less likely to change
184 their mind if their initial choice was correct and if it was made with high confidence
185 (Figure 2C; main effect of evidence difference – z=-3.06, p<0.01, main effect of confidence
186 – z=-10.12, p<0.001).

Experiment 2

191 While the results from the first experiment showed an effect of biased sampling on
192 subsequent choice, it was not clear whether this effect arises from biased evidence
193 accumulation caused by differential exposure to the perceptual evidence, or if the
194 sampling choices themselves drive the effect. In other words, would the same choice bias
195 appear if participants were passive recipients of biased sampling, or does the choice bias
196 require that participants make their own sampling decisions?

198 We addressed this question in a follow-up eye-tracking study (Figure 3A) in which we
199 introduced a control task, a 'fixed-viewing condition', in which participants did the same
200 task, but did not have the possibility to freely sample the patches in phase 2. Instead, the
201 dot patches were shown for a set amount of time. In one-third of trials, the patches were
202 shown an equal amount of time; in two-thirds of trials, one patch was shown three times
203 longer than the other. Each participant completed two sessions, one session involved free
204 sampling, similar to Experiment 1. The other involved a fixed-viewing control task.
205 Session order was pseudo-randomized between participants. Furthermore, presentation
206 of the visual stimuli and all the participants' responses were gaze-contingent. This meant
207 that during the sampling phase (phase 2), the dot patches were only presented when
208 participants fixated inside the patch. Furthermore, we hypothesised that sampling bias
209 might be stronger when more time to sample is available. Specifically, we expected that
210 participants might become more biased in their sampling throughout the sampling phase
211 as they gradually become more confident in their belief. Therefore, we manipulated the
212 length of phase 2 to be either 3000ms, 5000ms, or 7000ms. Again, participants were
213 sensitive to the difficulty of the given trials (Figure 3B) and were more accurate on the

214   second choice compared to the first choice (t_{17}=6.80, p<0.001).

215

216   We also replicated our main result in the free sampling condition showing that
217   participants spent more time viewing the patch they just chose (Figure 3C; t_{17}=3.52,
218   p<0.01). Furthermore, the size of this sampling time bias was proportional to the total
219   amount of sampling time available in study 2 (Figure S3.1), suggesting that there was no
220   particular period of time within the sampling phase where confirmatory information
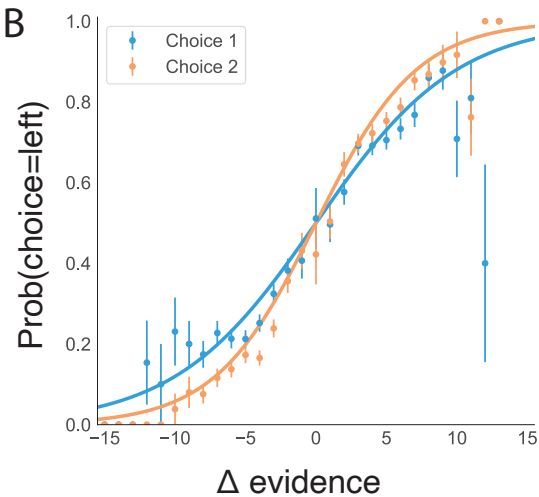221   sampling was more likely contrary to our expectation.
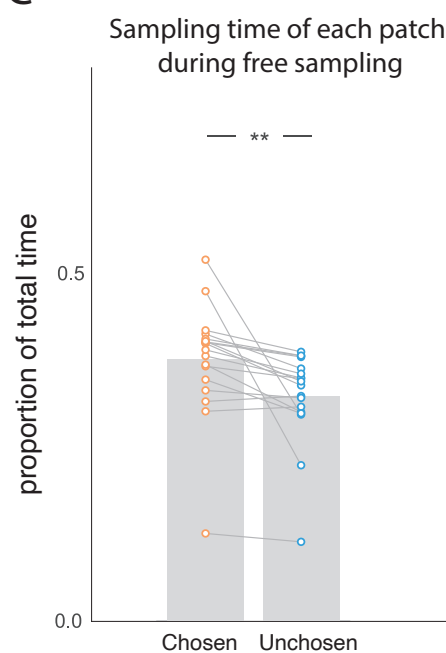
222

223
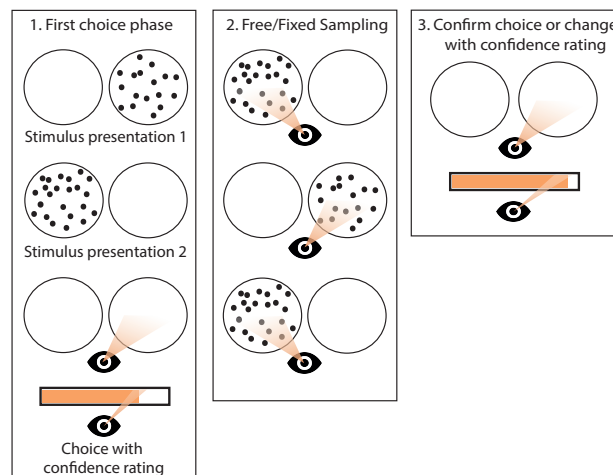
- Experiment 2 -

224   **Figure 3.** Task design and participant behaviour for experiment 2. (A) Task structure. Participants had to
225   choose which of two dot patches contained the most dots (phase 1) and rate their confidence in this choice.
226   Then participants were given 3000ms, 5000ms, or 7000ms to view the dots which they could allocate
227   between the two patches in whichever way they liked (phase 2) by looking inside the circles. Finally,
228   participants made a second choice about the same set of stimuli and rated their confidence again (phase 3).
229   (B) Participants effectively used the stimuli to make correct choices and improved upon their performance
230   on the second choice. This psychometric curve is plotting the probability of choosing the left option as a
231   function of the evidence difference between the two stimuli for each of the two choice phases. (C) In the
232   free sampling condition during the sampling phase (phase 2) participants spent more time viewing the
233   stimulus they chose on the preceding choice phase than the unchosen option. Data points represent
234   individual participants.

235

236

9

237 This new experiment also replicated the mediating effect of confidence on how
238 sampling bias affects subsequent choice. We again found a significant interaction
239 between choice and confidence (Figure 4A-B; $t_{16.97}=4.29$, $p<0.001$; see supplemental
240 materials S1) and replicated the main positive effects of choice and evidence difference
241 on sampling time difference (Figure 4B; main effect of choice: $t_{16.97}=2.90$, $p<0.01$; main
242 effect of evidence difference: $t_{16.97}=9.21$, $p<0.001$).
243

- Experiment 2 -

A Increased confidence leads to more bias

B Predicting sampling time difference

C Predicting change of mind

D Sampling bias only affects choice in free sampling trials
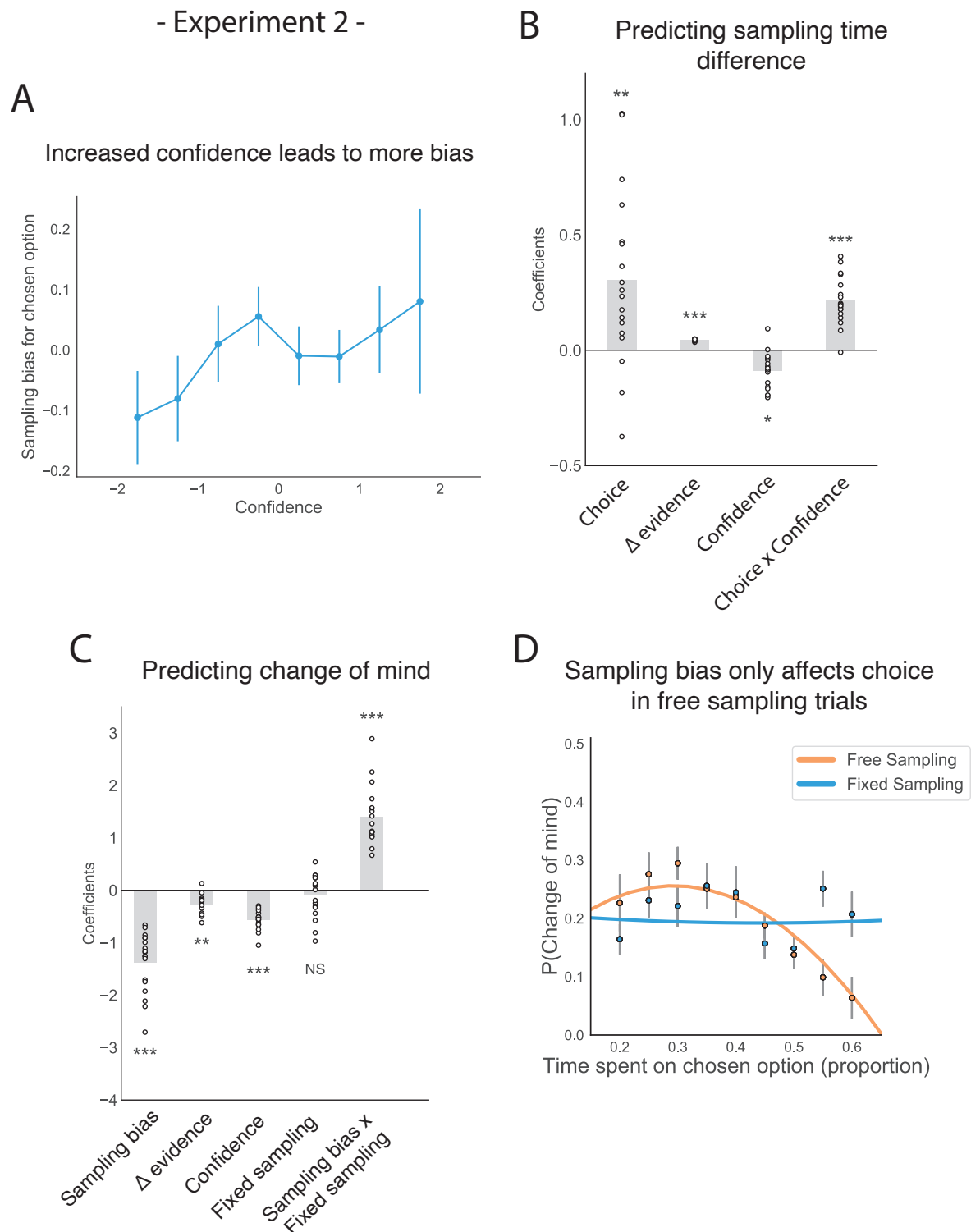
244
245

Figure 4. The effect of choice on sampling behaviour is mediated by confidence in experiment 2. Participants were less likely to change their mind if they showed a strong sampling bias for their initially chosen option in the sampling phase, but this was only the case in the free sampling condition. (A) Sampling bias in favour of the chosen option increases as a function of confidence in the initial choice. Confidence and sampling bias are both normalized in this plot. (B) There is a significant main effect of choice on sampling time difference, such that an option is sampled for longer if it was chosen, and a significant interaction effect of Choice x Confidence, such that options chosen with high confidence are sampled for even longer. (C) There is a main negative effect of sampling bias on change of mind, such that participants were less likely to change their mind in the second decision phase (phase 3) the more they sampled their initially chosen option in the free sampling phase (phase 2). The main effect of sampling bias on change of mind disappears in the fixed sampling condition, which can be seen by the positive interaction term Sampling bias x Fixed sampling which entirely offsets the main effect. The analysis includes a dummy variable 'Fixed Sampling' coding whether the trial was in the fixed-viewing condition. (B-C) Plotted are fixed-effect coefficients from hierarchical regression models predicting the sampling time (how long each patch was viewed in the sampling phase) difference between the left and right stimuli. Data points represent regression coefficients for each individual participant. (D) The probability that participants change their mind on the second choice phase is more likely if they looked more at the unchosen option during the sampling phase. The plot shows the probability that participants changed their mind as a function of the time spent sampling the initially chosen option during phase 2. The lines are polynomial fits to the data, while the data points indicate the frequency of changes of mind binned by sampling bias. Note that actual gaze time of the participants is plotted here for both task conditions. The same pattern can be seen when instead plotting the fixed presentation times of the stimuli for the fixed task condition (see Figure S4.1).

Similarly, we replicated the negative effect of sampling bias on subsequent change of mind (Figure 4C; z=-7.07, p<0.001; see S1) as well as the main negative effects of evidence difference and confidence on change of mind (Figure 4C; main effect of evidence difference: z=-2.71, p<0.01; main effect of confidence: z=-8.86, p<0.001).

