post-puberty experienced by the parents. Thus interesting exceptions, where adaptive explanations could be contemplated, would be those which failed to agree with this null. Another consequence of the model will be that (provided we have enough life history data for the relevant species) the 'generation time effect', where short-lived species tend to have higher rates of molecular evolution, can be re-examined in a more detailed and predictive way.

### REFERENCES

1. Thomas, G.W.C., Wang, R.J., Puri, A., Harris, R.A., Raveendran, M., Hughes, D., Murali, S., Williams, L., Doddapaneni, H., Muzny, D., *et al*. (2018). Reproductive longevity predicts mutation rates in primates. Curr. Biol. *28*, 3193–3197.

2. Cartwright, R.A., Hussin, J., Keebler, J.E., Stone, E.A., and Awadalla, P. (2012). A family-based probabilistic method for capturing de novo mutations from high-throughput short-read sequencing data. Stat. Appl. Genet. Mol. Biol. *11*, 11.

3. Kong, A., Frigge, M.L., Masson, G., Besenbacher, S., Sulem, P., Magnusson, G., Gudjonsson, S.A., Sigurdsson, A., Jonasdottir, A., Jonasdottir, A., *et al*. (2012). Rate of de novo mutations and the importance of father's age to disease risk. Nature *488*, 471–475.

4. Haldane, J.B.S. (1927). The mathematical theory of natural and artificial selection. Proc. Camb. Philos. Soc. *23*, 838–844.

5. Taddei, F., Radman, M., Maynard-Smith, J., Toupance, B., Gouyon, P.H., and Godelle, B. (1997). Role of mutator alleles in adaptive evolution. Nature *387*, 700–702.

6. Bjedov, I., Tenaillon, O., Gérard, B., Souza, V., Denamur, E., Radman, M., Taddai, F., and Matic, I. (2003). Stress-induced mutagenesis in bacteria. Science *300*, 1404–1409.

7. Martincorena, I., Neshasayee, A.S.N., and Luscombe, N.M. (2012). Evidence of non-random mutation rates suggests an evolutionary risk management strategy. Nature *485*, 95–98.

8. Couce, A., Caudwell, L.V., Feinauer, C., Hindré, T., Feugeas, J.-P., Weigt, M., Lenski, R.E., Schneider, D., and Tenaillon, O. (2017). Mutator genomes decay, despite sustained fitness gains, in a long-term experiment with bacteria. Proc. Natl. Acad. Sci. USA *114*, E9026–E9035.

9. Wielgoss, S., Barrick, J.E., Tenaillon, O., Wiser, M.J., Dittmar, W.J., Cruveiller, S., Chane-Woon-Ming, B., Médigue, C., Lenski, R.E., and Schneider, D. (2013). Mutation rate dynamics in a bacterial population reflect tension between adaptation and genetic load. Proc. Natl. Acad. Sci. USA *110*, 222–227.

10. Lynch, M., Ackerman, M.S., Gout, J.-F., Long, H., Sung, W., Thomas, W.K., and Foster, P.L. (2016). Genetic drift, selection, and the evolution of the mutation rate. Nat. Rev. Genet. *17*, 704–714.

# Selective Attention: A Plausible Mechanism Underlying Confirmation Bias

Genís Prat-Ortega[1,2] and Jaime de la Rocha[1,*]
[1]Institut d'Investigacions Biomèdiques August Pi i Sunyer, Barcelona, Spain
[2]Centre de Recerca Matemàtica, Campus de Bellaterra, Bellaterra, Spain
*Correspondence: jrochav@clinic.cat
https://doi.org/10.1016/j.cub.2018.08.024

**Our perception is strongly influenced by our experience of past stimuli and choices. A new study suggests that our attention is selectively deployed to those aspects of the sensory evidence which are consistent with our previous decisions, thus introducing a confirmation bias.**

Back in 1620, Francis Bacon noted that "The human understanding when it has once adopted an opinion draws all things else to support and agree with it". Although an unbiased assessment of the evidence lies at the core of the scientific method, it has long been suspected that it could be nothing more than a chimera: the mere act of opting for one hypothesis *versus* another may introduce choice-induced changes in the way we perceive the evidence. In particular, as pointed out by Bacon, subjects tend to neglect data that challenge their former conclusion. Understanding this so-called *confirmation bias* might not only help us to evaluate our experiments more

objectively, but may reveal fundamental aspects of how the brain processes sensory information to yield congruent percepts. Although there is ample evidence for the existence of confirmation biases at different cognitive levels [1], the underlying neural mechanisms have been elusive. A new study by Talluri, Urai and colleagues [2], reported in this issue of *Current Biology*, shows that categorical decisions bias the acquisition of new evidence by overweighting the evidence consistent with the decision and underweighting the inconsistent one, akin to the mechanism of selective attention.

