

Bayesian Causal Inference: A Unifying Neuroscience Theory

Ladan Shams

Departments of Psychology, BioEngineering,
and Neuroscience Interdepartmental Program
University of California, Los Angeles

Ulrik Beierholm

Department of Psychology
Durham University

Abstract

Understanding of the brain and the principles governing neural processing requires theories that are parsimonious, can account for a diverse set of phenomena, and can make testable predictions. Here, we review the theory of Bayesian causal inference, which has been tested, refined, and extended in a variety of tasks in humans and other primates by several research groups. Bayesian causal inference is a normative model and has accounted for human behavior in a vast number of tasks including unisensory and multisensory perceptual tasks, sensorimotor, and motor tasks, and has accounted for counter-intuitive findings. The theory has made novel predictions that have been tested and confirmed empirically, and recent studies have started to map the algorithms and neural implementation of the model in the human brain. The parsimony, the diversity of the phenomena that the model has explained, and its illuminating brain function at all three of Marr's levels of analysis make Bayesian causal inference a strong neuroscience theory. The progress made in understanding Bayesian causal inference in the brain has been due to the contributions and collaboration of many labs around the world specializing in different research topics and methods exploring different aspects of brain function. Overall, this highlights the importance of collaborative and multi-disciplinary research for the development of new theories in neuroscience.

1. Introduction

Neuroscience has been one of the fastest growing areas of science in the last two decades. While the proliferation of empirical findings in this area of research has been dizzying, parsimonious and unifying theories that probe the principles of cognitive function have been far and few in between. Here we propose Bayesian causal inference as a parsimonious and unifying theory in cognitive neuroscience and examine its evolution, successes, and limitations in that context.

In mid 2000s, in an attempt to account for auditory-visual perceptual phenomena ranging from integration to segregation a Bayesian model, known as Bayesian Causal Inference (BCI) model ¹, was proposed that involved a competition between two hypotheses, a common cause and independent causes. In the last 15 years, the model has been extended, refined, and adapted to account for a large number of perceptual and sensorimotor phenomena, and vast amounts of behavioral data. More recently the neural underpinnings of BCI have been the subject of extensive research. The model is mathematically equivalent to a few other models that had been proposed in other domains, and the core computation involved in BCI appears to be at work in a number of diverse perceptual and sensorimotor domains in humans and other species ². Bayesian causal inference is a computation that appears to be frequently employed in a variety of cognitive tasks and domains, and appears to have a long-standing evolutionary root². We therefore, henceforth refer to the core computation, which involves competitive priors, as Bayesian Causal Inference theory. We will refer to the Bayesian causal inference model of multisensory perception as a model within that theory which has been most extensively studied, provides an apt example of Bayesian Causal Inference theory, and will be the main focus of this review.