Once we confirmed all the main findings from the first experiment using this new setup, we were able to test our main hypothesis: does the effect of sampling bias on choice we identified require the participant to actively choose which information to sample? In other words, is the effect of confirmation bias on subsequent choice present when confirmatory evidence is passively presented to participants or does it require that confirmatory evidence is actively sought by the decision-maker? In the first case, we would expect to see the same effect in the fixed-viewing condition (in which asymmetric information is provided by the experimenter) as in the free-sampling condition. In the second case, we would expect that the effect of biased information sampling on the subsequent choice disappears in the fixed-viewing condition.

In line with the second prediction, we found that in the fixed-viewing condition, contrary to the free-sampling condition in this experiment and in the first experiment, the amount of time spent viewing the patches in the sampling phase did not significantly affect subsequent choice. In a multinomial hierarchical logistic regression predicting change of mind from the first choice phase to the second choice phase within a given trial, the main effect of sampling bias on change of mind was completely offset by the positive effect of the interaction term between sampling bias and a dummy variable that was set to 1 if a trial was in the fixed-viewing condition (Figures 4C-D; z=7.24, p<0.001; see supplemental materials S2). This means that there was no effect of sampling bias on change of mind in the fixed-viewing condition. To check that participants were engaging in this version of the task, we looked whether the number of saccades made *within* each patch during the

298   sampling phase was similar between the two tasks. We found that the number of saccades
299   was actually higher in the fixed-viewing condition than in the main experiment (t_{17}=-4.22,
300   p<0.001), which means participants were indeed exploring the information in this
301   condition. This confirms that active information sampling by the participant is crucial to
302   the effect of confirmation bias in sampling on future choices.

303

304   To further investigate how attention, when freely allocated, shapes the accumulation of
305   evidence and choice bias, we modelled the data from both viewing conditions using the
306   Gaze-weighted Linear Accumulator Model (GLAM; Molter et al., 2019; Sepulveda et al.,
307   2020; Thomas et al., 2019). GLAM belongs to the family of race models with an additional
308   modulation by visual attention (Figure 5A). It is an approximation of a widely used class
309   of models – the attentional Drift Diffusion Model (aDDM; Krajbich et al., 2010; Krajbich
310   and Rangel, 2011) in which the full dynamic sequence of fixations is replaced by the
311   percentage of time spent fixating the choice alternatives. Even-numbered trials were
312   used to fit the model while odd-numbered trials were used to test it. See the *Methods*
313   section for further details.

314

315   GLAM is defined by four free parameters: ν (drift term), γ (gaze bias), τ (evidence scaling),
316   and σ (normally distributed noise standard deviation). The model correctly captured the
317   reaction times (RT) and choice behaviour of the participants at group-level both in the
318   free-sampling (Figure 5B) and fixed-viewing conditions (Figure 5C). More specifically, we
319   found that the model predicted faster RTs when trial difficulty was low (|ΔDots| is high;
320   Figures 5B-C, top left). The model also reproduced overall choice behaviour as a function
321   of the number of dots in the patches (ΔDots = Dots$_{Left}$ – Dots$_{Right}$) in both conditions
322   (Figures 5B-C, top right). Furthermore, we found gaze allocation (ΔGaze = g$_{Left}$ – g$_{Right}$)
323   predicted the probability of choosing the correct patch in the free-sampling condition
324   (Figure 5C, bottom left). However, to properly test how predictive gaze allocation is of
325   choice, we must account for the effect of evidence (ΔDots) on choice. As such, we used the
326   gaze influence (GI) measure (Thomas et al., 2019), which reflects the effect of gaze on
327   choice after accounting for the effect of evidence on choice. GI is calculated taking the
328   actual choice (0 or 1 for right or left choice, respectively) and subtracting the probability
329   of choosing the left item as predicted by a logistic regression with ΔDots as a predictor
330   estimated from behaviour. The averaged residual choice probability reflects GI. We found
331   GI estimated purely from participant's behaviour was higher in the free-sampling than in
332   the fixed-viewing condition (comparing average GI by participant, free-sampling
333   condition: Mean = 0.148, SD = 0.169; fixed-viewing condition: Mean = 0.025, SD = 0.11;
334   t_{17} =2.202, p<0.05). This suggests the effect of visual attention on choice was higher in
335   the free-sampling condition. In line with this, the model also predicted a higher GI on
336   corrected choice probability in the free-sampling condition (comparing average GI by
337   individual model predictions, free-sampling condition: Mean = 0.112, SD = 0.106; fixed-
338   viewing condition: Mean = 0.041, SD = 0.055; t_{17} = 2.205, p<0.05; Figures 5B-C, bottom
339   right).

340

341   We then tested whether attention affected information integration more when
342   information was actively sought (i.e. the free-sampling condition) compared to when
343   information was given to the participants (i.e. the fixed-viewing condition). We compared

344   the parameters obtained from the individual fit in the free-sampling and fixed-viewing
345   conditions (Figure 5D). We found a significant variation in the gaze bias parameter (Mean
346   γ_{Free}= 0.81, Mean γ_{Fixed}= 0.96, t_{17}= -3.268; p<0.01), indicating a higher influence of gaze
347   on choice in the free-sampling condition. Note that during the fixed-viewing condition,
348   the parameter $\gamma \approx 1$ indicates that almost no gaze bias was present on those trials.
349   Conversely, there was no significant difference for the other parameters between the
350   conditions (Mean τ_{Free} =1.44, τ_{Fixed}=1.11, t_{17}=1.17; p=0.25, n.s.; Mean σ_{Free} =0.0077, Mean
351   σ_{Fixed}=0.0076, t_{17}=0.128; p=0.89, n.s.; ν_{Free}=8.07x10^{-5}, ν_{Fixed}=8.64x10^{-5}, t_{17}=-1.135; p =
352   0.27, n.s.). These results suggest that gathering information actively (i.e. free-sampling
353   condition) does not affect the overall speed at which information is integrated, but it
354   specifically modulates the likelihood of choosing the gazed-at option. Finally, to test that
355   the identified effect did not depend on a less variable gaze time range we resampled the
356   data from the free-sampling condition to match the gaze time range in the fixed-viewing
357   condition, and fitted the GLAM again. We replicated our finding even when the gaze time
358   range in the free-sampling condition was reduced to match that in the fixed-viewing
359   condition (Figure S5.1). For more in-depth analyses including model comparisons please
360   see the supplemental materials (S5).

361

362   Finally, we explored the idea that a sampling bias could arise from the use of previous
363   choice as evidence in its own right, in addition to previously acquired information.
364   Formalizing this idea, we developed an economic model that makes behavioural
365   predictions aligned with our findings (S6). In this model, after the first choice, the prior
366   that feeds into the subsequent active sampling phase is not the belief after the first
367   sampling phase, but rather it is a convex combination of this with the belief that the
368   individuals would hold if they only knew of the previous choice. Subsequent active
369   sampling then varies as a strictly increasing function of this prior. As a result, a sampling
370   bias in favour of the initially chosen option arises, and this bias varies as a function of
371   confidence in the initial choice, as seen in our data. This economic model provides a
372   formal description of the behaviour we observed. At the same time, it suggests that
373   seeking for confirmatory evidence might not arise from a simple low-level heuristic, but
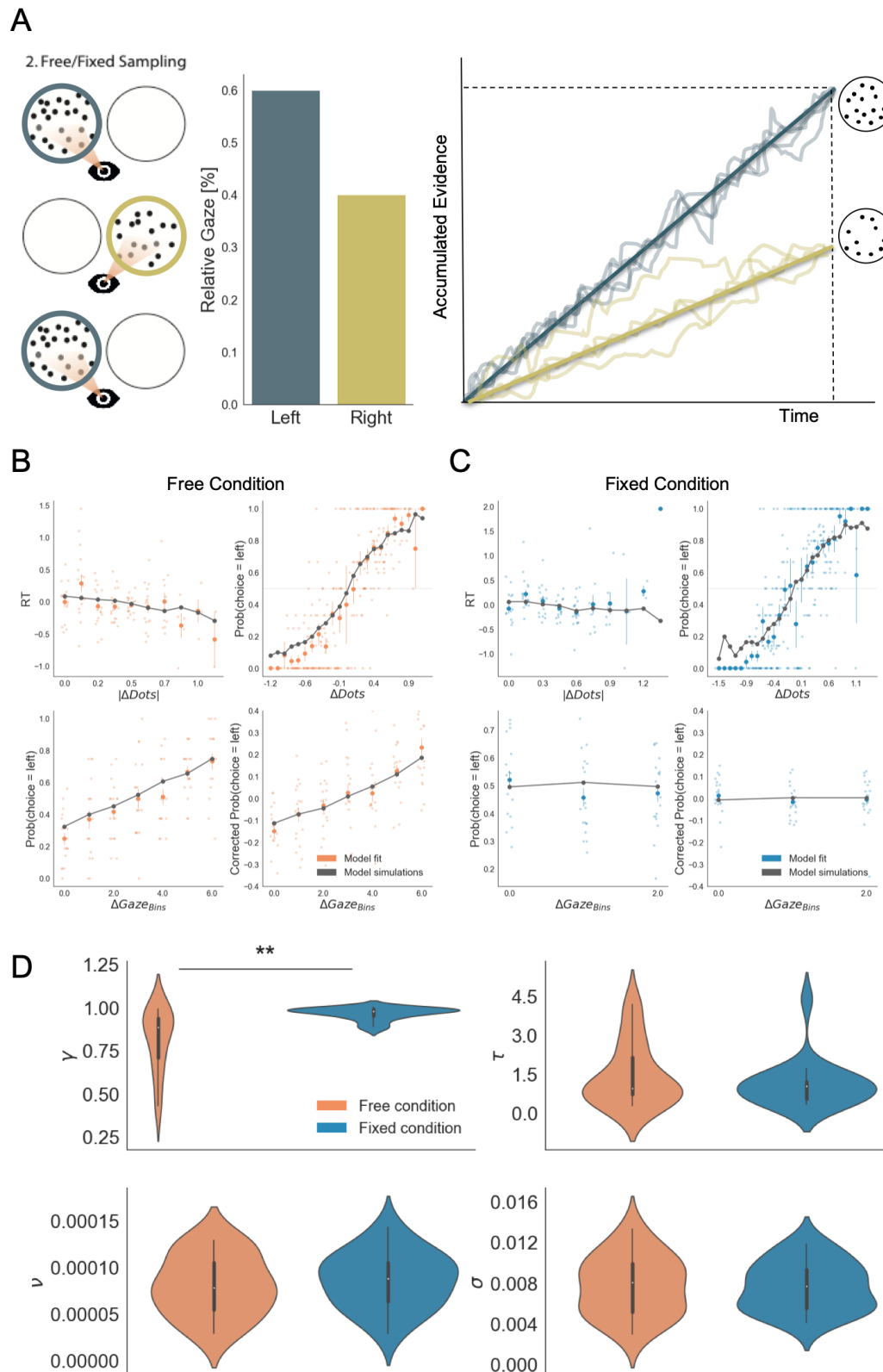374   is rooted in the way information is acquired and processed (for more details, see S6).

375

376

Figure 5. **Gaze impacted evidence accumulation (for the 2nd choice) more strongly in the free than in the fixed sampling condition.** (A) Free and fixed sampling condition trials were fitted separately using a Gaze-weighted Linear Accumulator Model (GLAM). In this model there are two independent accumulators for each option (left and right) and the decision is made once one of them reaches a threshold. Accumulation rate was modulated by gaze times, when gaze bias parameter is lower than 1 ($\gamma < 1$). In that

case, the accumulation rate will be discounted depending on γ and the relative gaze time to the items, within the trials. Gaze information from the free sampling trials, and presentation times from the fixed sampling trials were used to fit the models. The panel depicts an example trial: patch sampling during phase 2 (left panel) is used to estimate the relative gaze for that trial (central panel), and the resulting accumulation process (right panel) Notice GLAM ignores fixations dynamics and uses a constant accumulation term within trial (check *Methods* for further details). The model predicted the behaviour in free (B) and fixed (C) sampling conditions. The four panels present four relevant behavioural relationships comparing model predictions and overall participant behaviour: (top left) response time was faster (shorter RT) when the choice was easier (i.e. bigger differences in the number of dots between the patches); (top right) probability of choosing the left patch increased when the number of dots was higher in the patch at the left side (ΔDots = $Dots_{Left}$ – $Dots_{Right}$); (bottom left) the probability of choosing an alternative depended on the gaze difference ((ΔGaze = g_{Left} – g_{Right}); and (bottom right) the probability of choosing an item that was fixated longer than the other, corrected by the actual evidence ΔDots, depicted a residual effect of gaze on choice. Notice that in the free condition, the model predicted an effect of gaze on choice in a higher degree than in the fixed condition. Solid dots depict the mean of the data across participants in both conditions. Lighter dots present the mean value for each participant across bins. Solid grey lines show the average for model simulations. Data z-scored/binned for visualisation. (D) GLAM parameters fitted at participant level for free and fixed sampling conditions. Free sampling condition presented a higher gaze bias than fixed sampling, while no significant differences were found for the other parameters. γ: gaze bias; τ: evidence scaling; ν: drift term; σ: noise standard deviation. **: p<0.01

Discussion

In this work we have demonstrated that a form of confirmation bias exists in active information sampling, and not just in information weighting as previously shown. Using two novel experiments we showed that this effect is robust for simple perceptual choice. Critically we show that this sampling bias affects future choice. Furthermore, we demonstrated that this effect is only present in the free-sampling condition, showing that agency is essential for biased sampling to have an effect on subsequent choice.