The amount of available sensory information in our environment greatly

exceeds the limited processing capacity of our brains. It is for this reason that we are forced to choose which information we deem relevant to be scrutinized at any given moment and which one may not be so important. Selective attention is the way our brain controls this spotlight that targets certain aspects of the available information while neglecting others [3]. Visual attention can be directed to a particular part of the visual field, but also to certain values of a specific feature (what is termed feature-based attention [4]). When looking for your kid in a crowded beach, for instance, focusing your attention on the color of her swimming suit may help to find her
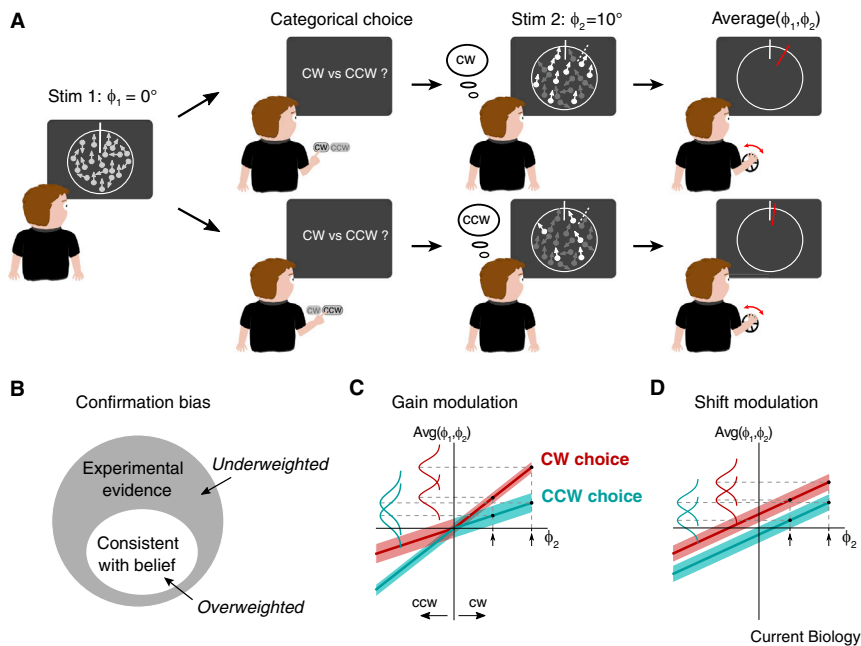
**Figure 1. Confirmation bias induced by a choice-consistent deployment of feature-based attention.**

(A) In Talluri *et al.* [2] subjects first viewed the random dot stimulus 1 and reported whether the direction of motion $\phi_1$ was tilted clockwise (CW) or counterclockwise (CCW) with respect to a reference line (solid white line). After subjects reported their CW *versus* CCW choice, in 50% of the trials a second stimulus was presented with direction $\phi_2$ drawn independently of $\phi_1$ (dashed line showing the direction $\phi_2$ was not displayed). After stimulus 2, subjects had to report their estimate of the average direction of both stimuli Avg($\phi_1,\phi_2$) (red line). Stimulus 2 had greater impact on Avg($\phi_1,\phi_2$) when it was congruent with the first choice (top row) than when it was incongruent (bottom row). Talluri *et al.* [2] suggest that the first choice introduced an attentional bias (see thought bubbles) such that feature-based attention was deployed to those directions consistent with that choice. Thus the dots moving towards the attended directions (consistent) were perceived more saliently (represented as whiter dots) than the dots moving towards the unattended (inconsistent evidence shown as darker dots). (B) Confirmation bias is the tendency to overweight the evidence consistent with our belief or hypothesis and underweight the inconsistent. (C) Schematic showing the estimated Avg($\phi_1,\phi_2$) as a function of $\phi_2$. This function describes the case in which $\phi_1 = 0$ so that Stimulus 1 has no systematic impact on Avg($\phi_1,\phi_2$). Filled bands represent variability in the estimation. A choice-induced selective gain modulation causes the slope of this transfer function to be different for consistent and inconsistent evidence (stimuli $\phi_2 > 0$ are consistent with CW choices but inconsistent with CCW). Consequently, the distributions of the estimated direction of motions for two different $\phi_2$ are more separable in the consistent than inconsistent case (inset distributions) as shown in Figure 2F in [2]. (D) In the choice-induced shift model, the first choice causes a consistent shift in the estimated Avg($\phi_1,\phi_2$). Although some subjects show traces of a shift effect (Figure 2A in [2]), this model cannot explain the increased separability in the distributions of the estimated Avg($\phi_1,\phi_2$) given $\phi_2$ as shown in (C).

