Gaze Bias Differences Capture Individual Choice Behavior

Armin W. Thomas[†,1,2,3], Felix Molter[†,2,3,4], Ian Krajbich[5], Hauke R. Heekeren[2,3], & Peter N. C. Mohr[*,2,3,4]

1 Technische Universität Berlin

2 Freie Universität Berlin

3 Center for Cognitive Neuroscience Berlin

4 WZB Berlin Social Science Center

5 Ohio State University

* Corresponding author

† Shared first authorship with equal contribution.

Author Note

Armin W. Thomas, Electrical Engineering and Computer Science, Technische Universität Berlin; Felix Molter, School of Business & Economics, Freie Universität Berlin; Ian Krajbich, Department of Psychology, Department of Economics, Ohio State University; Hauke R. Heekeren, Department of Education & Psychology, Freie Universität Berlin; Peter N. C. Mohr, School of Business & Economics, Freie Universität Berlin.

Ordering of shared first authorship was determined by the following coin toss procedure: A.W.T. brought three Canadian Dollar coins from which F.M. selected one. F.M. chose one of the coin's sides. F.M. flipped the coin and let it fall onto a hard wooden floor. The upper side of the coin then determined the ordering outcome.

Correspondence concerning this article should be addressed to Peter N. C. Mohr, School of Business & Economics, Freie Universität Berlin, Garystr. 21, 14195 Berlin, Germany. Email: peter.mohr@neuroeconomics-lab.de.

Abstract

How do we make simple consumer choices (e.g., deciding between an apple, an orange, and a banana)? Recent empirical evidence suggests a close link between choice behavior and eye movements at the group level, with generally higher choice probabilities for items that were looked at longer during the decision process. However, it is unclear how variable this effect is across individuals. Here, we investigate this question in a multialternative forced-choice experiment using a novel computational model that can be easily applied to the individual participant level. We show that a link between gaze and choice is present for most individuals, but differs considerably in strength, namely, the choices of some individuals are almost independent of gaze allocation, while the choices of others are strongly associated with gaze behavior. Accounting for this variability in our model allows us to explain and accurately predict individual differences in observed choice behavior.

Keywords: decision making, individual differences, eye movements, gaze bias, evidence accumulation

Main text word count: 6506

Gaze Bias Differences Capture Individual Choice Behavior

In everyday life, we are constantly confronted with simple consumer choices, such as whether to have an apple or a banana for breakfast, or which bottle of juice to buy at the supermarket. Traditional models describing this type of consumer choice assume that people assign a utility (or subjective value) to each available option and make utility-maximizing choices (Von Neumann & Morgenstern, 1945). Notably, choices are assumed to be based solely on option attributes, and thereby, are independent of information search processes during the decision. This assumption has recently been challenged by a variety of empirical findings showing that the allocation of gaze during the decision-making process also plays a significant role, as a longer gaze towards one option is regularly associated with a higher choice probability for that option (Armel, Beaumel, & Rangel, 2008; Cavanagh, Wiecki, Kochar, & Frank, 2014; Fiedler & Glöckner, 2012; Folke, Jacobsen, Fleming, & De Martino, 2016; Glöckner & Herbold, 2011; Konovalov & Krajbich, 2016; Krajbich & Rangel, 2011; Krajbich, Armel, & Rangel, 2010; Krajbich, Lu, Camerer, & Rangel, 2012; Pärnamets et al., 2015; Shimojo, Simion, Shimojo, & Scheier, 2003; Stewart, Gächter, Noguchi, & Mullett, 2015; Stewart, Hermens, & Matthews, 2016; Vaidya & Fellows, 2015). Further, external manipulation of an individual's gaze allocation changes choice probabilities accordingly (Armel et al., 2008; Pärnamets et al., 2015; Shimojo et al., 2003; Tavares, Perona, & Rangel, 2017).

These findings led to the development of novel computational models, which integrate eye movement data into the choice process and formalize the empirically observed association between gaze and choice (Ashby, Jekel, Dickert, & Glöckner, 2016; Cavanagh et al., 2014; Fisher, 2017; Krajbich & Rangel, 2011; Krajbich et al., 2010, 2012; Towal, Mormann, & Koch, 2013). These models are based on classic evidence accumulation models (Ratcliff, 1978; Ratcliff, Smith, Brown, & McKoon, 2016) and make the additional assumption that the

momentary rate of evidence accumulation depends on the decision maker's eye movements: Evidence accumulation for an option is assumed to be discounted by a constant factor while another item is fixated upon. Accounting for this gaze bias, these models provide a precise quantitative account of many aspects of simple consumer choice behavior at the group level.

While group level statistics are informative for some research questions (e.g., forecasting product sales in economic research), statements about the majority of people, or the "average person", are often unsuitable for understanding the choice behavior of an individual. Even worse, aggregate models of behavior can lead to false conclusions about true underlying individual processes (Grandy, Lindenberger, & Werkle-Bergner, 2017; Lewandowsky & Farrell, 2010): In a learning task, for example, the group level average learning curve would appear as a gradual, smooth function over time, even if all individuals showed abrupt, step-like learning curves (much like an epiphany), but with variable learning onsets across individuals (Hayes, 1953). This group level model, however, would not describe any individual of the group well, and the deduction that individual learning occurs smoothly would be false. It is thus crucial to understand and explain choice behavior at the individual level.

Similarly, previously reported group level models quantifying the association between gaze and choice specified a constant gaze bias for all individuals (e.g., Krajbich et al., 2010; Krajbich & Rangel, 2011), without testing the model performance on the individual level. It therefore remains to be shown whether an association between gaze and choice is present across individuals and whether the strength of such an association is constant. If, however, people's decisions were affected differently by looking behavior, we would find that the choices of some individuals are more biased by gaze, and therefore, more inconsistent with subjective value ratings than other individuals' choices. Imagine, for example, a choice between two bottles of orange juice in the supermarket: one has a slightly higher utility for the decision maker than the other, but it is also less visually salient. If this person's association of

gaze and choice behavior was strong, her choice would be biased towards the bottle that is attracting more of her gaze, even though it has lower utility. On the other hand, if the person's association is weak, she would then be able to select the option that is higher in utility, despite her gaze being attracted more towards the inferior option. Accordingly, if the strength of this association is variable across individuals, it is necessary to account for these differences to accurately predict individual choice behavior.

Here, we investigated whether the previously reported link between gaze and choice behavior is variable across individuals, using a novel computational model that can easily be applied to individual participant and multialternative choice data. With this model, we reaffirmed that an association between gaze and choice is present at the group level, and indeed, present for most individuals. The strength of this association, however, showed substantial variability. By accounting for this variability, we were able to explain and accurately predict empirically observed differences in individuals' choice behavior.

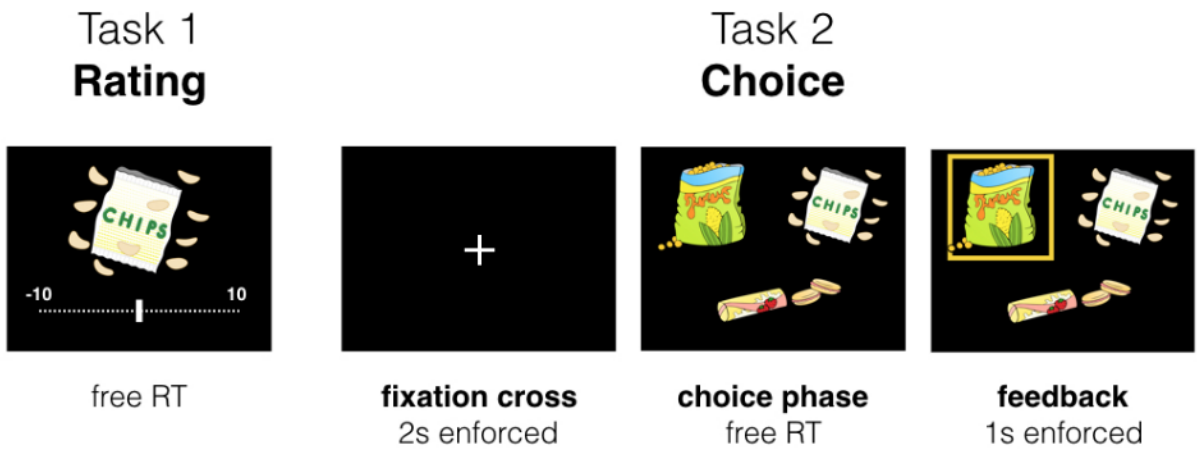
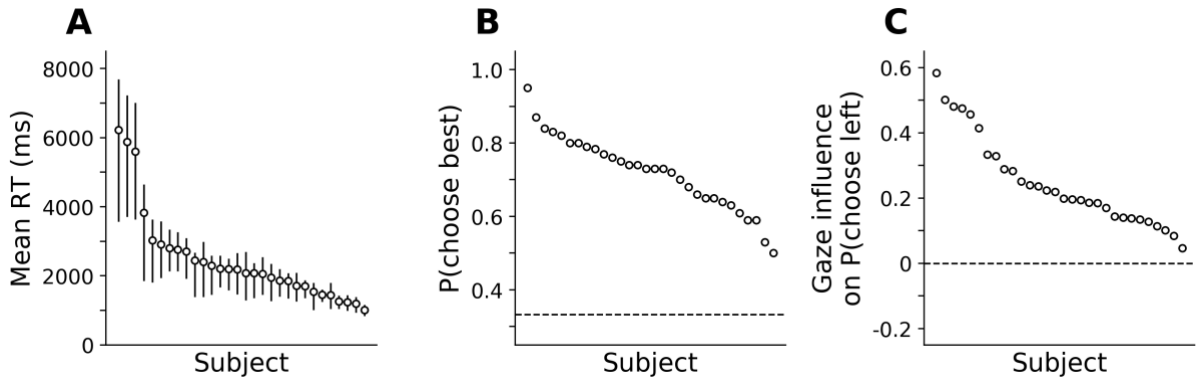


Figure 1: Experimental Paradigm. All the participants completed two tasks in a single session. **Task 1:** The participants rated all 70 snack foods items on a liking rating scale between -10 and 10, according to how much they would like to eat each item. **Task 2:** In each choice trial, the participants were required to maintain a central fixation for 2 s. Next, the participants were asked to choose the item that they would like to eat most from sets of 3 choice items, while their eye movements were being recorded. Choices were followed by 1 s of visual feedback.