Preference for confirmatory evidence has previously mostly been studied in the context of strongly held political, religious or lifestyle beliefs, and not in perceptual decision-making (Bakshy et al., 2015; Hart et al., 2009; Lord et al., 1979; Nickerson, 1998; Stroud, 2008; Wason, 1960, 1968). Our results, together with the recent work of others (Talluri, Urai et al., 2018; Palminteri et al., 2017; Lefebvre et al., 2020), show that confirmation bias in information search is present even in decisions where the beliefs formed by participants are not meaningful to their daily lives. This suggests that confirmation bias might be a fundamental property of information sampling, existing irrespective of how important the belief is to the agent.

Recent findings suggest that biases in information sampling might arise from Pavlovian approach, a behavioural strategy that favours approaching choice alternatives associated with reward (Hunt et al., 2016; Rutledge et al., 2015). Furthermore, the number of hypotheses an individual can consider in parallel is likely to be limited (Tweney et al., 1980). As such, it may be advantageous to first attempt to rule out the dominant hypothesis before going on to sample from alternative options. In this vein, the sampling bias we see could be the solution to an exploit-explore dilemma in which the decision-maker must decide when to stop 'exploiting' a particular hypothesis and instead 'explore' different ones.

434 A key novel feature of this task is that participants were able to freely sample information
435 between choice phases, providing a direct read-out of confirmation bias in the active
436 sampling decisions made by the participants. Previous accounts of confirmation bias in
437 perceptual choice have instead focused on altered weighting of passively viewed
438 information as a function of previous choice (Bronfman et al., 2015; Rollwage et al., 2020;
439 Talluri, Urai et al., 2018). However, from these findings it remained unclear to what extent
440 this bias manifests in the processing of information compared to active sampling of
441 information. Our findings show that active information sampling plays a key role in the
442 amplification of beliefs from one decision to the next, and that changes in evidence
443 weighting likely only account for part of observed confirmation bias effects.

445 We show that confidence modulated the confirmation bias effect, such that choices made
446 with higher confidence, led to increased sampling of the chosen option and an increased
447 likelihood of choosing the same option again in the second choice phase. This shows that
448 the strength with which a belief is held determines the size of the confirmation bias in
449 active information sampling. Confidence has previously been shown to affect the
450 integration of confirmatory evidence as reflected in MEG recordings of brain activity
451 during evidence accumulation (Rollwage et al., 2020). Furthermore, recent work in
452 economics and neuroscience have given theoretical and experimental proof of a
453 relationship between overconfidence, and extreme political beliefs (Ortoleva &
454 Snowberg, 2015; Rollwage et al., 2018). Our results suggest altered information sampling
455 could be the missing link between these two phenomena. Specifically, given that we have
456 shown that increased confidence leads to increased confirmation bias, it follows that
457 overconfidence in a belief would lead to increased sampling of confirmatory evidence in
458 line with that belief, which in turn would lead to even higher confidence.

460 We also show that it is not sufficient that agents freely make a choice for that choice to
461 bias future decisions. We do not see an effect of the first choice on subsequent choices
462 when the intermediate sampling phase is fixed by the experimenter, suggesting the
463 integration of information into beliefs is dependent on whether the agent has actively
464 chosen to sample this information. In other words, a sense of agency appears to impact
465 the extent to which new information influences future choice (Hassall et al., 2019;
466 Chambon et al., 2020), and making a choice may not necessarily lead to confirmation bias
467 if it is not followed by active decisions to sample evidence in line with that choice. Our
468 results are in line with Chambon et al. (2020), in which confirmation bias in learning rates
469 was only present when participants were able to freely choose between stimuli and not
470 when these choices were fixed by the experimenter. In their task, though, choice and
471 information gain were not separated by design, meaning information sampling was only
472 possible from outcome feedback after a choice was made. Our results expand on these
473 findings, by showing that choice commitment also affects subsequent decisions to
474 sample, and that the resulting biased sampling influences future choices. This means that
475 active information sampling is likely to play a central role in the propagation of
476 confirmation bias across decisions, as seen in our descriptive economic model (S6).

478 Additionally, the results from the GLAM model show that, in line with previous studies
479 (Krajbich et al., 2010; Krajbich & Rangel, 2010; Sepulveda et al., 2020; Tavares et al.,
480 2017; Thomas et al. 2019), a specific boost in the accumulation of evidence of the visually
481 attended items was found in the free sampling condition. Conversely, a disconnection
482 between an item's sampling time and evidence accumulation was found in the fixed

16

483   condition (i.e. the absence of gaze bias in GLAM implies that visual fixations did not affect
484   evidence integration when the sampling was not controlled by the decision-maker). One
485   explanation for this result is that attentional allocation itself is directed towards the
486   options that the participants reckon more relevant for the task to be performed
487   (Sepulveda et al., 2020). In our experiment the goal of the task was theoretically identical
488   for the first and second choice (i.e. to find the patch with more dots). However, in
489   agreement with the ideas of our descriptive economic model, it could be that participants
490   perceived the goal of the second choice to be slightly different from the goal of the first:
491   in the second case they had to verify whether their initial choice was correct, as well as
492   finding the patch with the most dots. This immediately bestowed the chosen option with
493   higher relevance and then more attention (consequently boosting evidence
494   accumulation). On the other hand, since in the fixed sampling condition participant's
495   attention allocation is not necessarily associated with their goals, the difference in display
496   time of the items is ignored or cannot be consistently integrated in their decision process.
497   An alternative hypothesis is that gaze allocation itself is used as additional evidence in
498   favour of a choice alternative in a way similar to previous choice as in our descriptive
499   model (S6). This would mean that when an agent has previously attended to a choice
500   alternative, this is used as evidence in favour of that option in and of itself. Therefore,
501   when gaze allocation is not under the agent's control, as in the fixed-viewing condition, it
502   is not used to inform choice.

503

504   Our findings imply that agency plays a clear role in evidence accumulation, and
505   consequently in confirmation bias. It also suggests that common experimental designs in
506   which information is provided by the experimenter and passively sampled by the
507   participant might not be an ecological way to study decision-making. These tasks mask
508   the potentially large effect of active information search on belief formation and choice.
509   Pennycook et al. (2021) recently showed that people were less likely to share false
510   information online if they had been asked to rate the accuracy of a headline just
511   previously. It may therefore be possible to reduce confirmation bias in information
512   search in a similar way by priming participants to attend more to accuracy instead of
513   confirmatory evidence. More complex behavioural tasks are required to improve our
514   understanding of the different drivers of information sampling and how sampling in turn
515   guides future choice (Gottlieb & Oudeyer, 2018).

516

517   To summarise our findings, we observed that participants sampled more information
518   from chosen options in a perceptual choice paradigm and that this sampling bias
519   predicted subsequent choice. Asymmetric sampling in favour of the chosen alternative
520   was stronger the higher participants' confidence in this choice. Furthermore, the effect of
521   information on subsequent choice was only seen in a version of the task where
522   participants could sample freely, suggesting agency plays an important role in the
523   propagation of strongly held beliefs over time. These findings suggest that confirmatory
524   information processing might stem from a general information sampling strategy used to
525   seek information to strengthen prior beliefs rather than from altered weighting during
526   evidence accumulation only, and that active sampling is essential to this effect. Biased
527   sampling may cause a continuous cycle of belief reinforcement that can be hard to break.
528   Improving our understanding of this phenomenon can help us better explain the roots of
529   extreme political, religious and scientific beliefs in our society.

530

Methods

Participants

Experiment 1. 30 participants took part in this study. We excluded two participants from the analysis, because they gave the highest possible confidence rating on more than 75% of trials. Participants received a £10 show-up fee in addition to a monetary reward between £0 and £6, which could be gained in the task. Participants had normal or corrected-to-normal vision, and no psychiatric or neurological disorders. We obtained written informed consent from all participants before the study. This experiment was approved by the University of Cambridge Psychology Research Ethics Committee.

Experiment 2. 23 participants completed this experiment, of which two were excluded from the analysis, because they gave the highest possible confidence rating on more than 75% of trials. Another three participants were excluded because their confidence was a poor predictor of their accuracy in the task (lower than two standard deviations under the mean coefficient predicting accuracy in a logistic regression). Participants were reimbursed £30 for their time as well as an additional amount between £0 and £20 that could be gained in the task. Participants had normal or corrected-to-normal vision, and no psychiatric or neurological disorders. We obtained written informed consent from all participants before the study. This experiment was approved by the University of Cambridge Psychology Research Ethics Committee.

Behavioural Task

Experiment 1. The computer task used in this study consisted of 200 trials. Participants were asked to make binary choices between dot patches using the arrow keys on the keyboard. They then reported their confidence in this decision on a continuous rating scale. In the subsequent sampling phase, participants were given the opportunity to look at the patches again for 4000ms. By using the left and right arrow keys they could alternate as often as they wanted between viewing either dot patch within the given time. Afterwards, participants were prompted to make a second choice for the same pair of patches and give another confidence judgment. Participants were constantly reminded of their initial choice by a green tick next to the chosen patch. Participants received no feedback about the correctness of their choice. None of the choices made or confidence ratings given were time-constrained. One of the patches always contained 50 dots, and the other a variable amount of dots. We calibrated the dot difference between the two patches such that accuracy level within each participant was maintained at 70% throughout the task using a one-up two-down staircase procedure (Fleming & Lau, 2014) The task was programmed using Python 2.7.10 and PsychoPy (Peirce, 2007).

Experiment 2. A similar task was used in the second experiment, except that in this task participants' responses were elicited through eye movements. To make a choice they looked at one of the patches and to rate their confidence they looked at a position inside a rating scale. The sampling phase between the two choices randomly varied in this task, and was either 3000, 5000, or 7000ms. This was not cued to the participants, so at the start of a sampling phase they did not know how much time they would be given to sample the patches. Sampling was gaze-contingent such that the dots were only visible when the participant fixated inside a patch and participants could only see one patch at a time. Furthermore, we introduced a control condition in which participants were not free to sample the circles however they liked during the sampling phase. Instead, in one-third

580   of trials the patches were shown for an equal amount of time each and in two-thirds of
581   trials one patch was shown three times longer than the other (50% of trials the left was
582   shown longer, 50% of trials the right was shown longer). Participants were constantly
583   reminded of their initial choice by the circle surrounding the chosen patch changing color.
584   The first experiment conducted using this task was a behavioural experiment in which
585   participants recorded their responses using the arrow keys on a computer. In the second
586   experiment, participants responded using eye movements. In the second experient,
587   participants took part in two sessions, each consisting of 189 trials. In one session they
588   performed the main task and in the other the control condition of the task. The order of
589   these two sessions was pseudo-random. This experiment was programmed using the SR
590   Research Experiment Builder version 1.10.1630 (*SR Research Experiment Builder*, 2017).
591
592   *Eye-tracking*
593   In experiment 2, we recorded eye movements at a rate of 1,000Hz using an EyeLink 1000
594   Plus eye-tracker. Areas of Interest (AI) for the eye tracking analyses were pre-defined as
595   two squares centered on the gaze-contingent circles in the experiment. The sides of the
596   squares were the same as the diameter of the gaze-contingent circles. For each decision
597   period we derived the total dwell time in each AI from the eye-tracking data. The
598   computer used in this experiment had a screen size of 68.58 x 59.77cm and participants
599   were seated 60cm away from the screen.
600
601   *Analyses*
602   We studied the effect of choice on the time spent on each of the two stimuli using paired
603   sample t-tests on the mean sampling times spent on each stimulus from each participant.
604   Trials with the shortest sampling phase length of 3000ms in experiment 2 were excluded
605   from all analyses, because it became apparent that this time was too short for participants
606   to be able to saccade to each circle more than once.
607
608   **Hierarchical models.** Hierarchical regression models were conducted using the lme4
609   package in R (Bates et al., 2015; Gelman & Hill, 2006). We computed degrees of freedom
610   and p-values with the Kenward-Roger approximation, using the package pbkrtest
611   (Halekoh & Hojsgaard, 2014). We predicted the sampling time difference between the
612   two circles using a hierarchical linear regression model. To predict choice in the second
613   choice phase, hierarchical logistic regressions were used predicting the log odds ratio of
614   picking the left circle on a given trial. Confidence and sampling time were z-scored on the
615   participant level. For detailed results and model comparisons, see supplemental
616   materials S1-S2, S7.
617
618   **Attentional model - GLAM.** The Gaze-weighted Linear Accumulator Model (Thomas et
619   al., 2019; Molter et al., 2019) is part of the family of linear stochastic race models in which
620   different alternatives (i; left or right) accumulate evidence (E_i) until a decision threshold
621   is reached by one of them, determining the chosen alternative. The accumulator for an
622   individual option was defined by the expression:
623
624   $$E_i(t) = E_i(t-1) + vR_i + \epsilon_t \ with \ \epsilon_t \sim N(0, \sigma) \ and \ E_i(t=0) = 0 \quad (1)$$
625
626   A drift term (v) controlled the speed of relative evidence (R_i) integration and noise was
627   integrated with normal distribution (zero-centered and with standard deviation, σ). R_i
628   expressed the amount of evidence that was accumulated for item i at each time point t.