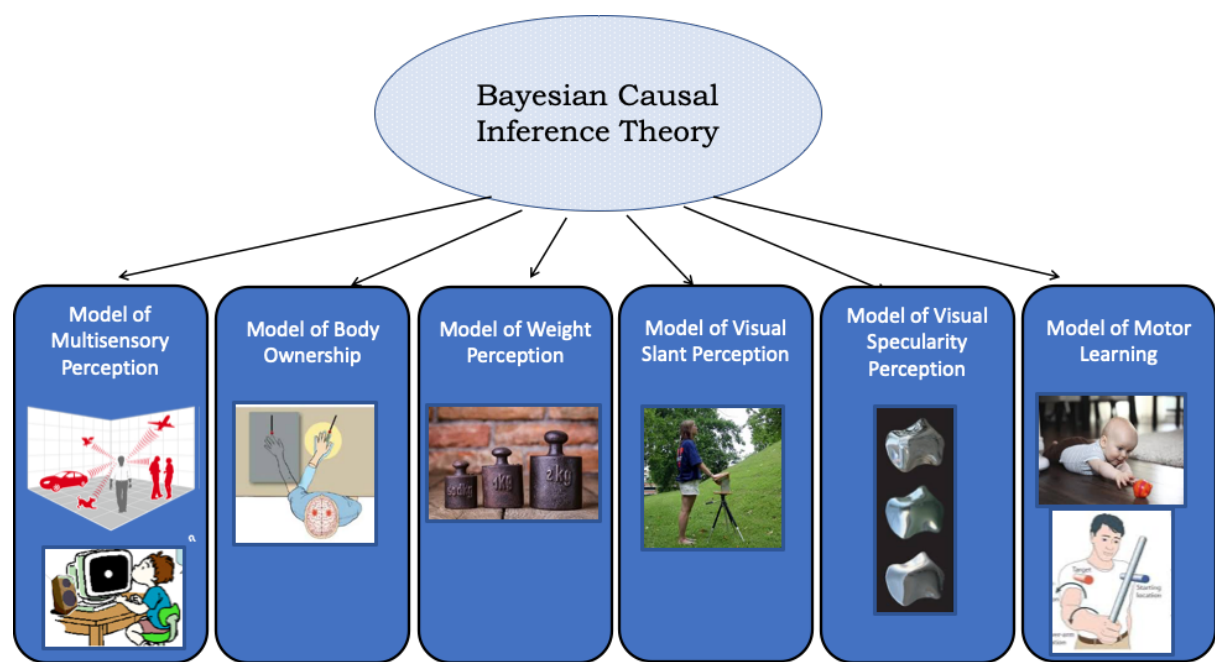


Figure 1. Bayesian Causal Inference theory. Various models employing Bayesian inference and competitive priors on the causal structure have been proposed to account for a variety of tasks. The models are essentially computationally equivalent and fall under Bayesian Causal Inference Theory.

In the following sections, we will review studies of BCI within the framework suggested by David Marr ³, assessing the theory at a computational level, followed by algorithmic/representational level, followed by implementational level. While the initial

studies and the vast majority of studies to date have focused on the computational level of analysis, in the last few years there has been an increasing number of studies that have probed the representations and brain mechanisms of BCI.

2. Computational level of analysis:

Here we will describe BCI using two examples in the perceptual domain, however, the same core computations apply to other tasks and domains of processing as described later.

When faced with sensory stimuli, the nervous system has to estimate the events/sources of sensory inputs and has to overcome two challenges: a) determining the causal structure of the signals, and b) the identity of the source(s) that gave rise to the stimuli. However, these two challenges are intertwined, and cannot be addressed separately and independently.

For example, if we see a talking face and hear speech, the speech perception system has to determine whether the two signals originated from the same source (the auditory and visual signals were produced by the same person), or whether they came from different sources (e.g., the muted video from TV, and the sound of someone speaking in the room). The degree of similarity/discrepancy (in time, space, content) between the auditory and visual signals as well as our prior expectations about the structure of the world affect the interpretation of the causal structure (whether they originate from the same source or not), but likewise the causal structure is important as to whether the signals should be integrated or not (signals with different causes should not be integrated), and thus, to the estimation of perceptual variables (the content of speech).

In this simple example, there are only two causal structures (or hypotheses) that need to be examined, the common cause and independent causes (see Fig. 2a). These two hypotheses can be thought of as two priors (or models of the world) that compete to explain the sensory data. For this reason, BCI is a form of competitive priors model⁴⁻⁶, and mathematically almost equivalent to competitive priors models that had previously been explored in vision science^{4,5,7,8}.

As another example, consider the task of acting on objects in the environment which sometimes requires determining the slant of objects and surfaces. Determining the slant of an oval shape in the visual field requires the visual system to first infer the shape of the object, whether it was an oval object or a circular object that has given rise to the oval projection (see Fig. 2b). The degree of similarity/discrepancy between the binocular disparity cue and shape cue, as well as our prior expectations about the world (how common oval vs. circular objects are) affect the interpretation of the causal structure (oval vs. circle), but likewise the causal structure is important as to whether the binocular disparity and shape signals should be integrated or not (aspect ratio of the oval shape is only informative about slant if the source is a circle). Therefore, here again causal structure (object form) priors/hypotheses compete to

explain the sensory data, and inference based on sensory stimuli leads to the estimation of causal structure, and conversely, the inferred casual structure influences the estimation of perceptual variables (slant).