Results

Data set & Task overview

To investigate individual differences in the influence of gaze allocation on simple economic choice behavior, we used a previously published, prototypical data set (Krajbich & Rangel, 2011) that we obtained from the original authors in a preprocessed format (see Methods for full details). In the corresponding experiment, hungry participants made choices between three snack food items, without time restrictions (Figure 1). Participants also gave a liking rating for each of the 70 snack food items that were used in the experiment. During the choice task, the participants' eye movements were continuously recorded using an eye tracker. The data set included 30 participants, who performed 100 choice trials each.



137

138 *Figure 2*: Individual differences in the three behavioral metrics: response time,
139 probability of choosing the highest rated item and influence of gaze on choice probability. **A**:
140 Mean individual response times (error bars denote first and third quartile). **B**: Mean
141 probability of choosing the item with the highest liking rating. The dashed horizontal line
142 indicates chance level accuracy. **C**: Individual influence of gaze on choice probability (mean
143 increase in choice probability for an item that is fixated longer than the others, after correcting
144 for the influence of item value). The data points are sorted from high to low in each panel.

145

146 **Individual differences in the data**

147 We analyzed three metrics for individual differences, namely, (i) the participants'
148 response time, (ii) the mean probability of choosing the item with the highest liking rating,
149 and (iii) the influence of gaze allocation on choice probability (mean increase in choice
150 probability for an item that was fixated on longer than the others, after correcting for the
151 influence of the item value). We found that participants differed considerably in all of the
152 three metrics (Figure 2):

153 The participants' mean response times ranged from 1006 to 6217 ms, with mean ± s.d.
154 = 2462 ± 1298 ms (Figure 2A), while their probabilities of choosing the highest rated item in
155 a trial ranged from 50.00% to 95.00%, with mean ± s.d. = 71.94% ± 10.01% (Figure 2B).

We also probed the relationship between individual allocation of gaze and choice. Previous work in simple choice tasks has shown that individuals are, on average, more likely to choose an option when they spent more time fixating on it, relative to the others (Armel et al., 2008; Cavanagh et al., 2014; Folke et al., 2016; Krajbich & Rangel, 2011; Krajbich et al., 2010). Here, we devised an individual measure to quantify the relationship between gaze allocation and choice for each individual: following previous work (Krajbich & Rangel, 2011; Krajbich et al., 2010), for each participant we first estimated the probability of choosing the left item in a choice set using logistic regression, based on its relative item value (the difference between the item's value and the mean value of all other items in that trial) and the range between the other items' value. We then subtracted this estimated probability from the empirically observed choice (either 1 if the left item was chosen, or 0 otherwise). Finally, we averaged the resulting "residual" choice probability for trials in which the left item had a positive and negative final gaze advantage (computed as the difference in the fraction of the total fixation time that the participants spent fixating on the left item and the average fraction that they spent fixating on the others). The difference between these two described the average difference in choice probability for the items with a positive versus negative final gaze advantage, when corrected for the influence of relative item value on choice probability and the other items' range of values. We found that individual scores on this measure ranged from 0.05 to 0.58, with mean ± s.d. = 0.25 ± 0.14 (Figure 2C). Notably, all the participants showed positive scores, indicating an overall positive relationship between gaze allocation and choice. We did, however, find strong variation in this measure.

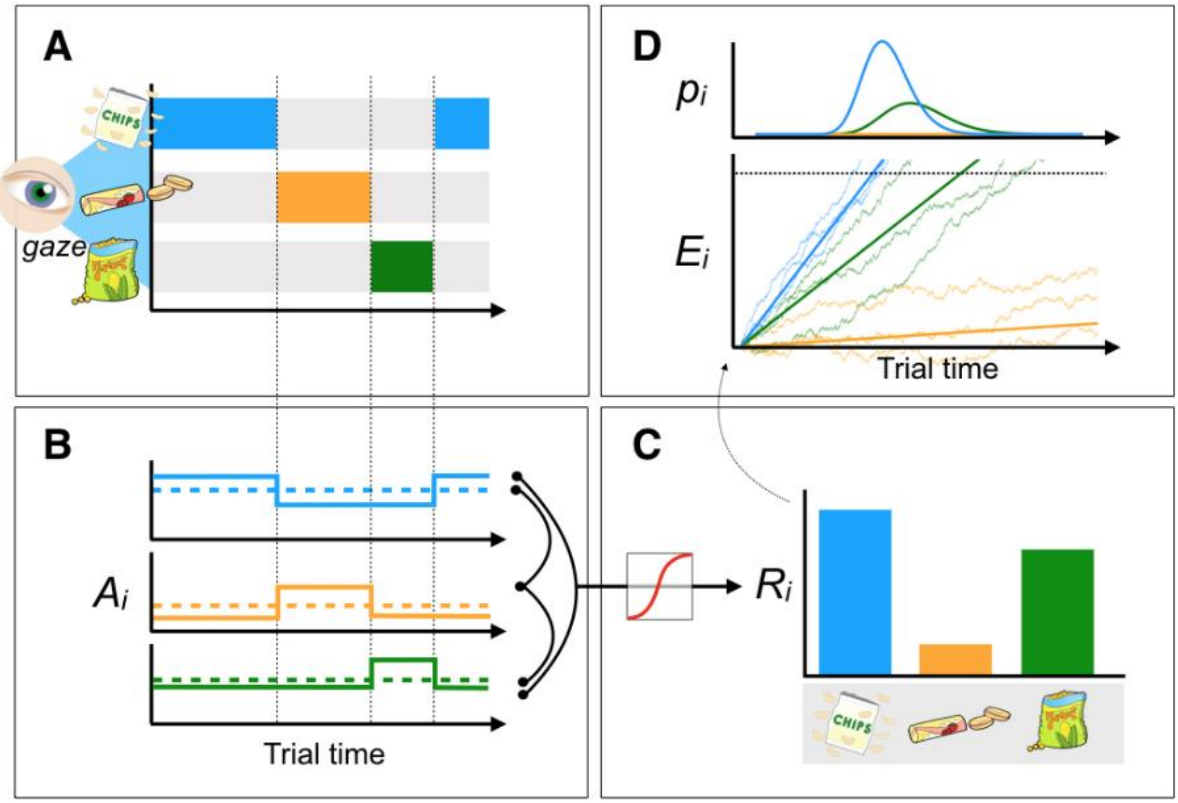


Figure 3: Gaze-weighted linear accumulator model (GLAM). The GLAM describes the influence of gaze allocation on the decision process, in the form of a linear stochastic race: It assumes that individuals accumulate evidence in favor of each item *i* and make a choice as soon as the relative evidence E_i in favor of one item reaches a choice threshold (**D**). Importantly, the speed of the accumulation process is dependent on the distribution of visual gaze during the decision (**A** & **B**). For each option in the choice set, an absolute evidence signal A_i is computed. The magnitude of this signal is dependent on the allocation of visual gaze, with lower magnitudes for options that are momentarily not fixated on. Absolute evidence signals are transformed into relative decision signals R_i (indicating relative item preferences) by (i) computing the average absolute evidence signal for each item in the trial (dashed lines in **B**), (ii) and then computing the difference between each of these averages and the maximum of the respective other two. (iii) The GLAM assumes an adaptive representation of these relative evidence signals that is maximally sensitive to small differences in the relative decision signals. To this end, a sigmoid transform is applied (**C**). The resulting scaled

relative evidence signals determine the drift terms R_i in the stochastic race (**D**). The stochastic race provides first-passage time distributions p_i, describing the likelihood of each item being chosen at each time point. See Methods for a more detailed model description.

Modeling individual differences in simple economic choice

The behavioral and eye tracking data suggest substantial variability in the extent to which gaze affects a participant's choice behavior (Figure 2C). The computational mechanism relating gaze and choice patterns, however, cannot be inferred from descriptive analyses alone. In addition, conclusive quantitative evidence for or against the presence of a mechanism that biases choices depending on the distribution of gaze has yet to be provided at the individual level. We therefore adopted a principled computational modeling approach to investigate whether a formalized gaze bias mechanism, in conjunction with individual gaze patterns, can improve model predictions of individual choice and response time data.