629   This was calculated as follows. We denote by g_i, the relative gaze term, calculated as the
630   proportion of time that participants observed item i:

631
632

633    $g_i = \dfrac{DT_i}{DT_1 + DT_2}$ (2)

634

635   with DT as the dwelling time for item i during an individual trial. Let r_i denote the
636   evidence (in this study evidence corresponded to number of dots presented for each
637   option) for item i. We can then define the average absolute evidence for each item (A_i)
638   during a trial:

639

640    $A_i = g_i r_i + (1 - g_i)\gamma r_i$ (3)

641

642   This formulation considers a multiplicative effect of the attentional component over the
643   evidence, capturing different rates of integration when the participant was observing
644   item i or not (unbiased and biased states, respectively). The parameter γ was the gaze
645   bias parameter: it controlled the weight that the attentional component had in
646   determining absolute evidence. When $\gamma = 1$, accumulation was the same in biased and
647   unbiased states, i.e. gaze did not affect the accumulation of evidence. When $\gamma < 1$, A_i was
648   discounted for the biased condition, resulting in higher probability of choosing items that
649   had been gazed at longer. When $\gamma < 0$, the model assumed a leak of evidence when the item
650   was not fixated. Therefore, the relative evidence of item i, R_i^*, was given by:

651

652    $R_i^* = A_i - \max_{j \neq i} A_j = A_i - A_j \rightarrow R_{right}^* = -R_{left}^*$ (4)

653

654   Since our experiment considers a binary choice, while the original formulation of the
655   model (Thomas et al., 2019) proposed for more than two alternatives, R_i^* was reduced to
656   subtract the average absolute evidence of the other item. Therefore, for the binary case,
657   R_i^* for one item was the additive inverse of the other. For example, if the left item had the
658   lower evidence, we would have $R_{left}^* < 0$ and $R_{right}^* > 0$. The difference in signals captured
659   by R_i^* was scaled by a logistic transformation. The model assumed an adaptive
660   representation of the relative decision signals, which is maximally sensitive to marginal
661   differences in absolute decision signals:

662

663    $R_i = \dfrac{1}{1 + \exp(-\tau R_i^*)}$ (5)

664

665   The temperature parameter τ of the logistic function controlled the sensitivity of the
666   transformation. Larger values of τ indicate stronger sensitivity to small differences in
667   absolute evidence (A_i). Given that R_i represents an average of the relative evidence across
668   the entire trial, the drift rate in E_i can be assumed to be constant, which enables the use
669   of an analytical solution for the passage of time density (for details see Thomas et al.,
670   2019, Molter et al., 2019). Notice that unlike the aDDM (Krajbich et al., 2010), GLAM does
671   not deal with the dynamics of attentional allocation process in choice. In summary, the
672   GLAM model considers four free parameters: ν (drift term), γ (gaze bias), τ (evidence
673   scaling), and σ (normally distributed noise standard deviation).

674 The model fit with GLAM was implemented at a participant level in a Bayesian framework
675 using PyMC3 (Salvatier et al., 2016). Uniform priors were used for all the parameters:

676 $v \sim \text{Uniform}(1^{-6}, 0.01)$
677 $\gamma \sim \text{Uniform}(-1, 1)$
678 $\sigma \sim \text{Uniform}(1^{-6}, 5)$
679 $\tau \sim \text{Uniform}(0, 5)$
680
681 We fitted the model for each individual participant and for free and fixed sampling
682 conditions in experiment 2 separately. To model participants' behaviour, we used as
683 input for GLAM the reaction times (RT) and choices obtained from phase 3, and relative
684 gaze for left and right alternatives for each trial during sampling phase 2. For fixed
685 sampling trials the presentation times of the dot patches were used to calculate the
686 relative gaze time. For both conditions, model fit was performed only on even-numbered
687 trials using Markov-Chain-Monte-Carlo sampling, we used implementation for No-U-
688 Turn-Sampler (NUTS), four chains were sampled, 1000 tuning samples were used, and
689 2000 posterior samples were used to estimate the model parameters. The convergence
690 was diagnosed using the Gelman-Rubin statistic ($|\hat{R} - 1| < 0.05$) for the four parameters
691 (v, γ, σ, and τ). Considering all the individual models (18 participants), we found
692 divergences in ~16% of the estimated parameters. We removed the participants that
693 presented divergent parameters (4 participants) to check whether the results we found
694 were driven by these data. The significantly higher gaze bias in free-viewing condition
695 was maintained even after removing these participants (see supplemental materials,
696 S5.4). Model comparison was performed using Watanabe-Akaike Information Criterion
697 (WAIC) scores available in PyMC3, calculated for each individual participant fit.
698
699 To check how well the model replicates the behavioural effects observed in the data
700 (Palminteri et al., 2017), simulations for choice and RT were performed using
701 participants' odd trials, each one repeated 50 times. For each trial, number of dots and
702 relative gaze for left and right items were used together with the individually fitted GLAM
703 parameters to simulate the trials. Random choice and RT (within a range of the minimum
704 and maximum RT observed for each particular participant) were set for 5% of the
705 simulations, replicating the contaminating process included in the model as described by
706 Thomas et al., 2019.
707
708
709 *Acknowledgements*
714

References

Bakshy, E., Messing, S., & Adamic, L. A. (2015). Exposure to ideologically diverse news and opinion on Facebook. *Science*, *348*(6239), 1130–1132. https://doi.org/10.1126/science.aaa1160

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious Mixed Models. *ArXiv*. http://arxiv.org/abs/1506.04967

Bennett, W. L., & Iyengar, S. (2008). A New Era of Minimal Effects? The Changing Foundations of Political Communication. *Journal of Communication*, *58*(4), 707–731. https://doi.org/10.1111/j.1460-2466.2008.00410.x

Bogacz, R., Wagenmakers, E.-J., Forstmann, B. U., & Nieuwenhuis, S. (2010). The neural basis of the speed–accuracy tradeoff. *Trends in Neurosciences*, *33*(1), 10–16. https://doi.org/10.1016/j.tins.2009.09.002

Boldt, A., Blundell, C., & De Martino, B. (2019). Confidence modulates exploration and exploitation in value-based learning. *Neuroscience of Consciousness*, *2019*(1), niz004. https://doi.org/10.1093/nc/niz004

Bronfman, Z. Z., Brezis, N., Moran, R., Tsetsos, K., Donner, T., & Usher, M. (2015). Decisions reduce sensitivity to subsequent information. *Proceedings of the Royal Society B: Biological Sciences*, *282*(1810), 20150228. https://doi.org/10.1098/rspb.2015.0228

Carrasco, M., Ling, S., & Read, S. (2004). Attention alters appearance. *Nature Neuroscience*, *7*(3), 308–313. https://doi.org/10.1038/nn1194

Chambon, V., Théro, H., Vidal, M., Vandendriessche, H., Haggard, P., & Palminteri, S. (2020). Information about action outcomes differentially affects learning from self-determined versus imposed choices. *Nature Human Behaviour*, *4*(10), 1067–1079. https://doi.org/10.1038/s41562-020-0919-5

740   Desender, K., Murphy, P., Boldt, A., Verguts, T., & Yeung, N. (2019). A Postdecisional Neural

741   Marker of Confidence Predicts Information-Seeking in Decision-Making. *The Journal*

742   *of Neuroscience*, *39*(17), 3309–3319. https://doi.org/10.1523/JNEUROSCI.2620-

743   18.2019

744   Eliaz, K., & Schotter, A. (2007). Experimental Testing of Intrinsic Preferences for

745   NonInstrumental Information. *American Economic Review*, *97*(2), 166–169.

746   https://doi.org/10.1257/aer.97.2.166

747   Fleming, S. M., & Lau, H. C. (2014). How to measure metacognition. *Frontiers in Human*

748   *Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00443

749   Folke, T., Jacobsen, C., Fleming, S. M., & Martino, B. D. (2017). 1 Explicit representation of

750   confidence informs future 2 value-based decisions. *Nature Human Behaviour*,

751   *1*(0002).

752   Gelman, A., & Hill, J. (2006). *Data Analysis Using Regression and Multilevel/Hierarchical*

753   *Models*. Cambridge University Press.

754   Gesiarz, F., Cahill, D., & Sharot, T. (2019). Evidence accumulation is biased by motivation: A

755   computational account. *PLOS Computational Biology*, *15*(6), e1007089.

756   https://doi.org/10.1371/journal.pcbi.1007089

757   Gottlieb, J., & Oudeyer, P.-Y. (2018). Towards a neuroscience of active sampling and

758   curiosity. *Nature Reviews Neuroscience*, *19*(12), 758–770.

759   https://doi.org/10.1038/s41583-018-0078-0

760   Halekoh, U., & Hojsgaard, S. (2014). A Kenward-Roger Approximation and Parametric

761   Bootstrap Methods for Tests in Linear Mixed Models—The R Package pbkrtest.

762   *Journal of Statistical Software*, *59*(9), 1–30.

763   Hart, W., Albarracín, D., Eagly, A. H., Brechan, I., Lindberg, M. J., & Merrill, L. (2009). Feeling

764   validated versus being correct: A meta-analysis of selective exposure to information.

765   *Psychological Bulletin*, *135*(4), 555–588. https://doi.org/10.1037/a0015701

766   Hassall, C. D., Hajcak, G., & Krigolson, O. E. (2019). The importance of agency in human

767   reward processing. *Cognitive, Affective, & Behavioral Neuroscience*, *19*(6), 1458–

768   1466. https://doi.org/10.3758/s13415-019-00730-2

769   Hunt, L. T., Malalasekera, W. M. N., de Berker, A. O., Miranda, B., Farmer, S. F., Behrens, T.

770   E. J., & Kennerley, S. W. (2018). Triple dissociation of attention and decision

771   computations across prefrontal cortex. *Nature Neuroscience*, *21*, 1471–1481.

772   Hunt, L. T., Rutledge, R. B., Malalasekera, W. M. N., Kennerley, S. W., & Dolan, R. J. (2016).

773   Approach-Induced Biases in Human Information Sampling. *PLOS Biology*, *14*(11),

774   e2000638. https://doi.org/10.1371/journal.pbio.2000638

775   Kobayashi, K., & Hsu, M. (2019). Common neural code for reward and information value.

776   *Proceedings of the National Academy of Sciences*, *116*(26), 13061–13066.

777   https://doi.org/10.1073/pnas.1820145116

778   Kobayashi, K., Ravaioli, S., Baranès, A., Woodford, M., & Gottlieb, J. (2019). Diverse motives

779   for human curiosity. *Nature Human Behaviour*, *3*(6), 587–595.

780   https://doi.org/10.1038/s41562-019-0589-3

781   Krajbich, I., & Rangel, A. (2011). Multialternative drift-diffusion model predicts the

782   relationship between visual fixations and choice in value-based decisions.

783   *Proceedings of the National Academy of Sciences*, *108*(33), 13852-13857.

784   Krajbich, I., Armel, C., & Rangel, A. (2010). Visual fixations and the computation and

785   comparison of value in simple choice. *Nature Neuroscience*, *13*(10), 1292-1298.