probability) and second, the increase in spiking activity had a time-course consistent with a top-down signal and specifically reflected a gain increase, a signature of feature-based attention. This finding suggests that, as subjects accumulate evidence towards one categorical decision, their attention is directed towards the incoming evidence consistent with that decision. Thus, it is the early evidence which triggers a gain increase in the response of consistent neurons, activating a positive reinforcement feedback that results in a primacy effect: the early evidence in the trial has a larger impact on the decision [9–12]. Both categorical decisions and selective attention require a competitive selection. A biologically inspired model implemented both the decision dynamics and the top-down attention-like modulation using the same winner-take-all dynamics between neuronal populations, thus proposing that an attentional bias is intimately linked to perceptual selection [10].

Optimal decoding arguments have also been used, independently of attention, to explain the existence of perceptual biases caused by a previous categorical decision [13]. Participants were asked to make two consecutive reports on the same stimulus: first, whether the motion direction of a random dot stimulus was clockwise (CW) or counterclockwise (CCW), and second, an estimation of the motion direction [13]. Estimated directions were consistently biased towards the category of the former choice. This effect could also be explained using a Bayesian *self-consistent* observer which defined a prior excluding the directions that were inconsistent with the former choice [14]. Had subjects made the estimation of motion direction independently of the categorical choice, they would have been more accurate at the cost of sometimes being inconsistent (for example, responding CW and then reporting –5o). Accuracy in these decisions was traded down to maintain consistency across the attributes of the *same* stimulus. But what are the limits of this type of consistency? As explained next, revealing experiments in Talluri *et al.* [2] suggest that the brain also seeks for consistency across *different* stimuli.

quicker. This is because those neurons in the visual cortex selective to the attended color will respond more vigorously when presented with an object matching that color [5]. Similarly, neurons in the visual area V5/MT increase (decrease) the gain of their response function when their preferred (anti-preferred) motion direction is attended [6]. To quantify these attention-induced changes in response gain, subjects are cued to direction to which they should attend. But where is the attention directed when experimenters do not cue what to attend? A common implicit assumption is that, if not exogenously cued, attention varies

randomly and uniformly from one trial to the next. Averaging our measurements over many trials would in principle get rid of any systematic bias introduced by selective attention. The new findings of Talluri *et al.* [2] might put this assumption into question (see below).

The relation between choices and attention has been debated for some time [7]. A seminal study [8] supporting a link between perceptual choices and attention found that, during a disparity discrimination task, first, neurons responded more strongly when the subject's choice matched their preferred disparity (a relation termed choice

In the new study, Talluri et al. [2] convincingly characterize a confirmation bias in a novel task in which subjects seek choice consistency across judgments over independent sources of evidence; the authors propose a mechanism, gain modulation via feature-based selective attention, underlying this confirmation bias, and show that the same biasing mechanism could be at play in high-level non-perceptual tasks.

Talluri et al. [2] first carried out a perceptual task in which subjects had to make a CW versus CCW categorical choice on the motion direction $\phi_1$ of Stimulus 1 (Figure 1A center) and then view Stimulus 2 with direction $\phi_2$ and report the average of the two directions (Avg($\phi_1,\phi_2$); Figure 1A, right). Their main finding was that subjects, when estimating the average, gave a higher weight to Stimulus 2 when it was consistent with the first choice than when it was inconsistent. This is in a nutshell the definition of confirmation bias (Figure 1B).

To illustrate this result, we sketched the estimate of Avg($\phi_1,\phi_2$) that subjects reported as a function of $\phi_2$ for the cases in which the first choice was CW or CCW (Figure 1C). The slope of these curves represents the weights of the second stimulus (higher slope for consistent stimuli). Talluri et al. [2] made a clever analysis to uncover the mechanism underlying the choice-induced bias on the perception of $\phi_2$: they computed the separability of the distributions of the reported Avg($\phi_1,\phi_2$) for two different $\phi_2$ (see arrows in Figure 1C). As predicted by a choice-induced gain modulation model, the distributions were more separable in the consistent than in the inconsistent case because the means were further apart and the variances were smaller (Figure 1C insets). Importantly this finding cannot be reproduced by models in which a shift in Avg($\phi_1,\phi_2$) underlies the confirmation bias [13] (Figure 1D). Together the new findings of Talluri et al. [2] provide strong evidence for the presence of a selective gain modulation, independently of the presence or absence of other mechanisms affecting Stimulus 2 (for example, a shift seems at play in some subjects).