We propose a new model called the Gaze-weighted Linear Accumulator Model (GLAM; Figure 3) that is inspired by the attentional Drift Diffusion Model (aDDM) proposed by Krajbich et al. (2010; also, Krajbich & Rangel, 2011; Krajbich et al., 2012). The GLAM assumes accumulation of evidence in favor of each item, that is modulated by gaze behavior: While an item is not fixated on, accumulation occurs at a rate discounted by the gaze bias parameter γ. A choice is made as soon as evidence in favor of one item reaches a decision threshold. The GLAM is rooted in the class of linear stochastic race models (Tillman, 2017; Usher, Olami, & McClelland, 2002). These models naturally generalize to choice scenarios with more than two items and remain analytically tractable, allowing for more complex applications (e.g., embedding in a hierarchical Bayesian framework).

In addition to the gaze bias parameter γ, the GLAM includes a general velocity parameter *v*, a noise parameter σ and a scaling parameter τ (see Methods for full model implementation details).

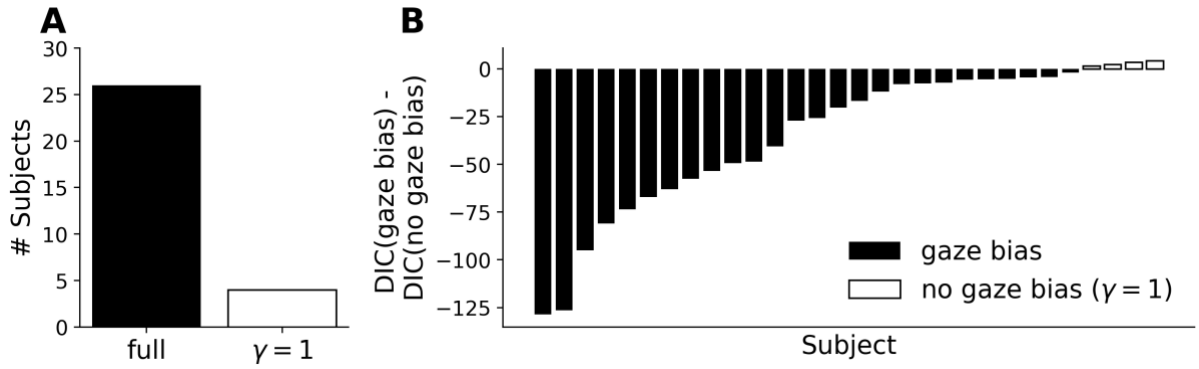


Figure 4: Model comparison between the full GLAM model and a no-gaze-bias GLAM (γ = 1) variant. **A:** Individual best fitting models, given by the lowest DIC score. Twenty-six of 30 participants (87%) were better described by the full model that includes a gaze bias mechanism. **B:** Individual differences in the Deviance Information Criteria (DIC) between the full and (γ = 1) model. Negative differences in the model DIC scores indicate better fits of the full model.

Individual model comparison

We fitted and compared two GLAM variants to the response time and choice data of each participant to gauge the evidence in favor of the previously described gaze bias mechanism and to quantify its strength on an individual level:

1. A *full* GLAM variant with free parameters *v*, γ, σ, τ. This model allowed the gaze bias parameter γ to vary freely between the individuals.
2. A *no-gaze-bias* GLAM variant, where the gaze bias parameter γ was fixed to 1 (resulting in no influence of gaze on the accumulation process)

The two models differ in their complexity: The full model has one more free parameter and can therefore be expected to provide a better absolute fit to the data. We used the Deviance Information Criterion (DIC; Spiegelhalter, Best, Carlin, & Van Der Linde, 2002; Spiegelhalter, Best, Carlin & Van Der Linde, 2014) to perform model comparisons at the individual level as it includes a penalty for model complexity: more complex models are only preferred only if their added complexity is justified by an improvement in absolute fit. Lower DIC scores indicate a better model fit accounting for differences in model complexity.

The full model fitted 26 of 30 (87%) participants better than the no-gaze-bias model (Figure 4A, B). The mean ± s.e.m. difference in the DIC scores between the full and no-gaze-bias models was -34.22 ± 6.84 (Figure 4B).

Individual estimates of the gaze bias parameter γ in the full model ranged from -0.93 to 0.81, with a mean ± s.d. = 0.20 ± 0.39 (Figure S2). Notably, the individual estimates covered a wide range of possible values between $\gamma = -1$ (strong gaze bias) to $\gamma = 1$ (no gaze bias). With a strong gaze bias, the GLAM leaks evidence for an item, while another is fixated on, whereas evidence accumulation is independent of gaze allocation when no gaze bias is present.

Taken together, the individual model comparison revealed that most participants' behavior was better described by a model that includes the gaze bias mechanism. Importantly, the extent to which the accumulation process was influenced by gaze, as captured by individual gaze bias (γ) estimates, showed non-trivial individual differences.

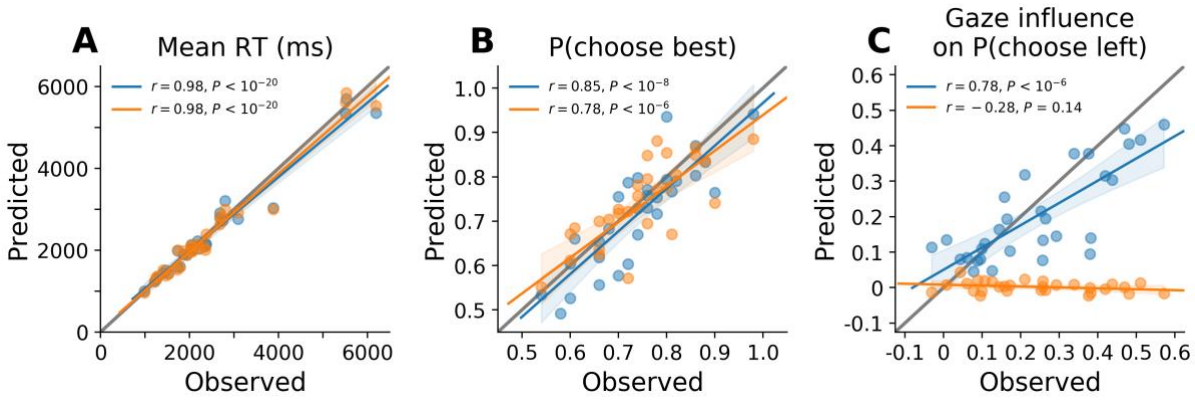


Figure 5: Out of sample correlations between the observed and predicted individual behavior in the odd experiment trials (the predictions were based on parameters estimated from even experiment trials). **A:** mean response time. **B:** probability of choosing the highest rated item. **C:** influence of gaze on choice probability. Model predictions are simulated from hierarchically estimated parameter estimates. Blue color indicates predictions from the full GLAM, whereas orange indicates predictions from a restricted GLAM variant with no gaze bias ($\gamma = 1$).

GLAM predicts individual choice behavior

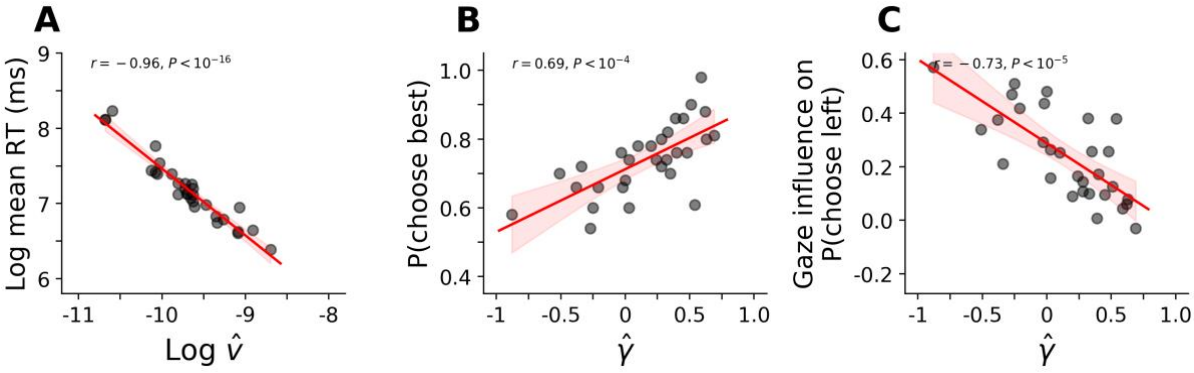
We found that in a relative model comparison the full GLAM best describes the data of most participants, when compared to a restricted variant with no gaze bias ($\gamma = 1$; see Figure 4). However, this analysis did not take into account whether the GLAM also accurately predicts individuals' behavior on an absolute level. To test this, we again used both model variants to simulate response data for each individual participant. This time, however, we split the data into even- and odd-numbered trials. We then used all the even trials to estimate the model parameters (training). Subsequently, we predicted the choices and response times for all the odd-numbered trials (test). The purpose of this out of sample prediction was to validate the individually estimated parameters, by comparing the GLAM's predictions to response data that did not inform the parameter estimates. To compensate for the resulting loss of

training data we fitted both GLAM variants using a hierarchical Bayesian framework (Kruschke, 2014; Wiecki, Sofer, & Frank, 2013). Here, individual model parameters and their group distributions are simultaneously estimated from the data. This is desirable as individual parameter estimates are informed by their distribution at the group level, thereby capitalizing on the information that is shared across individuals. With the hierarchical parameter estimates we then tested whether individual behavioral patterns across the three metrics are accurately predicted by the two models (see Figures 2 & 5).