786   Lefebvre, G., Summerfield, C., & Bogacz, R. (2020). A normative account of confirmatory

787   biases during reinforcement learning. *bioRxiv*. doi:

788   https://doi.org/10.1101/2020.05.12.090134

789   Lord, C. G., Ross, L., & Lepper, M. R. (1979). Biased Assimilation and Attitude Polarization:

790   The Effects of Prior Theories on Subsequently Considered Evidence. *Journal of*

791   *Personality and Social Psychology*, *37*(11), 2098–2109.

792   Meyer, R. J., & Shi, Y. (1995). Sequential Choice Under Ambiguity: Intuitive Solutions to the

793   Armed-Bandit Problem. *Management Science*, *41*(5), 817–834.

794   https://doi.org/10.1287/mnsc.41.5.817

795   Molter, F., Thomas, A. W., Heekeren, H. R., & Mohr, P. N. (2019). GLAMbox: A Python

796   toolbox for investigating the association between gaze allocation and decision

797   behaviour. *PloS One*, *14*(12), e0226428.

798   Nickerson, R. S. (1998). Confirmation Bias: A Ubiquitous Phenomenon in Many Guises.

799   *Review of General Psychology*, *2*(2), 175–220.

800   Ortoleva, P., & Snowberg, E. (2015). Overconfidence in Political Behavior. *American*

801   *Economic Review*, *105*(2), 504–535.

802   Palminteri, S., Lefebvre, G., Kilford, E. J., & Blakemore, S.-J. (2017). Confirmation bias in

803   human reinforcement learning: Evidence from counterfactual feedback processing.

804   *PLOS Computational Biology*, *13*(8), e1005684.

805   https://doi.org/10.1371/journal.pcbi.1005684

806   Palminteri, S., Wyart, V., & Koechlin, E. (2017). The importance of falsification in

807   computational cognitive modeling. *Trends in cognitive sciences*, *21*(6), 425-433.

808   Pariser, E. (2012). *The Filter Bubble: What The Internet Is Hiding From You*. Penguin Press.

809   Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience*

810   *Methods*, 6.

811   Pennycook, G., Epstein, Z., Mosleh, M., Arechar, A.A., Eckles, D., & Rand, D.G. (2021).

812   Shifting attention to accuracy can reduce misinformation online. *Nature*.

813   https://doi.org/10.1038/s41586-021-03344-2

814   Pouget, A., Drugowitsch, J., & Kepecs, A. (2016). Confidence and certainty: Distinct

815   probabilistic quantities for different goals. *Nature Neuroscience*, *19*(3), 366–374.

816   https://doi.org/10.1038/nn.4240

817   Rabbitt, P. M. (1966). Errors and error correction in choice-response tasks. *Journal of*

818   *Experimental Psychology*, *71*(2), 264–272.

819   Rodriguez Cabrero, J. A. M., Zhu, J.-Q., & Ludvig, E. A. (2019). Costly curiosity: People pay a

820   price to resolve an uncertain gamble early. *Behavioural Processes*, *160*, 20–25.

821   https://doi.org/10.1016/j.beproc.2018.12.015

822   Rollwage, M., Dolan, R. J., & Fleming, S. M. (2018). Metacognitive Failure as a Feature of

823   Those Holding Radical Beliefs. *Current Biology*, *28*(24), 4014-4021.e8.

824   https://doi.org/10.1016/j.cub.2018.10.053

825   Rollwage, M., Loosen, A., Hauser, T. U., Moran, R., Dolan, R. J., & Fleming, S. M. (2020).

826   Confidence drives a neural confirmation bias. *Nature Communications*, *11*(1), 2634.

827   https://doi.org/10.1038/s41467-020-16278-6

828   Rutledge, R. B., Skandali, N., Dayan, P., & Dolan, R. J. (2015). Dopaminergic Modulation of

829   Decision Making and Subjective Well-Being. *Journal of Neuroscience*, *35*(27), 9811–

830   9822. https://doi.org/10.1523/JNEUROSCI.0702-15.2015

831   Salem-Garica, N., Palminteri, S., & Lebreton, M. (2021). The computational origins of

832   confidence biases in reinforcement learning. *PsyArXiv*. doi: 10.31234/osf.io/k7w38

833   Salvatier, J., Wiecki, T. V., & Fonnesbeck, C. (2016). Probabilistic programming in Python

834   using PyMC3. *PeerJ Computer Science*, *2*, e55.

835   Schulz, E., & Gershman, S. J. (2019). The algorithmic architecture of exploration in the

836   human brain. *Current Opinion in Neurobiology*, *55*, 7–14.

837   https://doi.org/10.1016/j.conb.2018.11.003

838   Schwartenbeck, P., Passecker, J., Hauser, T. U., FitzGerald, T. H., Kronbichler, M., & Friston,

839   K. J. (2019). Computational mechanisms of curiosity and goal-directed exploration.

840   *ELife*, *8*, e41703. https://doi.org/10.7554/eLife.41703

841   Sepulveda, P., Usher, M., Davies, N., Benson, A., Ortoleva, P., & De Martino, B. (2020). Visual

842   attention modulates the integration of goal-relevant evidence and not value. *ELife*, *9*,

843   e60705. https://doi.org/10.1101/2020.04.14.031971

844   Sharot, T. (2011). The optimism bias. *Current Biology*, *21*(23), R941–R945.

845   https://doi.org/10.1016/j.cub.2011.10.030

846   *SR Research Experiment Builder* (1.10.1630). (2017). [Computer software]. SR Research Ltd.

847   Steyvers, M., Lee, M. D., & Wagenmakers, E.-J. (2009). A Bayesian analysis of human

848   decision-making on bandit problems. *Journal of Mathematical Psychology*, *53*(3),

849   168–179. https://doi.org/10.1016/j.jmp.2008.11.002

850   Stroud, N. J. (2008). Media Use and Political Predispositions: Revisiting the Concept of

851   Selective Exposure. *Political Behavior*, *30*(3), 341–366.

852   https://doi.org/10.1007/s11109-007-9050-9

853   Talluri, B. C., Urai, A. E., Tsetsos, K., Usher, M., & Donner, T. H. (2018). Confirmation Bias

854   through Selective Overweighting of Choice-Consistent Evidence. *Current Biology*,

855   *28*(19), 3128-3135.e8. https://doi.org/10.1016/j.cub.2018.07.052

856   Tarantola, T., Folke, T., Boldt, A., Pérez, O.D., & De Martino, B. (2021). Confirmation bias

857   optimizes reward learning. *bioRxiv*. doi: https://doi.org/10.1101/2021.02.27.433214

858   Tavares, G., Perona, P., & Rangel, A. (2017). The attentional drift diffusion model of simple

859   perceptual decision-making. *Frontiers in neuroscience*, *11*, 468.

860   Thomas, A. W., Molter, F., Krajbich, I., Heekeren, H. R., & Mohr, P. N. (2019). Gaze bias

861   differences capture individual choice behaviour. *Nature Human Behaviour*, *3*(6), 625-

862   635.

863   Tolcott, M. A., Marvin, F. F., & Lehner, P. E. (1989). Expert decision-making in evolving

864   situations. *IEEE Transactions on Systems, Man, and Cybernetics*, *19*(3), 606–615.

865   https://doi.org/10.1109/21.31066

866   Tweney, R. D., Doherty, M. E., Worner, W. J., Pliske, D. B., Mynatt, C. R., Gross, K. A., &

867   Arkkelin, D. L. (1980). Strategies of Rule Discovery in an Inference Task. *Quarterly*

868   *Journal of Experimental Psychology*, *32*(1), 109–123.

869   https://doi.org/10.1080/00335558008248237

870   van Lieshout, L. L. F., Vandenbroucke, A. R. E., Müller, N. C. J., Cools, R., & de Lange, F. P.

871   (2018). Induction and Relief of Curiosity Elicit Parietal and Frontal Activity. *The*

872   *Journal of Neuroscience*, *38*(10), 2579–2588.

873   https://doi.org/10.1523/JNEUROSCI.2816-17.2018

874   Vandormael, H., Castañón, S. H., Balaguer, J., Li, V., & Summerfield, C. (2017). Robust

875   sampling of decision information during perceptual choice. *Proceedings of the*

876   *National Academy of Sciences*, *114*(10), 2771–2776.

877   Wang, M. Z., & Hayden, B. Y. (2019). Monkeys are Curious about Counterfactual Outcomes.

878   *BioRxiv*.

879   Wason, P. C. (1960). On the failure to eliminate hypotheses in a conceptual task. *Quarterly*

880   *Journal of Experimental Psychology*, *12*(3), 129–140.

881   https://doi.org/10.1080/17470216008416717

882   Wason, P. C. (1968). Reasoning about a Rule. *Quarterly Journal of Experimental Psychology*,

883   *20*, 273–281.

884   Wilson, R. C., Geana, A., White, J. M., Ludvig, E. A., & Cohen, J. D. (2014). Humans use

885   directed and random exploration to solve the explore–exploit dilemma. *Journal of*

886   *Experimental Psychology: General*, *143*(6), 2074–2081.

887   https://doi.org/10.1037/a0038199

888   Yang, S. C.-H., Wolpert, D. M., & Lengyel, M. (2016). Theoretical perspectives on active

889   sensing. *Current Opinion in Behavioral Sciences*, *11*, 100–108.

890   https://doi.org/10.1016/j.cobeha.2016.06.009

891   Yeung, N., & Summerfield, C. (2012). Metacognition in human decision-making: Confidence

892   and error monitoring. *Philosophical Transactions of the Royal Society B: Biological*

893   *Sciences*, *367*(1594), 1310–1321. https://doi.org/10.1098/rstb.2011.0416

894

Supplemental Material

S1: Hierarchical regression predicting sampling time difference

Hierarchical Regression Model Predicting Sampling Time Difference, Experiment 1

Predictor	Coefficient	SE	t-value	p-value
Intercept	-0.27	0.03	-8.48	<0.0001
Choice	0.51	0.05	9.64	<0.0001
Dot Difference	0.05	0.004	13.02	<0.0001
Confidence	-0.08	0.03	-2.88	0.007
Choice x Confidence	0.19	0.04	5.26	<0.0001

This model was run using data from experiment 1. Confidence was z-scored per participant. It shows that choice is a significant predictor of sampling time difference between the chosen and unchosen circles, such that participants were likely to view the circle they chose for a longer time in the sampling phase. Furthermore, the interaction between choice and confidence was a significant predictor, signifying that the more confident participants were in their choice, the more biased their sampling was in favour of that choice. Dot difference also significantly predicts sampling time: participants sampled the circle containing the most dots for longer. Note that this variable shares some variance with the choice variable.

Hierarchical Regression Model Predicting Sampling Time Difference, Experiment 2

Predictor	Coefficient	SE	t-value	p-value
Intercept	-0.15	0.06	-2.63	0.017
Choice	0.30	0.10	2.90	0.01
Dot Difference	0.04	0.005	9.21	<0.0001
Confidence	-0.09	0.04	-2.43	0.03
Choice x Confidence	0.21	0.05	4.29	0.0005

We also ran this model using data from experiment 2, excluding control condition trials in which sampling time difference was fixed. Confidence was z-scored per participant. Again, choice, the interaction term between choice and confidence and dot difference were all significant predictors of the difference in sampling time between the two circles.

S2: Hierarchical logistic regression predicting change of mind

Hierarchical Logistic Regression Model Predicting Change of Mind, Experiment 1

Predictor	Coefficient	SE	z-value	p-value
Intercept	-1.18	0.21	-5.65	<0.0001
Sampling Bias	-1.24	0.08	-11.0	<0.0001
Dot Difference	-0.24	0.11	-3.06	0.002
Confidence	-0.58	0.06	-10.12	<0.0001

This model was run using data from experiment 1. Using a hierarchical logistic regression we predicted the log odds ratio that the participant changed their mind in the second decision phase on a given trial. Sampling bias and confidence were z-scored per participant. Convergence issues were addressed by square-root-transforming dot difference. Sampling bias is defined as the difference between the amount of time the initially chosen vs unchosen circles were sampled. This variable negatively predicts change of mind, meaning the more the chosen circle was sampled, the less likely participants were to change their mind. Absolute dot difference, which is a measure of trial difficulty, negatively predicts change of mind, such that participants were less likely to change their mind when the trial was easier. Finally, confidence was a negative predictor of change of mind; the higher confidence in the initial choice, the less likely participants were to change their mind.