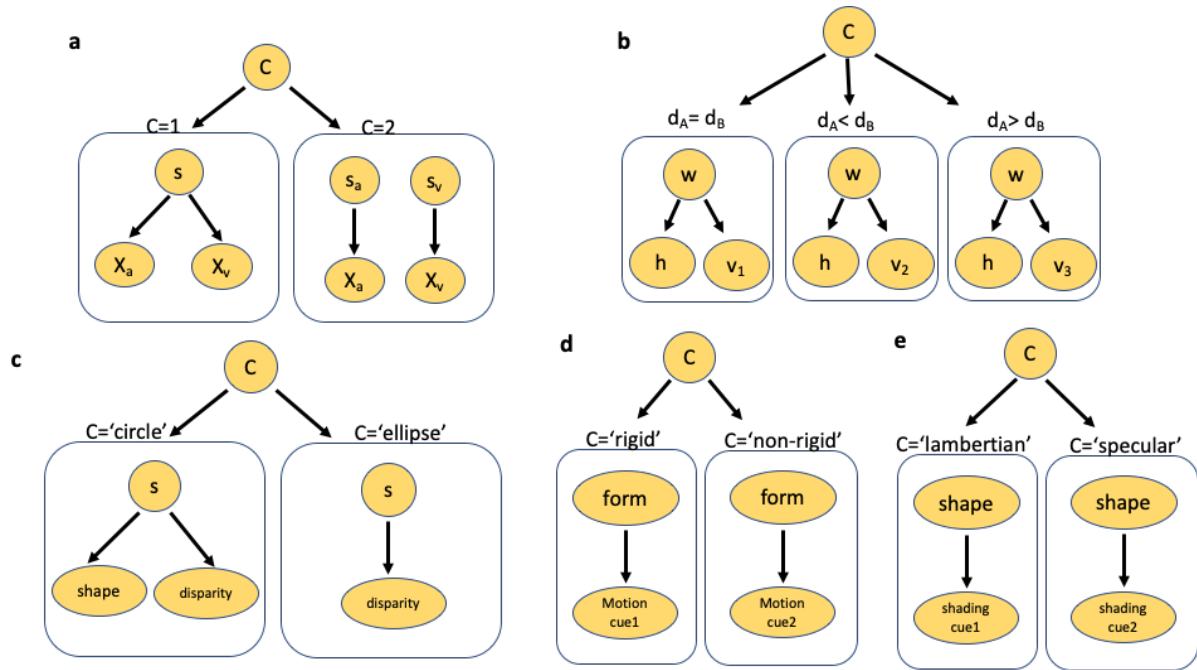


Figure 2. Generative model of five different BCI models. a) The generative model of BCI model of bisensory perception¹. b) The generative model of weight perception in size-weight illusion paradigm⁹. c) The generative model of visual slant perception⁷. d) The generative model of object form perception in the structure-from-motion paradigm⁵. e) The generative model of shape perception in the shape-from-shading paradigm⁴.

BCI is a normative model that addresses both the problem of causal inference (which causal structure/hypothesis generated the stimuli) and the problem of integration/estimation of sensory stimuli, in a unified and coherent fashion using Bayesian inference. The probability of each causal structure (e.g., $C=1$, see Fig. 2a or $C=\text{"circle"}$, see Fig. 2b) is computed based on the prior probability of that causal structure (e.g., the expectation of a common cause, or the expectation of a circle) and the similarity/congruency between the sensory signals, according to Bayes Rule. The estimate of each sensory source (e.g., what was said in the video/audio, or the slant of the object) is computed based on the probability of the causal structures (which determine the extent to which the signals should be integrated) and the reliability of the sensory signals (the more precise/reliable signal would play a more important role), again according to Bayes rule.

Therefore, the core computations that constitute the Bayesian Causal Inference theory can be summarized as follows: a) competition between priors (causal structures) to account for sensory data, b) estimation of hidden perceptual/cognitive variables based on the inferred causal structure, c) computation of both (a) and (b) using Bayes rule.

A natural starting point for understanding human perception is therefore to hypothesize that the human nervous system has evolved a strategy similar to BCI in solving perceptual problems. This hypothesis can be tested by comparing the behavior of BCI with that of human observers in perceptual tasks (Fig. 2, Fig. 3).

Below, we review the studies that have done exactly that, quantitatively or qualitatively comparing human observer data with predictions of BCI in a variety of tasks, and in a variety of sensory conditions in each task. We classify the studies by the nature of the perceptual task they have tackled, and briefly summarize the findings in each section.