Given that feature-based attention can cause an increase in the response gain of those neurons tuned to the attended motion direction, Talluri et al. [2] suggest that the "decision acts like a cue for selective [feature] attention": after a CW choice, attention is directed to CW directions (i.e. $\phi_2 > 0$). The dots in Stimulus 2 moving with a CW direction would then elicit a stronger response in MT neurons and consequently would have a larger net contribution in the estimate than the dots moving with a CCW direction (Figure 1A). Finally, Talluri et al. [2] showed that the same selective overweighting of consistent evidence occurs in a numerical averaging task, suggesting that perceptual decisions and high-level judgments might share similar mechanisms underlying the confirmation bias.

Talluri et al. [2] have made a very important contribution to understand decision biases, but a number of questions remain open. Although their analyses reveal the existence of a choice-induced gain modulation, the relationship between this 'perceptual gain' and the gain of the response function of sensory neurons [6,8,9] is far from being straightforward. Noticeably, the analyses revealed a concomitant decrease of perceptual noise for consistent stimuli (Figure S2A in [2]). This is reminiscent of the decrease in inter-neuronal noise correlations caused by attention, an effect that might dominate attention-induced perceptual improvements [15] and that could explain the increase in accuracy found in congruent trials [16].

Future work using mechanistic models should attempt to bridge the gap between attention-induced changes in neural populations and the perceptual biases found in their data. Moreover, consistency between stimuli 1 and 2 in the design in [2] could result from requiring subjects to average them. But was this a necessary condition to see a choice-induced gain change on Stimulus 2? Future experiments should explore the conditions that define the 'congruence window' and asses whether it could extend beyond one trial. There is growing evidence showing that stimulus perception is attracted towards the stimuli presented in preceding trials [17] (but see [18]) and that this bias can be fitted by a labeled-line model implementing a gain increase of the tuning curves centered at the previously reported stimulus [17], suggestive of feature-based attention [8].

Understanding if sequential effects are related to changes in selective attention will ultimately require linking choice serial dependencies and the simultaneous sequential changes in the activity of sensory neurons, a task which has proven difficult so far [19]. Is this choice-induced selective gain categorical or can it be modulated by, for example, choice confidence [20]? Would subjects behave rationally and reverse the 'sign' of attention if external feedback proved them wrong in their former decision?

These and other questions are waiting to be answered. Original contributions like Talluri et al. [2] are opening new and exciting lines of research that will necessitate imaging and physiology studies, followed by mechanistic modeling, in order to start teasing apart the underlying neural mechanisms. Meanwhile, we will remain attentive to further progress, but seemingly and unfortunately, more attentive to those reports that agree with our current view of the world.

### REFERENCES

1. Nickerson, R.S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. Rev. Gen. Psychol. 2, 175–220.

2. Talluri, B.C., Urai, A.E., Tsetsos, K., Usher, M., and Donner, T.H. (2018). Confirmation bias through selective overweighting of choice-consistent evidence. Curr. Biol. 28, 3128–3135.

3. Carrasco, M. (2011). Visual attention: the past 25 years. Vision Res. 51, 1484–1525.

4. Maunsell, J.H.R., and Treue, S. (2006). Feature-based attention in visual cortex. Trends Neurosci. 29, 317–322.

5. Bichot, N.P., Rossi, A.F., and Desimone, R. (2005). Parallel and serial neural mechanisms for visual search in macaque area V4. Science 308, 529–534.

6. Martinez-Trujillo, J.C., and Treue, S. (2004). Feature-based attention increases the selectivity of population responses in primate visual cortex. Curr. Biol. 14, 744–751.

7. Krug, K. (2004). A common neuronal code for perceptual processes in visual cortex? Comparing choice and attentional correlates in V5/MT. Philos. Trans. R. Soc. Lond. B 359, 929–941.

8. Treue, S., and Martínez Trujillo, J.C. (1999). Feature-based attention influences motion processing gain in macaque visual cortex. Nature *399*, 575–579.

9. Nienborg, H., and Cumming, B.G. (2009). Decision-related activity in sensory neurons reflects more than a neuron's causal effect. Nature *459*, 89–92.