Hierarchical Logistic Regression Model Predicting Change of Mind, Experiment 2

Predictor	Coefficient	SE	z-value	p-value
Intercept	-1.01	0.28	-3.62	0.0002
Sampling Bias	-1.40	0.20	-7.07	<0.0001
Dot Difference	-0.26	0.09	-2.72	0.007
Confidence	-0.57	0.06	-8.86	<0.0001
Fixed Sampling	-0.10	0.14	-0.67	0.50
Sampling Bias x Fixed Sampling	1.44	0.20	7.24	<0.0001

A similar model to that presented was run on the data from experiment 2. Because this experiment included a control condition in which sampling was fixed, we included a dummy variable in this model coding whether the trial was in the control condition or not. Sampling bias, dot difference, and confidence were all significant negative predictors of change of mind. However, an interaction term between 'Fixed Sampling' (the control condition) and sampling bias was significantly positive, and approximately the same size as the main effect of sampling bias on change of mind. This means that in control condition trials, the effect of sampling bias on change of mind disappears.

958 ## S3: Sampling Phase Length in Experiment 2
959
960 In experiment 2, the length of the sampling phase (phase II) was varied to study the
961 emergence of a sampling bias over time. The sampling phase randomly varied and was
962 either 3000, 5000 or 7000ms. In the experiment, participants were forced to look at each
963 circle at least once during the sampling phase, otherwise the trial was invalidated. We
964 decided to exclude trials with the shortest sampling phase length of 3000ms, because it
965 became apparent that participants were not able to allocate this time beyond looking at
966 each patch once.
967
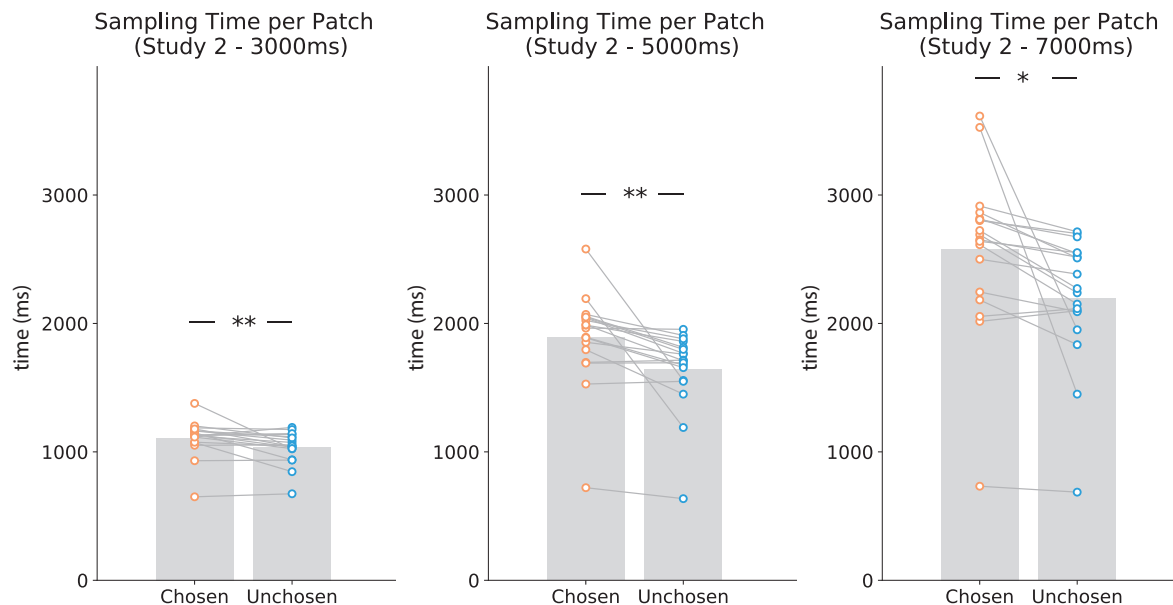
968
969
970 **Figure S3.1**: Mean sampling time spent viewing the initially chosen and unchosen patch in the sampling
971 phase in studies 2 for each sampling phase length. Data points represent individual participants.
972
973

32

974 S4: Change of mind as a function of stimulus presentation time in the fixed sampling
975 condition
976
977

Sampling bias only affects choice in free sampling trials
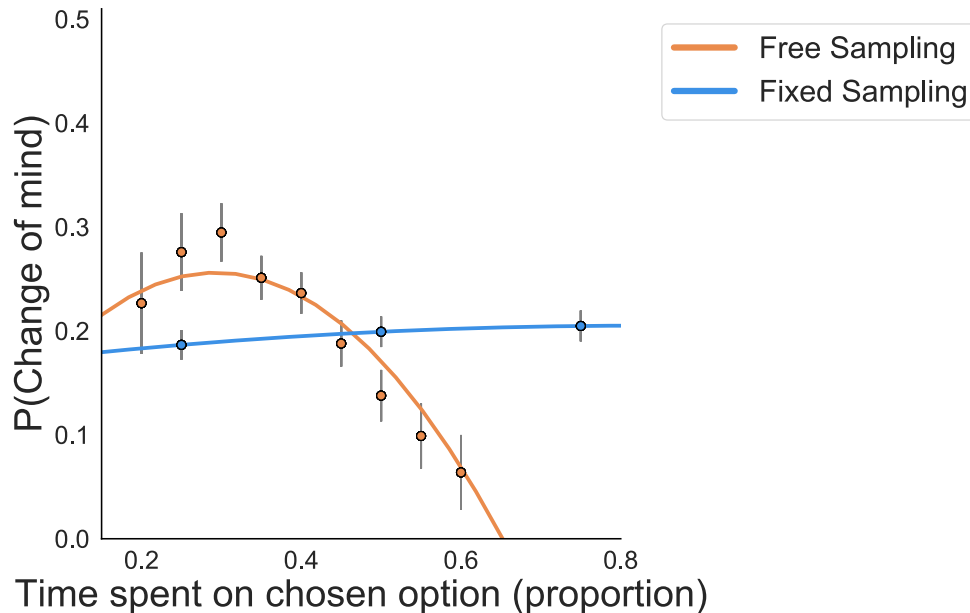
978
979
980 **Figure S4.1** In Figure 4D, we plotted the probability of changes of mind as a function of actual gaze time by
981 participants in the fixed viewing condition. Even though stimulus presentation was fixed, participants could
982 still choose to saccade back to the central fixation cross before the end of stimulus presentation. Therefore,
983 actual gaze time is a better reflection of information processing in this condition. Here, we are plotting
984 stimulus presentation time in the fixed sampling condition instead. The same lack of effect of sampling bias
985 on changes of mind can be seen here in the fixed sampling condition as in Figure 4.
986

S5. GLAM modelling
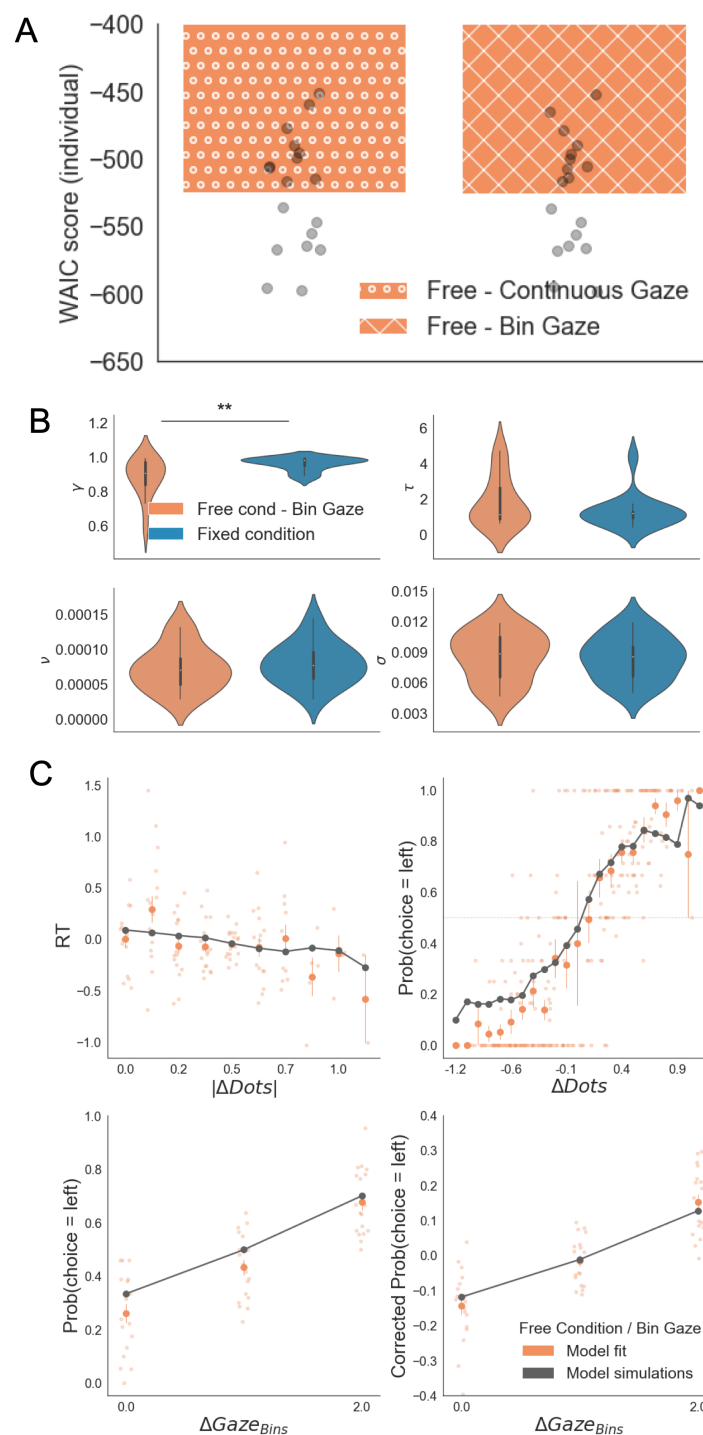
Figure S5.1. GLAM results with down-sampled gaze information in the free sampling condition. To check whether the results reported in the main text are an artefact of the low variability in relative gaze in fixed sampling trials, we reduced the variability of free sampling trials to only 3 bins. This replicated the number of bins in the gaze allocation presented in our fixed-viewing trials. GLAM was fitted using these data. (A) Model comparison between the GLAM models in the free sampling condition fitted using original gaze data (continuous gaze) and the down sampled gaze information (bin gaze). We found no significant difference between the model fit scores in both cases (Mean $WAIC_{Free/Continuous}$ = -524.58; Mean $WAIC_{Free/Bin}$= -525.12; t_{17} = 1.46, p = 0.16, ns). Model fitting was performed at a participant level. WAIC scores are presented using log scale. A higher log-score indicates a model with better predictive accuracy. (B) GLAM

999 parameters resulting from Bin Gaze model fit. We replicated the findings in the main text, with a higher
1000 gaze bias (lower γ parameter) in the free sampling condition (Mean γ $_{Free}$= 0.81, Mean γ $_{Fixed}$= 0.96, t$_{17}$= -
1001 3.268; p<0.01). No significant differences were found for the other parameters. γ: gaze bias; τ: evidence
1002 scaling; ν: drift term; σ: noise standard deviation. (C) The Bin Gaze model replicated the main behavioural
1003 relationships in a similar way to the original continuous gaze model. The four panels present four relevant
1004 behavioural relationships comparing model predictions and overall participant behaviour: (top left)
1005 responses time were faster (shorter RT) when the choice was easier (i.e. bigger differences in the number
1006 of dots between the patches); (top right) probability of choosing the left patch increased when the number
1007 of dots was higher in the patch on the left side (ΔDots = Dots$_{Left}$ – Dots$_{Right}$); (bottom left) the probability of
1008 choosing an alternative depended on the gaze difference (ΔGaze = g$_{Left}$ – g$_{Right}$); and (bottom right) the
1009 probability of choosing an item that was fixated longer than the other, corrected by the actual evidence
1010 (ΔDots), depicted a residual effect of gaze on choice. Solid dots depict the mean of the data across
1011 participants. Lighter dots present the mean value for each participant across bins. Solid grey lines show the
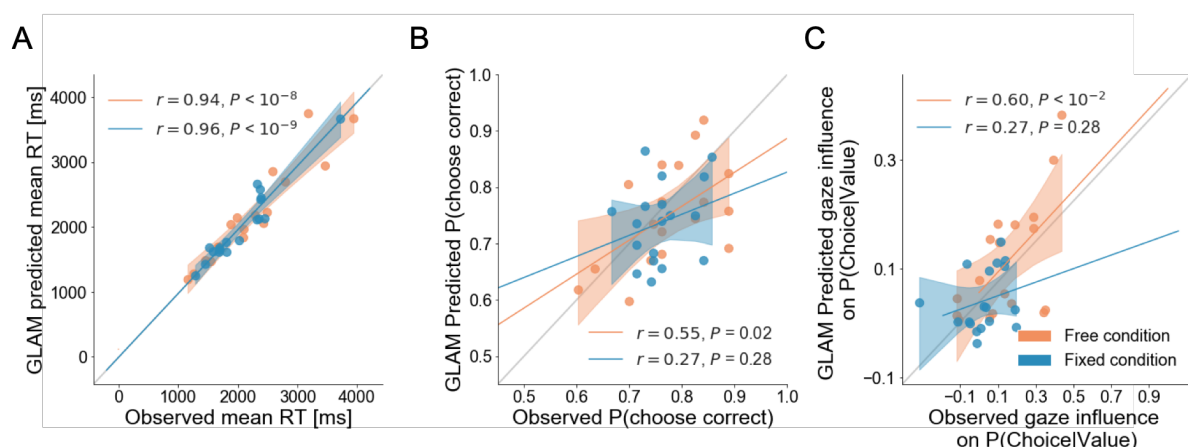1012 average for model simulations. Data are binned for visualization. **: p<0.01.