2.1 Spatial perception

Spatial localization of objects, as well as auditory-visual interactions in this process as exemplified by the ventriloquist illusion^{10,11}, have been studied extensively (e.g., ^{12–18}). Early computational models of multisensory localization had focused on the limited range of cases in which the senses deviated little from each other, and therefore perceptual cues were always fused into a single unified percept. However, when the auditory and visual signals differ in location substantially there is an absence of integration and human data exhibits a spectrum of phenomena ranging from segregation to integration. The BCI model explains this by the integration of perceptual cues when they are inferred to be causally linked (e.g. when proximal), and segregation of the signals when likely to originate from different sources (e.g. when far apart). Multiple studies of auditory-visual spatial localization have now been performed in which data from observers was compared with BCI^{1,19–23}, and a BCI model with only 4 free parameters could account for the observers' data remarkably well (e.g., accounting for 97% of variance in 250 data points)¹.

In addition to accounting well for the observers' perceived location, BCI also makes predictions about the judgement of unity (common cause). This allows BCI to also account¹ for a phenomenon that had been previously considered counter-intuitive and puzzling, namely, “negative bias”¹⁴, by showing how splitting data based on reports of unity can lead to bias in location responses. Since trials with perceived independent causes tend to be trials in which a larger disparity exists between the encoding of the auditory and visual inputs, limiting the analysis to these trials would lead to an apparent negative bias.

In a more recent study²⁴, observers were asked to report not only the auditory location, but also their judgment of common/independent cause. In accordance with BCI, the more reliable visual stimuli sharpened the window of common cause perception, and increased the visual bias of auditory localization when the discrepancy between the two was not large. Furthermore, in a later study it was found that the reliability of the visual stimuli was itself estimated over time, as modeled through the BCI model²⁵.

While these studies investigated spatial processing along the azimuth, a recent study examined the perception of verticality which relies on visual and vestibular information. The

study performed model comparison and reported that perception of verticality by human observers was best accounted for by the Bayesian Causal Inference model²⁶.

Slightly different variants of BCI used by different research groups or in perceptual tasks that differed from the original localization task used by^{1,19,20,22} showed that BCI could account for observer data in spatial tasks remarkably well^{24,27–29}, for both humans and monkeys.

Subsequently, several studies have used BCI as a quantitative tool to explore the influence of selective attention²³, frame of reference³⁰, aging³¹, adaptation³², and stability and generalization of auditory-visual binding²¹ in human spatial processing, again confirming that BCI captures human spatial processing remarkably well.

2.2 Temporal perception

BCI applies to temporal as well as spatial processing. A simple non-hierarchical* variant of BCI was originally developed to account for behavioral data in a temporal numerosity judgment task³³. This is the task of reporting the number of flashes and beeps presented simultaneously to observers, in which the sound-induced flash illusion often occurs when the number of flashes and beeps are not the same^{34,35}, for example, when a single flash is accompanied with two brief beeps, leading to the percept of two flashes.

The hierarchical model¹(Fig. 2a) was later tested also on the temporal numerosity task and shown to account for the data very well both in bisensory conditions^{21,23,36} and trisensory conditions³⁷ capturing interactions among all combinations of flashes, beeps, and taps on the finger.

In a study of visual-tactile interactions in temporal numerosity judgment where the deviation between number of flashes and taps was limited to one a non-hierarchical approximation to BCI was shown to account well for the observers responses³⁸.

A related task to temporal numerosity judgement is temporal rate discrimination or categorization. A non-hierarchical variant of BCI was shown to account for observers' auditory-visual rate discrimination^{2,39}. In a recent study⁴⁰ using auditory-visual temporal rate categorization, the participants' responses were compared with predictions of three models, including the traditional forced fusion model, and the BCI. BCI could qualitatively and quantitatively account for observers' data and outperformed the other models.

* Non-hierarchical versions do not involve a hierarchy of inference (inference about the structure followed by inference about the sensory variables). Instead of explicit representation and inference of causal structure, they assume a prior on the sensory variables that would capture the two causal scenarios by the mixture of the two priors on the sensory variables:

$$P(X_A, X_V) = N(X_A - X_V, \sigma) * w + \dots N(X_V | 0, \sigma_0) N(X_A | 0, \sigma_0) (1 - w)$$

Although not yet quantitatively studied, given that the sound-induced flash illusion has been reported to occur in other species such as rodents⁴¹, it is likely that BCI also operates in the sensory processing systems of lower mammals.