We found that the full GLAM variant accurately predicted individual differences in response times ($\beta = 1.06$, $t(28) = 26.17$, $P < 10^{-20}$; Figure 5A). Similarly, individual differences in the probability of choosing the highest rated item were predicted precisely ($\beta = 0.75$, $t(28) = 8.57$, $P < 10^{-8}$; Figure 5B). Lastly, we also found that the full GLAM predicted individual differences in the influence of gaze on choice probability well ($\beta = 0.98$, $t(28) = 6.69$, $P < 10^{-6}$; Figure 5C).

The restricted GLAM variant with no gaze bias predicted the participants' individual response times and the probability of choosing the highest rated item similarly well (RT: $\beta = 1.02$, $t(28) = 27.07$, $P < 10^{-20}$, Figure 5A; P(choose best): $\beta = 0.75$, $t(28) = 6.53$, $P < 10^{-6}$, Figure 5B). However, the model failed to predict the influence of gaze on the participants' choices ($\beta = -2.97$, $t(28) = -1.53$, $P = 0.14$; Figure 5C), resulting in no correlation between the predicted and empirical data in our gaze influence measure.

These results showed that the full GLAM outperformed the restricted model variant in accurately predicting the participants' empirical choices, as it also captured empirical choice patterns that are driven by gaze and not solely by the items' liking rating.



297

298 *Figure 6*: Correlations between individuals response behavior in the odd-numbered
299 trials and the model parameters estimated from the even-numbered trials. **A:** Response time
300 and *v* (plotted on a log-log-scale). **B:** Probability of choosing the highest rated item and γ. **C:**
301 Influence of gaze on choice probability (mean increase in choice probability for an item that
302 is fixated on longer than the others, when corrected for the influence of the item's relative
303 value and the range of the other items' values) and γ.

304

305 **GLAM explains individual choice behavior**

306 We found that the GLAM accurately predicted individuals' response behavior. Next,
307 we tested whether the individual model parameters are able to explain variability in the
308 participants' choice behavior. Here, we used standard OLS regressions to predict the three
309 behavioral metrics in the odd-numbered trials from the individual GLAM parameters that
310 were previously estimated hierarchically from the even-numbered trials (see Figure 6). We
311 found that *v* (velocity parameter; see Methods for details) scaled logarithmically with the
312 participants mean response time ($\beta = -0.89$, $t(28) = -18.87$, $P < 10^{-16}$; Figure 6A). We did not
313 find a meaningful relationship between the individual σ estimates and the probability of
314 choosing the highest rated item ($\beta = -1.59$, $t(28) = -0.19$, $P = 0.85$), even though the σ
315 parameter determines the magnitude of noise in the accumulation process (see Methods for
316 details). However, we found that γ (gaze bias) estimates predicted the participants'

probabilities of choosing the highest rated item ($\beta = 0.18$, $t(28) = 4.98$, $P < 10^{-4}$; Figure 6B), so that stronger gaze biases (smaller γ) were associated with more choices that were inconsistent with the item ratings. This relationship can be explained as follows: the gaze bias parameter γ allows the model to bias the choice process according to the distribution of gaze between items: with a strong gaze bias, the model's predictions are strongly dependent on the distribution of gaze, and a gaze distribution that is random with respect to the items' liking ratings then leads to random choices. On the other hand, the model's predictions are independent of gaze when no gaze bias is present. The model then neglects gaze and predicts choices solely driven by liking ratings. Lastly, as expected, we also found that γ estimates predicted the participants' individual scores of gaze influence on choice probability ($\beta = -0.31$, $t(28) = -5.68$, $P < 10^{-5}$; Figure 6C).

Discussion

Here, we investigated individual differences in the influence of gaze allocation on simple economic choice behavior by analyzing a previously published data set, where individuals made choices between three snack food items. We found that individuals showed an overall positive relationship between gaze and choice (longer gaze increases choice probability), but that the strength of this relationship was highly variable across individuals. To better understand the computational mechanism underlying this effect and its variability, we proposed a new model called the Gaze-weighted Linear Accumulator Model (GLAM). It assumes that individuals accumulate evidence in favor of each available item and make a choice as soon as the cumulative evidence for one item reaches a choice threshold. Importantly, the accumulation process is biased by gaze behavior, with discounted accumulation rates for unattended items. We found that the GLAM accurately predicts individuals' choice and response time data and does so better than a model that does not assume any influence of gaze allocation on choice. We also found that the GLAM's gaze bias

343 estimates reliably explained individual differences in choice behavior, namely, the strength of
344 the individuals' association of gaze and choice behavior and the individuals' probability of
345 choosing the highest rated item in a choice set (stronger gaze biases were generally associated
346 with more choices that were inconsistent with item ratings).

347 With the GLAM, we have provided a model that captures individual choice behavior
348 in simple economic choice tasks with multiple alternatives with high predictive accuracy by
349 integrating information about the individuals' allocation of gaze. It is statistically and
350 computationally tractable, making it readily extendable to novel choice tasks and research
351 questions.

352 Our individual model comparison revealed the added value of a gaze bias mechanism
353 in decision models. The large majority of the participants were better described by the full
354 model, compared to a restricted variant without any influence of gaze on choice. One reason
355 for this superior performance is that the GLAM's use of the individual trial gaze data allows
356 the model to make different predictions across otherwise identical choice sets. In this way, the
357 model is able to explain variance in behavior that would otherwise be attributed to unspecified
358 decision noise. As decision making can be seen as a stochastic process (Rieskamp, 2008),
359 choices across identical trials with high difficulty (where ratings for the available items are
360 very similar, for example, 3, 2 and 2 for the left, middle and right item, respectively) can be
361 assumed to vary. A stochastic choice model without a gaze bias mechanism will make
362 probabilistic, but identical predictions for two such trials (Gluth & Rieskamp, 2017).
363 Leveraging a gaze bias mechanism, however, allows a model to make trial specific
364 predictions, and these, relying on the generally positive relationship between gaze and choice,
365 will have higher accuracy. We expect the gaze bias mechanism to be especially relevant in
366 high-difficulty trials, where item rating information by itself provides little evidence in favor
367 of any one alternative.

Naturally, the inclusion of gaze information into choice models leads to the question of what drives gaze during choice. This question has received considerable attention in previous research, confirming influences of item surface size (Lohse, 1997; Wedel & Pieters, 2008), position (Chandon, Hutchinson, Bradlow, & Young, 2009) and saliency (Itti & Koch, 2000) on gaze. Still, a formal integration with computationally formalized choice models was achieved only recently (Towal et al., 2013). We believe that the development of generative models for the fixation process itself, and their integration with choice models, has the potential to largely improve existing models.

Our analyses also confirmed the need for individual model fits: we found substantial variability across individuals in the influence of gaze on choice that was hidden in the group level analyses. Given that the influence of gaze on choice is variable between participants, a single gaze bias parameter γ for the whole group would not fit all the individuals well and would therefore result in inferior predictive performance of the model: participants whose link between gaze allocation and choice behavior is weaker than the group average would falsely be predicted to make choices less consistent with their value ratings, and driven more by looking behavior. Predictions for the participants' choices with a stronger link than the group average, on the other hand, would not contain enough influence of gaze. Accounting for individual differences in the link between gaze allocation and choice behavior opens important avenues for future research, focusing on the specific determinants of these differences. For example, are these differences best characterized as a trait (stable within a person, but variable between persons), state (variable within a person, between different situations or contexts) or both (variable between persons and contexts) (see Peters & Büchel, 2011, for a similar discussion in the context of delay discounting)?

Despite a wealth of findings exploring the computational mechanisms underlying simple choice behavior and its link to visual fixations (e.g., Shimojo et al., 2003; Armel et al. 2008; Krajbich et al., 2010; Krajbich & Rangel, 2011; Towal et al., 2013; Cavanagh et al.,

2014; Fisher, 2017), most of this work, and the associated computational frameworks (e.g., Ratcliff et al., 2016), is difficult to extend to complex choice scenarios (i.e., involving more than two choice alternatives). Here, we have shown that the GLAM captures individuals' choice behavior well in choice situations with only few choice alternatives. However, the GLAM naturally extends to choices involving many more options, as we mostly encounter in our everyday lives. Imagine standing in front of a vending machine to buy a snack. These machines can easily store up to 20 items. We assume that in these multialternative choice situations, both gaze and individual differences will play an even more prominent role: individuals, when confronted with large choice sets, do not always look at all available items (e.g., Reutskaja, Nagel, Camerer, & Rangel, 2011). A choice model that considers individuals' liking values only will therefore fail in accurately predicting individuals' choice behavior. A model that includes information about individuals' gaze distribution during decision formation will, on the other hand, outperform such naive models, because it will better account for the set of items that individuals actually consider for a choice. In addition, we assume that behavioral differences between individuals to increase with increasing choice set size. For example, we assume that some individuals may look at only a few of the available items, before making a choice, while others may spend a long time searching for the most highly valued option (as indicated in Reutskaja et al., 2011). To understand whether there is a common choice mechanism underlying these different types of choice behavior, it is necessary to test the ability of a model to capture individual choice patterns.