Figure S5.2. Individual out-of-sample GLAM predictions for behavioural measures in free and fixed sampling conditions. The correlations between observed data and predictions of the model for individual (A) mean RT, (B) the probability of choosing the correct patch, and (C) the gaze influence in choice probability are presented. In the fixed sampling condition, the correlation between the performance (probability of correct) of individual participants and the model predictions was found not statistically significant, indicating the model was not completely accurate in predicting participant-level performance. However, the model captures group-level performance (as depicted in Figure 5B), since predicted trials had higher than chance accuracy and a similar range of performance as observed trials (accuracy is between 0.6-0.9 for observed and predicted). Regarding the gaze influence measure (residual effect of gaze on choice, once the effect of evidence is accounted for), the free sampling model predicts this effect significantly at the participant level, but the fixed sampling model did not. Since in the fixed sampling model, in practical terms there is no gaze bias (γ ≈1), we expected the model would have trouble predicting any residual gaze influence. Dots depict the average of observed and predicted measures for each participant. In the free sampling condition the model prediction correlated significantly with observed accuracy and gaze influence, at the participant-level. Lines depict the slope of the correlation between observations and predictions. Dots indicate the average measure for each participant's observed and predicted data. Mean 95% confidence intervals are represented by the shadowed region. All model predictions are simulated using parameters estimated from individual fits for even-numbered trials.
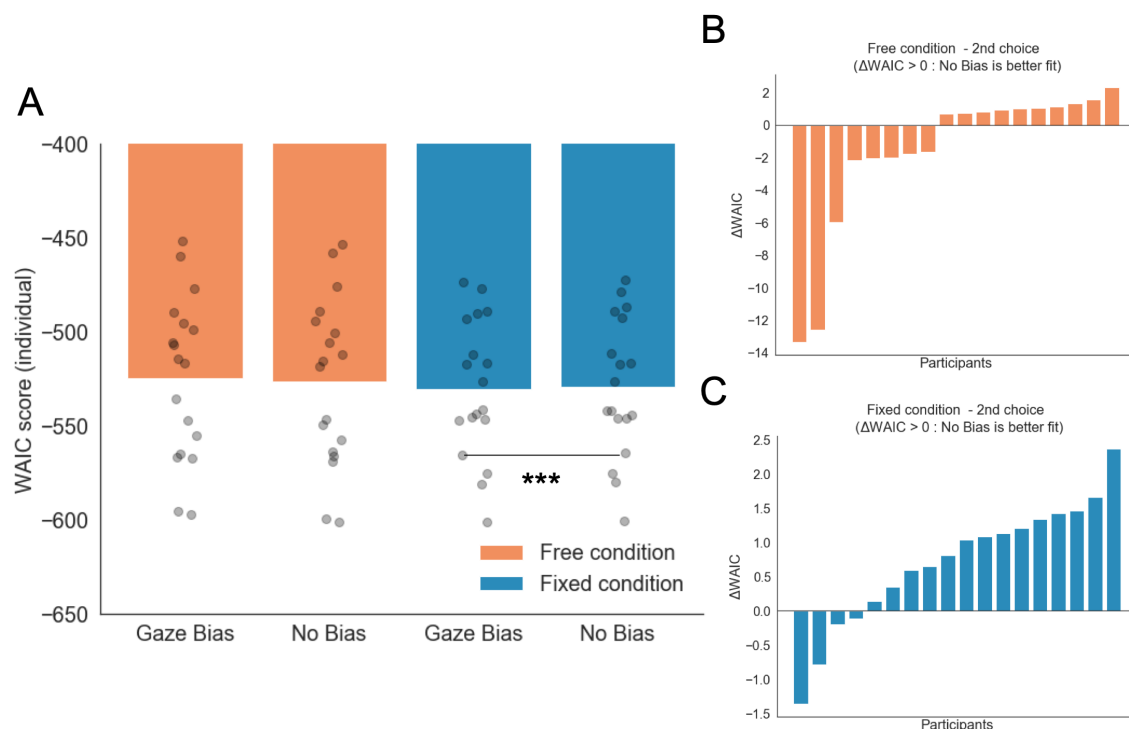
Figure S5.3. GLAM model comparison for free and fixed sampling conditions. (A) WAIC scores for free and fixed sampling models did not report significant difference in fit between the conditions (mean $WAIC_{Free}$ = -524.58; mean $WAIC_{Fixed}$ = -530.004; t_{17} = 0.98, p = 0.34, ns). As an additional check, we fitted new models for both conditions without the gaze bias (no bias, γ = 1). We found that in the fixed gaze condition, the no bias model was the most parsimonious model (Mean $WAIC_{Fixed/GazeBias}$ = -530.004; Mean $WAIC_{Fixed/NoBias}$ = -529.29; t_{17} = 3.304, p < 0.01). No differences were found between the gaze bias and no bias models in the free sampling condition (Mean $WAIC_{Free/GazeBias}$ = -524.58; Mean $WAIC_{Free/NoBias}$ = -526.23; t_{17} = 1.537, p = 0.14, ns). (B-C) Differences in WAIC score between Gaze Bias – No Bias (ΔWAIC) models were calculated for each individual participant and experimental condition. This corroborated that the No Bias model has a better fit in the fixed sampling condition only (C). WAIC scores are presented using log scale. A higher log-score indicates a model with better predictive accuracy. ***: p<0.001.
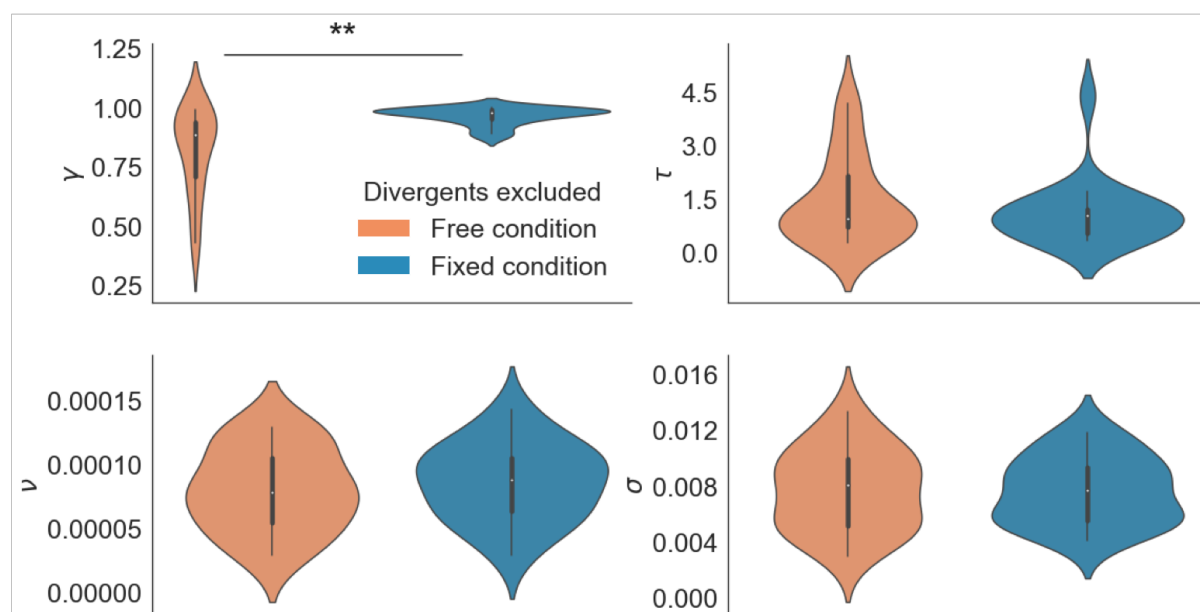
Figure S5.4. GLAM parameters for free and fixed sampling conditions. Participants for which parameter estimation did not converge were removed from the analysis (4 participants). The results reported in the main text were still observed: a higher gaze bias (lower γ parameter) in the free sampling condition (Mean γ_{Free}= 0.786, Mean γ_{Fixed}= 0.961, t_{17}= -3.033; p<0.01). No significant differences were found for the other parameters. γ: gaze bias; τ: evidence scaling; ν: drift term; σ: noise standard deviation. **: p<0.01

1062 ## S6: A Descriptive Model of Confirmatory Information Processing

1063 We have designed the following descriptive economic decision-making model that can be
1064 used to capture the findings described above. There is a set of states of the world that
1065 denote the number of dots in each part of the screen, denoted $(\omega^L, \omega^R) \in \mathbb{R}^2_+$. In what
1066 follows, for any probability measure μ over $\mathbb{R}^2_+$, denote by $\mu(\omega^L > \omega^R)$ the probability
1067 that μ assigns to the event that ω^L is above ω^R; and by $BU(\mu, A)$ the Bayesian update of μ
1068 using information A. Subjects start with a symmetric prior μ_0.
1069
1070 **First Stage**: Subject go through a sampling phase in which they gather information about
1071 the number of dots in each screen. They get two noisy signals about each component of
1072 the state, x_1^L and x_1^R, for which for simplicity we assume normally distributed noise:
1073
1074 $$x_1^L = \omega^L + \epsilon_1^L$$
1075 $$x_1^R = \omega^R + \epsilon_1^R$$
1076
1077 where $\epsilon_1^L \sim N(0, \sigma_1^L)$, $\epsilon_1^R \sim N(0, \sigma_1^R)$. For generality, we may allow the variances to be
1078 different, in case individuals follow an adaptive search procedure that leads to
1079 asymmetric information acquisition. For simplicity, we assume here that they are the
1080 same, $\sigma_1^L = \sigma_1^R = \sigma^2$.
1081
1082 At the end of the information gathering stage, subjects form a posterior about the state of
1083 the world, $\widehat{\mu_1} = BU(\mu_0 | x_1^R, x_1^L)$. They are then asked to choose an action $a_1 \in \{L, R\}$ and a
1084 confidence level c. Following standard arguments, they choose $a_1 = L$ if $\widehat{\mu_1}(\omega^L > \omega^R) >$
1085 0.5, $a_1 = R$ if $\widehat{\mu_1}(\omega^L > \omega^R) < 0.5$, and randomize with equal probability between L and
1086 R if $\widehat{\mu_1}(\omega^L > \omega^R) = 0.5$ (a probability 0 event).
1087
1088 They report a confidence
1089
$$c = 2 * \left| \widehat{\mu_1}(\omega^L > \omega^R) - \frac{1}{2} \right|.$$
1090
1091 Note that $c = 1$ if $\widehat{\mu_1}(\omega^L > \omega^R) \in \{1, 0\}$, $c = 0$ if $\widehat{\mu_1}(\omega^L > \omega^R) = 0.5$.
1092
1093
1094 **Second Stage**: The key departure from the standard normative model of decision-making
1095 is that the beliefs that individuals have at the beginning of stage 2 are not the posteriors
1096 they obtained at the end of stage 1, but are also influenced by their choice. Denote $\widehat{\mu_1}^D$
1097 the belief over the state of the world that the individual would obtain if she only observed
1098 her choice – given the decision rule above – but did not take into account the signals x_1^L
1099 and x_1^R, i.e., $\widehat{\mu_1}^D = BU(\mu_0, a_1)$. We posit that the belief used by the individual at the
1100 beginning of the second stage, μ_2, is a convex combination of this belief and the posterior
1101 at the end of period 1, i.e.,
1102
1103 $$\mu_2 = \theta \widehat{\mu_1}^D + (1 - \theta) \widehat{\mu_1}$$
1104
1105 where $\theta \in [0, 1]$ indicates the degree of distortion: when $\theta = 0$ beliefs are not distorted;
1106 when $\theta = 1$, choices in the second stage only use the more limited information contained
1107 in the choice of stage 1, and not the full information coming from observed signals. The
1108 use of the linear structure is not relevant but simplifies the analysis.