Auditory-visual temporal processing has also been studied using speech stimuli and temporal asynchrony detection task, showing that in addition to the actual physical time discrepancy between the two signals, the perceptual experience of synchrony depends on the temporal acuity of the observers and the prior expectation of a common cause, and can be accounted for by a BCI model⁴².

Temporal processing in a unisensory setting also requires an inference about the grouping of the stimuli. For example, the perception of the time intervals in an auditory sequence would depend on the inferred causal structure (whether three brief sounds belong to the same event or not). Sawai et al⁴³ demonstrated that a Bayesian causal inference model can account for observers' perceptions of sound sequences.

2.3 Spatio-temporal perception

The need for causal inference is not limited to the situations where information about a given sensory variable is available from multiple cues. There are numerous situations both in unisensory and multisensory settings wherein the nervous system has to consider multiple causal structures (or competitive priors) for determining and estimating the attributes of objects (including one's own body) and events.

For example, interpretation of retinal displacement requires model inference. Yuille, Bulthoff and colleagues were the first to note that the perceptual system has to solve a model inference problem^{4,44}. They proposed a competitive priors model⁵ that accounted for the different interpretation of visual motion cue for structure depending on whether the object is rigid or non-rigid (see Fig. 2c).

More recently, studies of motion perception have shown that human perception can be accounted for by Bayesian models incorporating heavy-tailed priors^{8,45}, which have been shown to be computationally very similar to BCI². The perceptual system has to determine whether two local patches of motion were caused by the same object or different objects in order to determine whether or not to integrate the information across space and apply the same constraints to both patches.

2.3 Weight perception

BCI has been successful in explaining weight perception. A classic example is the size-weight illusion in which when lifting two objects that are identical in shape, mass, and apparent material, but different in size, the smaller object is perceived to be heavier than the larger object^{46,47}. This illusion had evaded theoretical explanation for decades and had been considered to be an "anti-Bayesian" illusion^{48,49}. Indeed a simple non-competitive prior

Bayesian model would predict the opposite of this phenomenon: the smaller object should be perceived as lighter.

Peters et al.⁹ showed that a competitive prior model, equivalent to BCI, in which priors on density relationship between the two objects compete to explain the sensory data, can qualitatively and quantitatively account for the illusion. Furthermore, the model accounts for findings of other studies such as the effect of training with small heavy objects on the illusion. Importantly, the model made a novel prediction that the prior expectations of the density relationship across individuals should correlate with the degree of illusion experienced, subsequently confirmed in experiments⁹.

Similarly in the material-weight illusion, when lifting two objects that have the same size, shape, and mass, the object that appears to be made of low-density material (such as styrofoam) is perceived to be heavier⁵⁰. Peters et al.⁵¹ showed that the same Bayesian model with competing priors on density relationship can also explain this illusion.

2.4 Body ownership perception

Research in the field of body perception over the last two decades has revealed that even the perception of body ownership and attributes are remarkably malleable and involves continuous processing of multisensory information such as visual, proprioceptive, tactile, and vestibular inputs^{52–60}.

In the Rubber-hand illusion⁵², the observer experiences ownership over a fake rubber hand positioned where one's own hand typically would be when the real hand is out of the view and stroked simultaneously (tactile input) with the visible stroking of the rubber hand. This illusion also produces a recalibration of the proprioceptive perception of the real hand.

Samad et al.⁶¹ offered the first computational account of the rubber-hand illusion through a BCI model with two competing causal structures, one in which all the sensory signals are caused by the same object, the observer's hand, and one in which the proprioceptive and tactile signals are caused by the real hand, and the visual signal is caused by the rubber hand. The outcome of causal inference computation depends on the spatial and temporal discrepancy among the signals and the prior probability of a common cause for the sensory signals. The model provided a parsimonious explanation for the known phenomena, but also predicted that the illusion could also be experienced in the absence of tactile input (stroking) confirmed through both subjective reports and skin conductance data. The model also predicted that the Rubber-hand illusion and accompanying proprioceptive drift would not occur if the rubber hand is placed at 30 cm or more from the real hand, consistent with behavioral data⁶². An analogous body ownership phenomenon, the Rubber-foot illusion, can similarly be accounted for by BCI⁶³. A conceptualization based on BCI has been proposed to account for several experimental findings regarding body ownership illusions⁶⁴.