Real life choices have another level of complexity often not considered in simple economic choice tasks; options comprise multiple, oftentimes orthogonal attributes. For example, each item in a vending machine is associated with a price that has to be considered. Consumer goods often also include other attributes (e.g., ratings by other consumers: De Martino, Bobadilla-Suarez, Nouguchi, Sharot, & Love, 2017; energy consumption, etc.), which are commonly displayed visually to the decision maker. Certain configurations of

option attributes can induce systematic shifts in preference, called context effects (Mohr, Heekeren, & Rieskamp, 2017; Soltani, De Martino, & Camerer, 2012; Trueblood, Brown, Heathcote, & Busemeyer, 2013). These preference shifts can vary considerably between individuals (Mohr et al., 2017). Notably, eye-tracking data about the identity and sequence of fixated attributes are predictive of choice in context effect settings (Noguchi & Stewart, 2014). Future research on these effects and their relationship to gaze requires a model that can be fitted on an individual basis and is applicable to multialternative choice scenarios. The GLAM provides a starting-point to explore these types of research questions in the future.

Like other existing decision making models (e.g., Usher and McClelland, 2001; Roe, Busemeyer & Townsend, 2001; Ratcliff et al., 2016; Krajbich et al., 2010), the GLAM also incorporates several assumptions about the neural computations underlying simple economic choices. It is necessary to evaluate the plausibility of these assumptions, next to the ability of a model to capture individuals' choice behavior. The GLAM assumes evidence accumulation towards a decision threshold and a fixation-dependent bias of this process. There is strong neural evidence for accumulation-to-bound processes during decision formation in a variety of choice tasks (e.g., Basten, Biele, Heekeren, & Fiebach, 2010; Philiastides & Sajda, 2007; Churchland, Kiani & Shadlen, 2008; Liu & Pleskac, 2011; O'Connell, Dockree & Kelly, 2012; Wyart, De Gardelle, Scholl & Summerfield, 2012; Polanía, Krajbich, Grueschow & Ruff, 2014; Lafuente, Jazayeri & Shadlen, 2015; for a review see Gold & Shadlen, 2001, and Heekeren, Marrett & Ungerleider, 2008). Recently, it was also shown that single-trial EEG components reflecting attention in simple perceptual decision-making tasks explain variance in single-trial evidence accumulation rates of the decision process (Nunez, Vandekerckhove, & Srinivasan, 2017) and that variability in these components can explain behavioral differences between individuals (Nunez, Srinivasan, & Vandekerckhove, 2015). Two recent studies also provided first empirical evidence that value-driven activity in the orbitofrontal cortex of monkeys is modulated by fixation location when they viewed reward-associated

visual cues in a free-viewing paradigm (Hunt, Malalasekera, Berker, Miranda, Farmer, Behrens & Kennerley, 2017; McGinty, Rangel, & Newsome, 2016). Together, these studies provided the first neurobiological evidence of the influence of visual fixations on the process of decision formation. Ultimately, a better understanding of these computations will be central to building holistic models of the choice process and for advancing existing choice frameworks. In addition, it might also help us to better understand the origin of behavioral variability that we observe within and between individuals.

The focus on individual differences in the relationship between gaze and choice behavior can also prove itself relevant in clinical research domains. Increasingly prevalent clinical conditions, such as type 2 diabetes and obesity, typically involve maladaptive decision-making between visually presented stimuli, often strategically positioned and designed to capture attention. Snack food items, for example, are advertised with bright, salient colors and placed prominently (e.g., at eye level, near the checkout in the supermarket), which could have adverse effects on individuals prone to making maladaptive food choices. Healthier diets (i.e., food choices) are both prevention and treatment for such diseases, and a better understanding of how individuals' decisions are impacted by looking behavior could help inform the search for predictors of clinical behavior and improve therapeutic approaches. In addition, individually tailored therapeutic approaches, based on a better understanding of individual response patterns, promise higher efficacy and, in turn, reduced health care spending.

Methods

Data, tasks, procedure & preprocessing

We reanalyzed a data set that was previously published in Krajbich and Rangel (2011). In the corresponding experiment, hungry participants made repeated choices between

multiple snack food items (e.g., Twix, Lays Chips, Skittles, etc.), while their eye movements were recorded.

The data set contains data from 30 Caltech students, who reported to regularly eat the snack foods that were used in the experiment and had no dietary restrictions. The participants received a show-up fee of $20 and one food item. The experiment was approved by Caltech's Human Subjects Internal Review Board.

All participants were asked not to eat for 3 hours prior to the experiment. In an initial liking rating task the participants indicated liking ratings between -10 to 10 for each of the 70 different snack food items using an on screen slider with a randomized starting point and free response time ("How much would you like to eat this at the end of the experiment?"; Figure 1, Task 1). These ratings were used as a measure of the value participants placed on each item. In the subsequent choice task the participants made choices between triplets of food items. The items were arranged in a triangular fashion on the screen (Figure 1, Task 2). In one half of the trials, this triangle pointed upwards (center option on top), in the other half it pointed downwards (center option at the bottom). Choices were indicated with free response times and using the left, down and right arrow keys on a keyboard. Each trial began with a 2 s forced fixation towards the center of the screen. A yellow feedback box was shown around the chosen item for 1 s after a choice was made. Lastly, the participants were required to stay for 30 min after the experiment, to eat a food item that they chose in one randomly selected choice trial. The participants performed 100 choice trials each.

The participants' eye movements were continuously recorded with a 50 Hz desktop-mounted Tobii eye tracker.

The data were obtained from the original authors in an already preprocessed format. The original preprocessing steps included the removal of trials with missing fixation data for more than 500 ms at the beginning or end of the trial, resulting in a total of 2966 remaining trials (mean ± s.e.m. number of trials dropped per participant was 1.1 ± 0.9). Rectangular

areas of interest (AOIs) were constructed around each food item in each trial and visual fixations were assigned to the corresponding item or coded as non-item fixations. If a non-item fixation was preceded and succeeded by fixations on the same item, the non-item fixation would also be assigned to this item. Other non-item fixations were not reassigned and discarded from all further analyses.

Gaze-weighted Linear Accumulator Model (GLAM) details

The GLAM belongs to the class of linear stochastic race models (Usher & McClelland, 2001). It assumes accumulation of noisy evidence in favor of each available alternative *i*, and that the choice is determined by the first accumulator that reaches a common boundary. In particular, we define the accumulated relative evidence E_i in favor of alternative *i*, as a stochastic process that changes at each point in time *t* according to:

$$E_i(t) = E_i(t-1) + v \times R_i + N(0,\sigma^2); E_i(0) = 0 \quad (1)$$

E_i consists of two separate components: a drift term R_i and zero-centered normally distributed noise with standard deviation σ. The overall speed of the accumulation process is governed by the velocity parameter *v*. The drift term R_i describes the average amount of relative evidence for item *i* that is accumulated at each point in time *t*. We define the relative evidence R_i^* as the difference in the stationary absolute evidence signal A_i of item *i* and the maximum absolute evidence of all other items *J*:

$$R_i^* = A_i - max_J(A_J) \quad (2)$$

The model's gaze bias mechanism is implemented in the absolute evidence signal A_i: Similar to the aDDM, the absolute evidence signals are assumed to be proportional to the value ratings r_i, and crucially, switch between two different states during the trial: an unbiased state, when an item is currently looked at, and a biased state, when gaze is directed

towards a different item. Therefore, on average, A_i is a linear combination of two terms that are weighted by the fraction g_i of the total fixation time that item *i* was fixated in the trial:

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

Here, γ ($-1 \leq \gamma \leq 1$) is the model's gaze bias parameter that determines the strength of the downweighting during the biased state. If $\gamma = 1$, there is no difference between the biased and unbiased state, producing no gaze bias. If $\gamma < 1$, the absolute evidence signal is discounted by the γ parameter, resulting in a gaze bias. If $-1 \leq \gamma < 0$, the sign of the evidence signal changes, thereby leaking evidence, when the item is not fixated. This leakage mechanism is supported by a recent empirical study (Ashby et al., 2016). The maximum amount of evidence that can be accumulated or leaked at each time point is symmetric in magnitude, as the γ parameter is bounded between -1 and 1.

Note that the range of possible R_i^* (equation (2)) depends on the participants' use of the item rating scale: if the ratings only cover a narrow range of possible values on the scale, the relative evidence values R_i^* will likewise be small, whereas they will be large if the participant utilizes the entire range of the rating scale. GLAM assumes an adaptive representation of the relative evidence signals that is compensating for the participants' use of the rating scale and thereby sensitive to marginal differences in the relative evidences, particularly to values close to 0 (where the absolute evidence signal for one item is only marginally different to the maximum of all others). To this end, a logistic transform $s(x)$, with scaling parameter τ is applied:

$$s(x) = \frac{1}{1+\exp(-\tau x)} \quad (4)$$

$$R_i = s(R_i^*) \quad (5)$$

The first passage time density $f_i(t)$ of a single linear stochastic accumulator E_i, with decision boundary *b*, is given by the Inverse Gaussian Distribution (Wald, 1973):

$$f_i(t) = \left[\frac{\lambda}{2\pi t^3}\right]^{1/2} exp\left\{\frac{-\lambda(t-\mu)^2}{2\mu^2 t}\right\} \text{ with } \mu = \frac{b}{vR_i} \text{ and } \lambda = \frac{b^2}{\sigma^2} \quad (6)$$

546 However, this density does not take into account that there are multiple accumulators
547 in each trial racing towards the same boundary. As soon as any of these accumulators crosses
548 the boundary a choice is made and the trial ends. For this reason, $f_i(t)$ must be corrected for
549 the probability that any other accumulator crosses the boundary first. The probability that a
550 single accumulator crosses the boundary prior to *t*, is given by its cumulative distribution
551 function $F_i(t)$:

552 $$F_i(t) = \Phi\left(\sqrt{\frac{\lambda}{t}}\left(\frac{t}{\mu} - 1\right)\right) + \exp\left(\frac{2\lambda}{\mu}\right)\Phi\left(-\sqrt{\frac{\lambda}{t}}\left(\frac{t}{\mu} + 1\right)\right), \quad (7)$$

553 where $\Phi(x)$ is the standard normal cumulative distribution function. Hence, the joint
554 probability $p_i(t)$ that accumulator E_i crosses *b* at time *t*, and that no other accumulator *j* has
555 reached *b* first, is given by:

556 $$p_i(t) = f_i(t) \prod_j (1 - F_j(t)). \quad (8)$$

557 We performed a parameter recovery study to rule out misspecifications of the model
558 and assert the validity of the parameters estimated from empirical data. All of the parameters
559 could be recovered to a satisfying degree (see Figure S1 for detailed results).