1109

1110 In the second stage, subjects have to decide the sampling strategy. We posit that the
1111 fraction of time spent on Left is determined (with noise) as a strictly increasing function
1112 of the belief μ_2: subjects spend more time on average on the area where they believe are
1113 more dots. In Sepulveda et al. (2020) a model is described that generates a related
1114 tendency. In particular, we posit that the fraction of time spent on the left circle, s, is
1115 determined by a standard logistic function

1116

1117
$$s = \frac{1}{1 + e^{-k(\mu_2(\omega^L > \omega^R) - 0.5)}}.$$

1118

1119 Through this sampling, individuals receive signals

1120

1121
$$x_2^L = \omega^L + \epsilon_2^L$$
1122
$$x_2^R = \omega^R + \epsilon_2^R$$

1123

1124 where $\epsilon_1^L \sim N(0, \frac{\sigma^2}{s})$, $\epsilon_1^R \sim N(0, \frac{\sigma^2}{1-s})$, so that better information is obtained about each
1125 dimension the more time spent contemplating it. Using this, subjects form a posterior
1126 $\widehat{\mu_2} = BU(\mu_2, |x_2^L, x_2^R)$.

1127

1128 Finally, subjects choose an action $a_2 \in \{L, R\}$ and a confidence level c. Following standard
1129 arguments, they choose $a_1 = L$ if $\widehat{\mu_2}(\omega^L > \omega^R) > 0.5$, $a_1 = R$ if $\widehat{\mu_2}(\omega^L > \omega^R) < 0.5$, and
1130 randomize with equal probability between L and R if $\widehat{\mu_2}(\omega^L > \omega^R) = 0.5$ (again, a 0
1131 probability event).

1132

1133 This model is reminiscent of the Second-Order Model of Fleming and Daw (2017), but
1134 with the difference that we are not introducing a choice-affecting-beliefs component at
1135 the time of reporting confidence, but only at the beginning of the second stage.

1136

1137 This model implies that, on average, subjects will sample more from the patch they had
1138 previously chosen; and that this effect is stronger the higher their confidence in their
1139 choice—which is what we see in the data.

1140

1141 When $\theta > 0$, the previous choice affects the sampling strategy even controlling for
1142 confidence—which is also what we find. Whenever $\widehat{\mu_1}^D(\omega^L > \omega^R) > \widehat{\mu_1}(\omega^L > \omega^R) >$
1143 0.5, the presence of $\theta > 0$ will lead to $\mu_2(\omega^L > \omega^R) > \widehat{\mu_1}(\omega^L > \omega^R) > 0.5$. This, in turn,
1144 will distort her later sampling strategy as well as her future beliefs: the individual will
1145 require even stronger information to change her mind. The opposite happens when
1146 $\widehat{\mu_1}(\omega^L > \omega^R) > \widehat{\mu_1}^D(\omega^L > \omega^R) > 0.5$: in this case $\widehat{\mu_1}(\omega^L > \omega^R) > \mu_2(\omega^L > \omega^R) > 0.5$,
1147 which implies that the individual acts as if she were less confident of her choice. Previous
1148 choice will affect the sampling strategy, but now less so.

1149

1150

1151

1152   S7: Model Comparisons

1153

1154   *Sampling Time Models*

1155   To investigate the effects of choice, dot difference, confidence, and response time on
1156   sampling we compared 6 hierarchical regression models. These models are presented
1157   below. The BIC scores for each model in each experiment are plotted in Figure S7.1. The
1158   best fitting models according to BIC-scores were Model 5 in experiment 1 and Model 3 in
1159   experiment 2. We chose to present Model 5 for both experiments as we were interested
1160   in the contribution of confidence to biased sampling.

1161

1162

1163   **Sampling Time Models**

1164

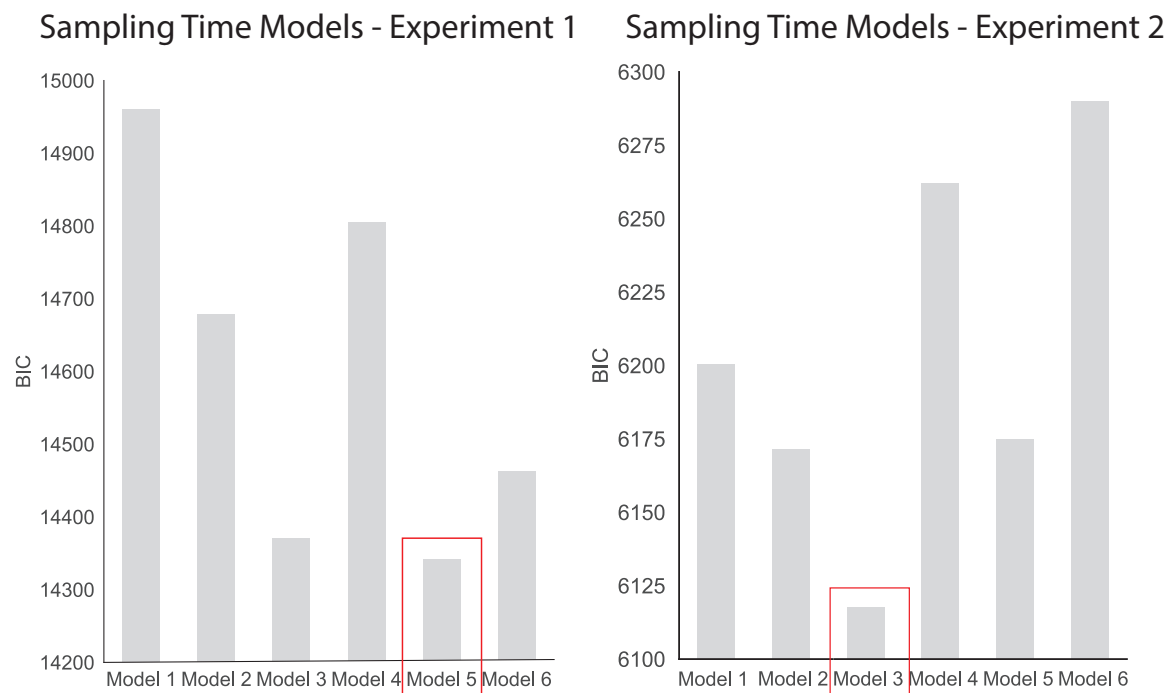| Models	Formula
1	Sampling Time Difference $\sim \mathcal{N}(\beta_0 + \beta_1[\text{Choice}] + \varepsilon)$
2	Sampling Time Difference $\sim \mathcal{N}(\beta_0 + \beta_1[\text{Dot Difference}] + \varepsilon)$
3	Sampling Time Difference $\sim \mathcal{N}(\beta_0 + \beta_1[\text{Choice}] + \beta_2[\text{Dot Difference}] + \varepsilon)$
4	Sampling Time Difference $\sim \mathcal{N}(\beta_0 + \beta_1[\text{Choice}] + \beta_2[\text{Confidence}] + \beta_3[\text{Choice} * \text{Confidence}] + \varepsilon)$
5	Sampling Time Difference $\sim \mathcal{N}(\beta_0 + \beta_1[\text{Choice}] + \beta_2[\text{Dot Difference}] + \beta_3[\text{Confidence}] + \beta_4[\text{Choice} * \text{Confidence}] + \varepsilon)$
6	Sampling Time Difference $\sim \mathcal{N}(\beta_0 + \beta_1[\text{Choice}] + \beta_2[\text{Dot Difference}] + \beta_3[\text{Confidence}] + \beta_4[\text{Reaction Time}] + \beta_5[\text{Choice} * \text{Confidence}] + \beta_6[\text{Choice} * \text{Reaction Time}] + \varepsilon)$

1165

1166

1167
1168 **Figure S7.1. BIC comparison of the sampling time models for experiments 1 and 2.**
1169 Model 5 fit the data from experiment 1 the best (BIC = 14340.8), whereas Model 3 was the best fit for the
1170 data in experiment 2 (BIC = 6117.8).
1171
1172

1173 *Change of Mind Models*
1174 To investigate the effects of dot difference, sampling, confidence, and response time on
1175 change of mind we compared 5 hierarchical regression models. These models are
1176 presented below. The BIC scores for each model in each experiment are plotted in Figure
1177 S7.2. Model 5 includes a dummy variable coding whether or not the trial was in the
1178 control condition or not (in which sampling was fixed). As such, this model is only
1179 applicable to experiment 2. The best fitting models according to BIC-scores were Model
1180 3 in experiment 1 and Model 5 in experiment 2.
1181
1182

Change of Mind Models

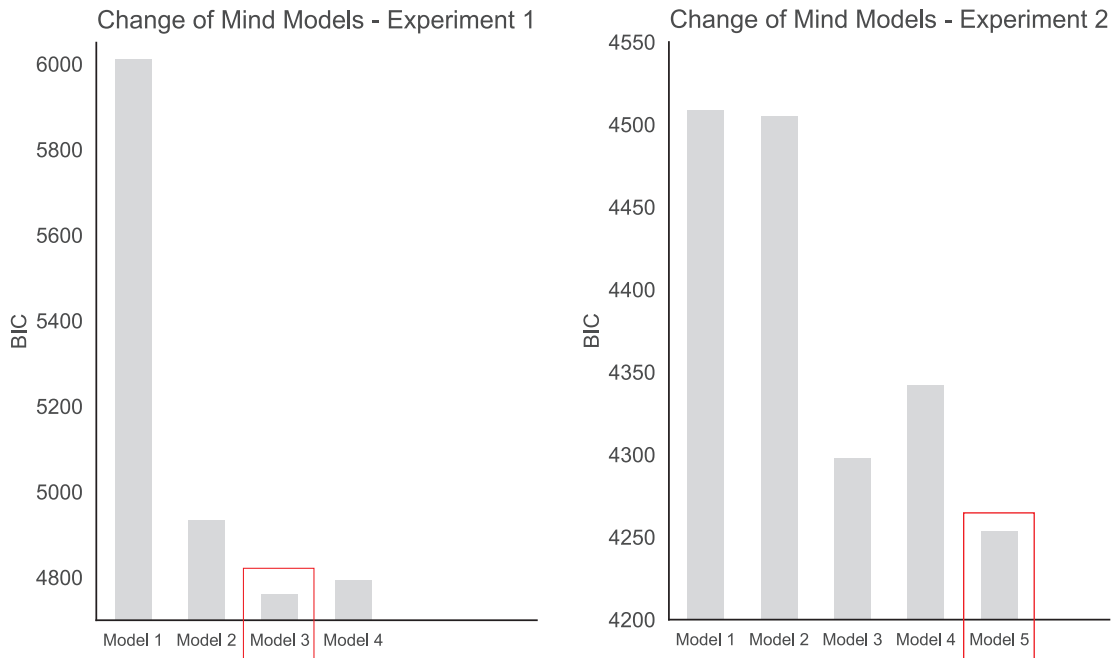| Models	Formula
1	Change of Mind $\sim$ logit^{-1}(β_0 + β_1[Dot Difference] + ϵ)
2	Change of Mind $\sim$ logit^{-1}(β_0 + β_1[Dot Difference] + β_2[Sampling Bias] + ϵ)
3	Change of Mind $\sim$ logit^{-1}(β_0 + β_1[Dot Difference] + β_2[Sampling Bias] + β_3[Confidence] + ϵ)
4	Change of Mind $\sim$ logit^{-1}(β_0 + β_1[Dot Difference] + β_2[Sampling Bias] + β_3[Confidence] + β_4[Reaction Time] + ϵ)
5	Change of Mind $\sim$ logit^{-1}(β_0 + β_1[Dot Difference] + β_2[Sampling Bias] + β_3[Confidence] + β_4[Fixed Sampling] + β_5[Sampling Bias * Fixed Sampling] + ϵ)

Figure S7.2. BIC comparison of the change of mind models for experiments 1 and 2.
Model 3 fit the data from experiment 1 the best (BIC = 4760.8), whereas Model 5 was the best fit for the data in experiment 2 (BIC = 4253.7).