Another phenomenon that is related to the perception of self is the sense of agency, in which one perceives one's own actions to have caused an observed event. A BCI model was shown to qualitatively account for empirical findings related to the sense of agency⁶⁵.

The illusions of body ownership have also been observed in monkeys, and BCI has been shown to account for behavioral responses from monkeys as well as those of human participants⁶⁶. These findings again suggest that the Bayesian causal inference is not unique to the human brain and has a longer evolutionary history.

2.5 Sensori-motor processing:

At any given moment, the visual system has to perform causal inference to determine how many causes exist for any changes in the retinal image, and accordingly, estimate the movement of objects in the scene and one's own eye/body movements (e.g., so that we can navigate accurately and not bump into things). Human postural behavior has been accounted for well by a causal inference model with competing priors of self vs. environment as causes of retinal displacement⁶⁷.

Despite the continual changes in the retinal image caused by frequent saccades, we perceive the world to be stable and generally not moving. Atsma et al.⁶⁸ showed that as quantitatively predicted by BCI, spatial constancy depends on the degree of consistency between pre-saccadic object location memory and post-saccadic visual input, and the two signals are integrated and/or segregated based on the posterior probability that they refer to the same position in the world (common cause).

When the eyes move, the motor system sends a copy of the command sent to the muscles to the sensory system, a signal known as the "efference copy."⁶⁹ When we are passively moved--for example, riding in a moving car--the retinal image changes and there is no efference copy of motor commands available for spatial updating. In a study of passive self motion, Perdreau et al.⁷⁰ found that as predicted by BCI, the perceptual system weighs the integration of the internally updated target position and of the visual feedback by the posterior probability that they correspond to a common position in the world. The data were accounted for by BCI and the account was superior to those of alternative models.

Another important perceptual task for mobile organisms is to determine the direction of heading during movement. Because visual and vestibular information may not always be originating from the same source (body movement), similarly to examples above, the nervous system has to first determine if the two sensory signals have a common source, and if so, to integrate them optimally to best estimate the heading direction and velocity. It has recently been shown that BCI can account for human heading perception⁷¹, as well as heading perception in monkeys^{72,73}.

Recent studies also support the idea that the process of determining the source of an error in any action is governed by Bayesian causal inference. The motor system typically receives feedback from the visual system. When faced with an error (e.g. the deviation from the target in reaching), the motor system has to determine whether the error is due to the motor system or due to other sources (e.g., change in the environment/target, etc.). If the error is due to the motor system, it needs to be corrected, but not otherwise. A study by Wei and Körding⁷⁴ in which the visual feedback of the observed error was manipulated reported a pattern of motor learning/correction consistent with that predicted by a model similar to the Bayesian causal inference. Participants showed no correction if the error was large (and thus the error was not attributed to the motor system; but instead attributed to the experimenter, for example). The largest correction occurred for the largest size error that could still be attributed to the motor system (not too large relative to the variability of the motor system). Motor adaptation has also been examined in a setting that allowed change in both the environment and the motor system (e.g., due to fatigue, etc.) and shown to follow BCI⁷⁵. Participants' behavior in several studies in which participants' movements were perturbed experimentally have been accounted for by BCI as well (see ⁷⁶).

2.7 Other perceptual and sensorimotor tasks

Visual slant perception has been accounted for by Bayesian models utilizing mixed priors^{7,77} or heavy-tailed likelihoods⁷⁸ that have been shown to be equivalent or very similar to BCI².

The oddity detection task is a task that can involve stimuli in one sensory modality or in multiple sensory modalities. Hospedales et al.⁷⁹ showed that a BCI-style model can nicely account for both unisensory and multisensory oddity detection findings that had been previously posed as a challenge to forced fusion maximum likelihood models of sensory integration. The human observer data showed that the detection of oddity depended on the degree of discrepancy between the components in each stimulus, and thus whether or not they are bound together.