560 Although the race framework deviates from the classical Drift-Diffusion Model
561 (DDM; Ratcliff, 1978; Ratcliff et al., 2016), which is known to implement an optimal
562 decision procedure in the sense of the sequential probability ratio test (Bogacz, Brown,
563 Moehlis, Holmes, & Cohen, 2006), it has reasonable benefits in the context of this paper: first,
564 the model naturally generalizes to choices between three alternatives, which is not trivial for
565 the classical DDM. Second, it generalizes to settings with even larger choice sets. Third, an
566 analytical solution for the first-passage time density of the linear stochastic race exists. This
567 solution enables a very fast and efficient parameter estimation, without the need to
568 numerically estimate densities from a large number of model simulations.

569

570

GLAM parameter estimation

Individual. For the individual model comparison, we first estimated the model parameters at the individual level. The full GLAM has four parameters (v, γ, σ, τ). The individual models were implemented in the Python library PyMC3 (Salvatier, Wiecki, & Fonnesbeck, 2016) and fitted using the default Markov Chain Monte Carlo No-U-Turn-Sampler (NUTS; Hoffman & Gelman, 2014). We parameterized the model so that the noise parameter σ was sampled proportionally to the velocity parameter v using a signal-to-noise variable *SNR* (note that this does not add a free parameter to the model, as σ is now fully determined by v and *SNR*). We placed uninformative, uniform priors between sensible limits on all model parameters:

$v \sim$ Uniform(1e-10, 0.01)

$\gamma \sim$ Uniform(-1, 1)

$SNR \sim$ Uniform(1, 500)

$\sigma = v \times SNR$

$\tau \sim$ Uniform(0, 5)

Further, we assumed a fixed 5% rate of error trials, which we model as a participant-specific uniform likelihood distribution $u_s(t)$. This error likelihood describes the probability of a random choice for any of the N available choice items at a random time point in the interval of empirically observed response times (cf. Ratcliff & Tuerlinckx, 2002; Wiecki et al., 2013):

$$u_s(t) = \frac{1}{N(\max(rt_s) - \min(rt_s))} \quad (9)$$

The resulting choice likelihood is then given by:

$$l_i(t) = 0.95 \times p_i(t) + 0.05 \times u_s(t) \quad (10)$$

For each individual model, the NUTS sampler was initialized using the default behavior in PyMC 3.2, followed by 500 tuning samples that were discarded. Finally, we drew 2000 posterior samples that we used to estimate the model parameters.

In addition, a restricted no-gaze-bias GLAM variant was also fit to the individual data. It was specified and fitted identically to the full model, but had the gaze-bias parameter γ fixed at 1.0. The reported parameter estimates are maximum a posteriori (MAP) estimates.

Hierarchical. We also estimated the GLAM parameters in a hierarchical Bayesian framework (Kruschke, 2014; Vandekerckhove, Tuerlinckx, & Lee, 2008; Wiecki et al., 2013). Here, the fit of participant level parameters is informed by the group distribution of parameters. Each parameter on the participant level is modeled as coming from a population distribution whose shape and location are also estimated from the data. We assumed that all participant level parameters are drawn from normal population distributions, which we bounded to sensible ranges:

$v \in$ (1e-10, 0.01), $SNR \in$ (1, 500), $\gamma \in$ (-1, 1), $\tau \in$ (0, 5).

We used ADVI (Kucukelbir, Ranganath, Gelman, & Blei, 2015) from the PyMC3 python library (Salvatier et al., 2016) to approximate the posterior distribution of the model parameters. In analogy to the individual models, we assumed a fixed 5% rate of participant-specific uniformly distributed error responses (see above). The reported parameter estimates are maximum a posteriori (MAP) estimates.

Model simulations

Choice and response time data was simulated from the GLAM according to the following procedures: each trial in the left-out data set, containing all the odd-numbered trials, was repeated 50 times. For every trial the model used the empirically observed item ratings and gaze distributions. With a fixed rate of 5% the simulation produced a random choice and a random response time between the participant's minimum and maximum observed response

times (cf. equations (9) & (10)). With a rate of 95% the choice and response time were simulated from the actual GLAM: for each item in the trial, a first passage time (FPT) was drawn according to the single-item first passage densities (equation (8)). The response time and choice were then determined by the item with the shortest FPT.

Availability of data, model and analysis code

All analyses and figures can be reproduced using the data set, scripts and GLAM resources that are available at http://www.github.com/glamlab/glam.

Software

All analyses were performed in Python, using the NumPy and SciPy (Van der Walt, Colbert & Varoquaux, 2011), Pandas (McKinney, 2010), Statsmodels (Skipper & Perktold, 2010), PyMC3 (Salvatier et al., 2016) and Theano (Theano Development Team, 2016) libraries. We used Matplotlib (Hunter, 2007) for visualization.

References

Armel, K. C., Beaumel, A., & Rangel, A. (2008). Biasing simple choices by manipulating relative visual attention. *Judgment and Decision Making, 3*(5), 396-403.

Ashby, N. J. S., Jekel, M., Dickert, S., & Glöckner, A. (2016). Finding the right fit: A comparison of process assumptions underlying popular drift-diffusion models. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42*(12), 1982-1993. http://doi.org/10.1037/xlm0000279

Basten, U., Biele, G., Heekeren, H. R., & Fiebach, C. J. (2010). How the brain integrates costs and benefits during decision making. *Proceedings of the National Academy of Sciences, 107*(50), 21767-21772. http://doi.org/10.1073/pnas.0908104107

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review, 113*(4), 700-765. http://doi.org/10.1037/0033-295X.113.4.700

Cavanagh, J. F., Wiecki, T. V., Kochar, A., & Frank, M. J. (2014). Eye tracking and pupillometry are indicators of dissociable latent decision processes. *Journal of Experimental Psychology: General, 143*(4), 1476-1488. http://doi.org/10.1037/a0035813

Chandon, P., Hutchinson, J. W., Bradlow, E. T., & Young, S. H. (2009). Does in-store marketing work? Effects of the number and position of shelf facings on brand attention and evaluation at the point of purchase. *Journal of Marketing, 73*(6), 1-17. http://doi.org/10.1509/jmkg.73.6.1

Churchland, A. K., Kiani, R., & Shadlen, M. N. (2008). Decision-making with multiple alternatives. *Nature Neuroscience, 11*(6), 693-702. http://doi.org/10.1038/nn.2123

Fiedler, S., & Glöckner, A. (2012). The dynamics of decision making in risky choice: An eye-tracking analysis. *Frontiers in Psychology, 3*. http://doi.org/10.3389/fpsyg.2012.00335

661 Fisher, G. (2017). An attentional drift diffusion model over binary-attribute choice. *Cognition, 168*, 34–45. http://doi.org/10.1016/j.cognition.2017.06.007

663 Folke, T., Jacobsen, C., Fleming, S. M., & De Martino, B. (2016). Explicit representation of confidence informs future value-based decisions. *Nature Human Behaviour, 1*. http://doi.org/10.1038/s41562-016-0002

666 Glöckner, A., & Herbold, A.-K. (2011). An eye-tracking study on information processing in risky decisions: Evidence for compensatory strategies based on automatic processes. *Journal of Behavioral Decision Making, 24*(1), 71–98. http://doi/10.1002/bdm.684

669 Gluth, S., & Rieskamp, J. (2017). Variability in behavior that cognitive models do not explain can be linked to neuroimaging data. *Journal of Mathematical Psychology, 76*, 104–116. https://doi.org/10.1016/j.jmp.2016.04.012

672 Gold, J. I., & Shadlen, M. N. (2001). Neural computations that underlie decisions about sensory stimuli. *Trends in Cognitive Sciences, 5*(1), 10–16. https://doi.org/10.1016/S1364-6613(00)01567-9

675 Grandy, T., Lindenberger, U., & Werkle-Bergner, M. (2017, April 11). When group means fail: Can one size fit all? *bioRxiv*. https://doi.org/10.1101/126490

677 Hayes, K. J. (1953). The backward curve: A method for the study of learning. *Psychological Review, 60*(4), 269-275. http://doi.org/10.1037/h0056308

679 Heekeren, H. R., Marrett, S., & Ungerleider, L. G. (2008). The neural systems that mediate human perceptual decision making. *Nature Reviews Neuroscience, 9*(6), 467–479. http://doi.org/10.1038/nrn2374

682 Hoffman, M. D., & Gelman, A. (2014). The No-U-turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research, 15*(1), 1593–1623.