10. Wimmer, K., Compte, A., Roxin, A., Peixoto, D., Renart, A., and de la Rocha, J. (2015). Sensory integration dynamics in a hierarchical network explains choice probabilities in cortical area MT. Nat. Commun. *6*, 6177.

11. Kiani, R., Hanks, T.D., and Shadlen, M.N. (2008). Bounded integration in parietal cortex underlies decisions even when viewing duration is dictated by the environment. J. Neurosci. *28*, 3017–3029.

12. Haefner, R.M., Berkes, P., and Fiser, J. (2016). Perceptual decision-making as probabilistic inference by neural sampling. Neuron *90*, 649–660.

13. Jazayeri, M., and Movshon, J.A. (2007). A new perceptual illusion reveals mechanisms of sensory decoding. Nature *446*, 912–915.

14. Stocker, A.A., and Simoncelli, E.P. (2007). A Bayesian model of conditioned perception. Adv. Neural Inf. Process. Syst. *2007*, 1409–1416.

15. Cohen, M.R., and Maunsell, J.H.R. (2009). Attention improves performance primarily by reducing interneuronal correlations. Nat. Neurosci. *12*, 1594–1600.

16. Bronfman, Z.Z., Brezis, N., Moran, R., Tsetsos, K., Donner, T., and Usher, M. (2015). Decisions reduce sensitivity to subsequent information.

Proc. Biol. Sci. *282*, available at: https://doi.org/10.1098/rspb.2015.0228.

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

18. Fritsche, M., Mostert, P., and de Lange, F.P. (2017). Opposite effects of recent history on perception and decision. Curr. Biol. *27*, 590–595.

19. Lueckmann, J.-M., Macke, J.H., and Nienborg, H. (2018). Can serial dependencies in choices and neural activity explain choice probabilities? J. Neurosci. *38*, 3495–3506.

20. Urai, A.E., Braun, A., and Donner, T.H. (2017). Pupil-linked arousal is driven by decision uncertainty and alters serial choice bias. Nat. Commun. *8*, 14637.

# Nervous System Development: Flies and Worms Converging on Neuron Identity Control

Paschalis Kratsios[1] and Oliver Hobert[2]

[1]Department of Neurobiology, University of Chicago, Chicago, IL 60637, USA
[2]Department of Biology, Howard Hughes Medical Institute, Columbia University, New York, NY 10027, USA
Correspondence: pkratsios@uchicago.edu (P.K.), or38@columbia.edu (O.H.)
https://doi.org/10.1016/j.cub.2018.08.032

**Distinct neuronal cell types display phenotypic similarities such as their neurotransmitter identity. Studies in worms and flies have revealed that this phenotypic convergence can be brought about by distinct transcription factors regulating the same effector genes in different neuron types.**

The phenotypic properties of the very diverse sets of neuronal cell types in an adult nervous system are encoded by the expression of neuron type-specific gene batteries. These 'nuts and bolts' genes, often referred to as terminal effector genes, code for neurotransmitter-synthesizing enzymes or transporters, neuropeptides, ion channels that set the resting potential of a neuron, metabotropic or ionotropic neurotransmitter receptors and many other proteins that define the structural and functional properties of a neuron [1]. What seems at first sight non-intuitive about such terminal identity features is the lack of specificity of each individual component of a given neuron type-specific terminal gene battery. Meaning, with some extremely rare exceptions (mostly sensory receptors that perceive specific sensory modalities), there is generally no such thing as a terminal effector gene that is uniquely expressed in one neuron type, but not in any other neuron type. Consider neurotransmitter identity as an example: Very distinct neuron types utilize the same neurotransmitter system; that is, the same enzymes and transporters dedicated to a specific neurotransmitter system are expressed in multiple different neuron types. If most genes are expressed in several distinct neuron types, how is the astounding phenotypic complexity in the nervous system encoded? As illustrated in Figure 1A, the answer lies in unique combinatorial patterns of gene expression, i.e., the neuron type-specific combinatorial expression of terminal effector genes defines unique phenotypic properties. Through such combinatorial coding, an almost infinite complexity of gene expression profiles and hence neuronal phenotypes can be generated, at least in theory. A recent study in *Cell* by Konstantinides *et al.* [2] defines such combinatorial profiles and further advances our understanding of the molecular mechanisms that control neuron type-specific terminal features, such as neurotransmitter identity.

The question of how the expression of neuron type-specific gene batteries is genetically specified has traditionally been addressed in many different organisms and several different parts of the nervous system in what one could call a 'cell-centric' approach: phenotypic features (both anatomical and molecular) of specific neuron types were studied by researchers interested in specific regions or cell types of the nervous system and