Hunt, L. T., Malalasekera, W. M. N., Berker, A. O. de, Miranda, B., Farmer, S. F., Behrens, T. E. J., & Kennerley, S. W. (2017, August 1). Triple dissociation of attention and decision computations across prefrontal cortex. *bioRxiv*. http://doi.org/10.1101/171173

Hunter, J. D. (2007). Matplotlib: a 2D graphics environment. *Computing in Science & Engineering*, *9*, 90-95. https://doi.org/10.1109/MCSE.2007.55

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research, 40*(10), 1489–1506. https://doi.org/10.1016/S0042-6989(99)00163-7

Konovalov, A., & Krajbich, I. (2016). Gaze data reveal distinct choice processes underlying model-based and model-free reinforcement learning. *Nature Communications, 7*. http://doi.org/10.1038/ncomms12438

Krajbich, I., & Rangel, A. (2011). Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions. *Proceedings of the National Academy of Sciences*, *108*(33), 13852–13857. http://doi.org/10.1073/pnas.1101328108

Krajbich, I., Armel, C., & Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature Neuroscience*, *13*(10), 1292–1298. http://doi.org/10.1038/nn.2635

Krajbich, I., Lu, D., Camerer, C., & Rangel, A. (2012). The attentional drift-diffusion model extends to simple purchasing decisions. *Frontiers in Psychology*, *3*. http://doi.org/10.3389/fpsyg.2012.00193

Kruschke, J. (2014). *Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan*. Academic Press.

Kucukelbir, A., Ranganath, R., Gelman, A., & Blei, D. (2015). Automatic variational inference in Stan. In *Advances in Neural Information Processing Systems* (pp. 568-576).

de Lafuente, V., Jazayeri, M., & Shadlen, M. N. (2015). Representation of accumulating evidence for a decision in two parietal areas. *Journal of Neuroscience*, *35*(10), 4306-4318. https://doi.org/10.1523/JNEUROSCI.2451-14.2015

Lewandowsky, S., & Farrell, S. (2010). *Computational modeling in cognition: Principles and practice*. Sage Publications.

Liu, T., & Pleskac, T. J. (2011). Neural correlates of evidence accumulation in a perceptual decision task. *Journal of Neurophysiology*, *106*(5), 2383-2398. http://doi.org/10.1152/jn.00413.2011

Lohse, G. L. (1997). Consumer eye movement patterns on yellow pages advertising. *Journal of Advertising*, *26*(1), 61–73. https://doi.org/10.1080/00913367.1997.10673518

De Martino, B., Bobadilla-Suarez, S., Nouguchi, T., Sharot, T., & Love, B. C. (2017). Social information is integrated into value and confidence judgments according to its reliability. *Journal of Neuroscience*, *37*(25), 6066-6074. https://doi.org/10.1523/JNEUROSCI.3880-16.2017

McGinty, V. B., Rangel, A., & Newsome, W. T. (2016). Orbitofrontal Cortex Value Signals Depend on Fixation Location during Free Viewing. *Neuron*, *90*(6), 1299-1311. https://doi.org/10.1016/j.neuron.2016.04.045

McKinney, W. (2010, June). Data structures for statistical computing in Python. In *Proceedings of the 9th Python in Science Conference* (Vol. 445, pp. 51-56). Austin, TX.

Mohr, P. N. C., Heekeren, H. R., & Rieskamp, J. (2017). Attraction Effect in Risky Choice Can Be Explained by Subjective Distance Between Choice Alternatives. *Scientific Reports*, *7*(1), 8942. http://doi.org/10.1038/s41598-017-06968-5

Noguchi, T., & Stewart, N. (2014). In the attraction, compromise, and similarity effects, alternatives are repeatedly compared in pairs on single dimensions. *Cognition*, *132*(1), 44-56. https://doi.org/10.1016/j.cognition.2014.03.006

737 Nunez, M. D., Srinivasan, R., & Vandekerckhove, J. (2015). Individual differences in
738 attention influence perceptual decision making. *Frontiers in Psychology, 8*.
739 http://doi.org/10.3389/fpsyg.2015.00018
740 Nunez, M. D., Vandekerckhove, J., & Srinivasan, R. (2017). How attention influences
741 perceptual decision making: Single-trial EEG correlates of drift-diffusion model
742 parameters. *Journal of Mathematical Psychology*, *76*, 117-130.
743 https://doi.org/10.1016/j.jmp.2016.03.003
744 O'Connell, R. G., Dockree, P. M., & Kelly, S. P. (2012). A supramodal accumulation-to-
745 bound signal that determines perceptual decisions in humans. *Nature*
746 *Neuroscience*, *15*(12), 1729-1735. http://doi.org/10.1038/nn.3248
747 Pärnamets, P., Johansson, P., Hall, L., Balkenius, C., Spivey, M. J., & Richardson, D. C.
748 (2015). Biasing moral decisions by exploiting the dynamics of eye gaze. *Proceedings*
749 *of the National Academy of Sciences*, *112*(13), 4170-4175.
750 http://doi.org/10.1073/pnas.1415250112
751 Peters, J., & Büchel, C. (2011). The neural mechanisms of inter-temporal decision-making:
752 understanding variability. *Trends in Cognitive Sciences*, *15*(5), 227-23.
753 https://doi.org/10.1016/j.tics.2011.03.002
754 Philiastides, M. G., & Sajda, P. (2007). EEG-informed fMRI reveals spatiotemporal
755 characteristics of perceptual decision making. *Journal of Neuroscience*, *27*(48),
756 13082-13091. https://doi.org/10.1523/JNEUROSCI.3540-07.2007
757 Polanía, R., Krajbich, I., Grueschow, M., & Ruff, C. C. (2014). Neural oscillations and
758 synchronization differentially support evidence accumulation in perceptual and value-
759 based decision making. *Neuron*, *82*(3), 709-720.
760 https://doi.org/10.1016/j.neuron.2014.03.014
761 Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*(2), 59.
762 http://doi.org/10.1037/0033-295X.85.2.59

763 Ratcliff, R., Smith, P. L., Brown, S. D., & McKoon, G. (2016). Diffusion decision model:
764 current issues and history. *Trends in Cognitive Sciences*, *20*(4), 260-281.
765 https://doi.org/10.1016/j.tics.2016.01.007
766 Ratcliff, R., & Tuerlinckx, F. (2002). Estimating parameters of the diffusion model:
767 Approaches to dealing with contaminant reaction times and parameter
768 variability. *Psychonomic Bulletin & Review*, *9*(3), 438-481.
769 https://doi.org/10.3758/BF03196302
770 Reutskaja, E., Nagel, R., Camerer, C. F., & Rangel, A. (2011). Search dynamics in consumer
771 choice under time pressure: An eye-tracking study. *American Economic Review*, *101*,
772 900-926. http://doi.org/10.1257/aer.101.2.900
773 Rieskamp, J. (2008). The probabilistic nature of preferential choice. *Journal of Experimental*
774 *Psychology: Learning, Memory, and Cognition*, *34*(6), 1446.
775 http://doi.org/10.1037/a0013646
776 Roe, R. M., Busemeyer, J. R., & Townsend, J. T. (2001). Multialternative decision field
777 theory: A dynamic connectionist model of decision making. *Psychological*
778 *Review*, *108*(2), 370. http://doi.org/10.1037/0033-295X.108.2.370
779 Salvatier, J., Wiecki, T. V., & Fonnesbeck, C. (2016). Probabilistic programming in Python
780 using PyMC3. *PeerJ Computer Science*, *2*, e55. http://doi.org/10.7717/peerj-cs.55
781 Seabold, S., & Perktold, J. (2010, June). Statsmodels: Econometric and statistical modelling
782 with Python. In *Proceedings of the 9th Python in Science Conference* (Vol. 57, p. 61)
783 Shimojo, S., Simion, C., Shimojo, E., & Scheier, C. (2003). Gaze bias both reflects and
784 influences preference. *Nature Neuroscience*, *6*(12), 1317-1322.
785 http://doi.org/10.1038/nn1150
786 Soltani, A., De Martino, B., & Camerer, C. (2012). A range-normalization model of context-
787 dependent choice: a new model and evidence. *PLoS Computational Biology*, *8*(7),
788 e1002607. https://doi.org/10.1371/journal.pcbi.1002607

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & van der Linde, A. (2014). The deviance information criterion: 12 years on. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *76*(3), 485-493. https://doi.org/10.1111/rssb.12062

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & van der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *64*(4), 583-639. https://doi.org/10.1111/1467-9868.00353

Stewart, N., Gächter, S., Noguchi, T., & Mullett, T. L. (2016). Eye movements in strategic choice. *Journal of Behavioral Decision Making*, *29*(2-3), 137-156. https://doi.org/10.1002/bdm.1901

Stewart, N., Hermens, F., & Matthews, W. J. (2016). Eye movements in risky choice. *Journal of Behavioral Decision Making*, *29*(2-3), 116-136. https://doi.org/10.1002/bdm.1854

Tavares, G., Perona, P., & Rangel, A. (2017). The Attentional Drift Diffusion Model of Simple Perceptual Decision-Making. *Frontiers in Neuroscience*, *11*, 468. https://doi.org/10.3389/fnins.2017.00468

Theano Development Team (2016). Theano: a Python framework for fast computation of mathematical expressions. *arXiv*.

Tillman, G. (2017, September 18). The Racing Diffusion Model of Speeded Decision Making. http://doi.org/10.17605/OSF.IO/XUWBK

Towal, R. B., Mormann, M., & Koch, C. (2013). Simultaneous modeling of visual saliency and value computation improves predictions of economic choice. *Proceedings of the National Academy of Sciences*, *110*(40), E3858-E3867. http://doi.org/10.1073/pnas.1304429110

Trueblood, J. S., Brown, S. D., Heathcote, A., & Busemeyer, J. R. (2013). Not just for consumers: Context effects are fundamental to decision making. *Psychological Science*, *24*(6), 901-908. https://doi.org/10.1177/0956797612464241

814 Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky,
815 competing accumulator model. *Psychological Review, 108*(3), 550.
816 http://doi.org/10.1037/0033-295X.108.3.550
817 Usher, M., Olami, Z., & McClelland, J. L. (2002). Hick's law in a stochastic race model with
818 speed–accuracy tradeoff. *Journal of Mathematical Psychology, 46*(6), 704-715.
819 https://doi.org/10.1006/jmps.2002.1420
820 Vaidya, A. R., & Fellows, L. K. (2015). Testing necessary regional frontal contributions to
821 value assessment and fixation-based updating. *Nature Communications, 6*, 10120.
822 https://doi.org/10.1038/ncomms10120
823 Vandekerckhove, J., Tuerlinckx, F., & Lee, M. (2008, July). A Bayesian approach to
824 diffusion process models of decision-making. In *Proceedings of the 30th Annual*
825 *Conference of the Cognitive Science Society* (pp. 1429-1434). Cognitive Science
826 Society.
827 Von Neumann, J., & Morgenstern, O. (1945). *Theory of games and economic behavior*.
828 Princeton, NJ: Princeton University Press.
829 Wald, A. (1973). *Sequential analysis*. Courier Corporation.
830 van der Walt, S., Colbert, S., C., & Varoquaux, G (2011). The NumPy Array: A Structure for
831 Efficient Numerical Computation. *Computing in Science & Engineering, 13*, 22-30.
832 http://doi.org/10.1109/MCSE.2011.37
833 Wedel, M., & Pieters, R. (2008). A review of eye-tracking research in marketing. In *Review of*
834 *marketing research* (pp. 123-147). Emerald Group Publishing Limited.
835 Wiecki, T. V., Sofer, I., & Frank, M. J. (2013). HDDM: Hierarchical Bayesian estimation of
836 the drift-diffusion model in Python. *Frontiers in Neuroinformatics, 7*.
837 http://doi.org/10.3389/fninf.2013.00014

838 Wyart, V., De Gardelle, V., Scholl, J., & Summerfield, C. (2012). Rhythmic fluctuations in
839 evidence accumulation during decision making in the human brain. *Neuron*, *76*(4),
840 847-858. https://doi.org/10.1016/j.neuron.2012.09.015

Acknowledgements

Peter N. C. Mohr is funded by the Freie Universität Berlin within the Excellence Initiative of the German Research Foundation (DFG). Further funding is provided by the WZB Berlin Social Science Center. Armin W. Thomas is funded by the Bernstein Center for Computational Neuroscience Berlin (BCCN). Felix Molter is supported by the International Max Planck Research School on the Life Course (LIFE). Ian Krajbich is funded by National Science Foundation Career Award 1554837.

Author contributions

A.W.T, F.M., H.R.H. and P.N.C.M. conceived the study. A.W.T. and F.M. designed the model and carried out the data analysis. A.W.T, F.M., I.K., H.R.H. and P.N.C.M. co-wrote the manuscript.

Competing interest statement

The authors declare no competing interests.

Supplementary Information

Parameter recovery

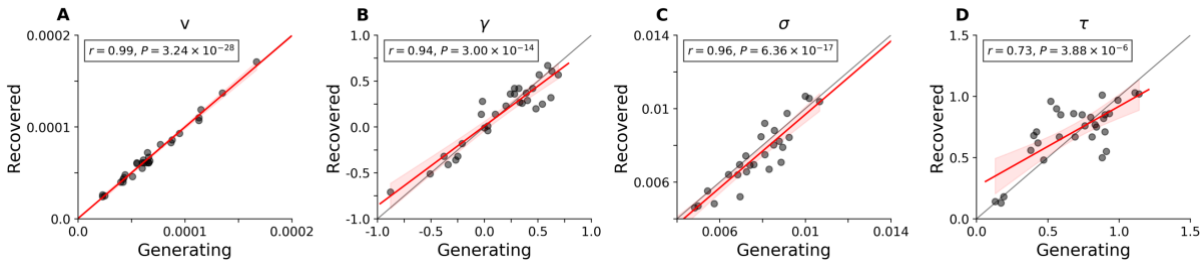


Figure S1: Results of a parameter recovery study of the GLAM. Parameter estimates were estimated from model simulated data sets. Panels **A** to **D** show relationships between generating and recovered parameters. All parameters could be recovered to a satisfying degree.

We performed a parameter recovery study to validate the parameter estimates. We simulated data using the GLAM and the corresponding hierarchically estimated individual parameter estimates. Each empirically observed even-numbered trial (i.e., set of item value ratings and relative gazes in training set trials) was simulated once, resulting in a GLAM-generated data set that matches the original training data in size and structure. We then performed the exact same hierarchical parameter estimation procedure as we did in the training data. Ideally, the recovered parameters should match the originally estimated ones. The correlations between generating and recovered parameters are displayed in Figure S1. The critical model parameters *v*, γ and σ were recovered very well (Figure S1 A, B, C, respectively). We found a significant linear relationship between generating and recovered τ parameters (Figure S1D), although they include some level of variance.

Parameter estimates

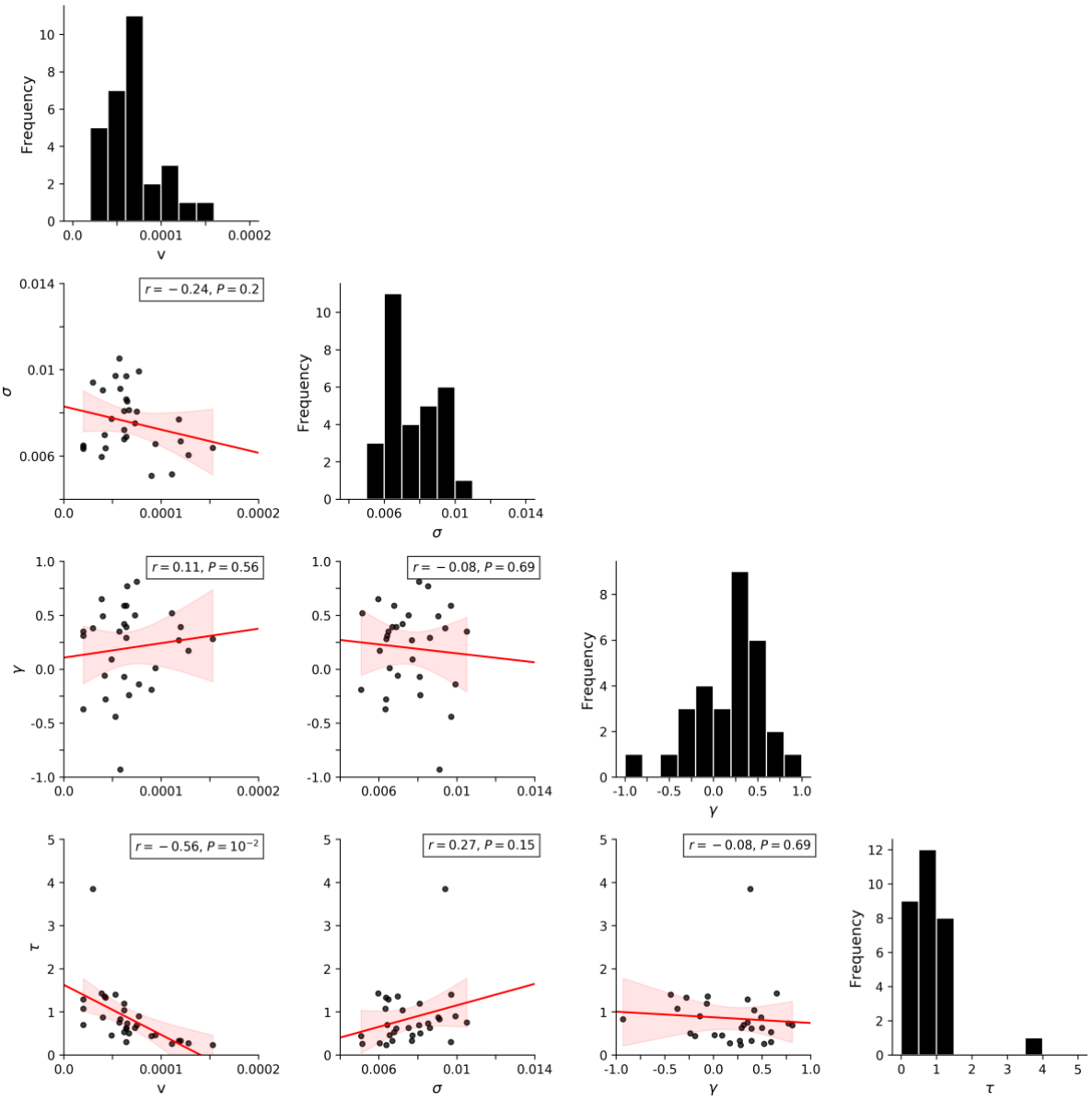


Figure S2: Parameter estimates and correlations between parameters. Estimates shown are from individual model fits.