BCI has also been employed to account for behavioral data in tasks related to speech perception, such as judgment of asynchrony between auditory and visual speech tokens⁴², the McGurk effect^{80–82}, and the identification of phonetic categories across speakers⁸³.

Research on animal navigation has indicated that many animals, including rats, hamsters, honeybees, and spiders, can exploit multiple sources of information by utilizing path integration (the ability to keep track of distance and direction of path traversed) and landmarks. The pattern of processing of the two cues qualitatively follows Bayesian causal inference, in that when the discrepancy between the two cues is small, a bigger weight is given to the landmark cue than path integration, and when the conflict between the two cues is large, the landmark cue seems to be ignored^{84,85}. This is consistent with a process of causal inference, deciding whether the landmark cue corresponds to the target or to another location².

2.8 Learning, adaptation, recalibration, and selective attention

While psychologists and neuroscientists have studied adaptation, recalibration, perceptual learning, and selective attention for decades, an understanding of how computation and neural processing is affected by these processes has remained elusive. Because Bayesian models, such as BCI, allow quantitative and rigorous characterization of the various components of perceptual and sensorimotor processing in each individual subject, the change in each of these components can be examined quantitatively in various settings and scenarios.

Odegaard et al.²³ investigated the influence of selective attention to visual or auditory stimuli in a spatial localization task and a temporal numerosity judgment task. Surprisingly, they found that selective attention only benefits the sensory modality that is already “good” (or reliable) at the task, and not the sensory modality that is weaker (or less reliable). Specifically, using BCI model fitting they found that in the spatial task, visual precision improves and in the temporal task auditory precision improves as a result of selective attention. The tendency to bind the stimuli (or perceive a common cause) did not seem to be affected by selective attention to a specific sensory modality. A more recent study.⁸⁶ also used BCI in an attempt to characterize the effect of selective attention to visual and tactile modalities on processing components.

The spatial recalibration of auditory map by vision is known as the ventriloquist aftereffect, and happens after repeated exposure to simple visual and auditory stimuli that are presented to the observer at a fixed spatial discrepancy. The same outcome can be due to either a shift in auditory representations of space (likelihood functions), or the prior distribution of the stimuli (priors), or a combination of the two. Wozny et al.³² investigated ventriloquist aftereffect in human observers using BCI and showed that the observer responses were most consistent with a shift in the likelihood functions, therefore, supporting a very low-level neural representation phenomenon.

Using the same experimental paradigm, however, manipulating both the spatial and temporal discrepancy between the visual and auditory stimuli during the adaptation phase, Odegaard et al.⁸⁷ investigated if there are sensory exposures that can modify the priors rather than likelihoods in the spatial localization task. Surprisingly, they discovered that repeated exposure to large auditory-visual spatial discrepancy resulted in an *increase*--instead of a decrease--in the prior expectation of a common cause. Using the BCI framework, they explain why such an adaptation exposure can enhance the tendency to bind, an effect that could potentially be exploited in clinical and educational applications.

The recalibration paradigm described above involves passive exposure to auditory-visual stimuli, no task, and no feedback during the exposure phase. In contrast, in perceptual learning paradigms, the observer actively performs a task during a training phase, and usually receives feedback on the accuracy of their responses. In a study using perceptual learning

paradigm, McGovern et al.⁸⁸ trained observers in an audio-visual simultaneity task, and examined their AV integration in a spatial localization task before and after the training session. They found two effects of training: the window of integration narrowed and there was an overall reduction in integration across the whole range of spatial discrepancies. They showed that a BCI model that included both spatial and temporal variables (an extension of Kording et al.¹ model) could quantitatively and qualitatively account for the findings well, by indicating an increase in temporal precision and a reduction in the tendency to bind subsequent to training.

Different sensory modalities encode space in different frames of reference (e.g. vision in eye-centered, audition in head-centered) and yet our perception of space is unified and coherent and independent of the modality of origin. However, several studies of spatial perception including those discussed above have suggested that the human perceptual system utilizes priors in the estimation of auditory-visual location. Therefore, this begs the question of which frame of reference the prior expectation of space is encoded in. Odegaard et al.³⁰ recently investigated this question using BCI and manipulating the direction of gaze of human observers in an auditory-visual localization task. The results of their quantitative BCI modeling suggested that the frame of reference is a combination of eye-centered and head-centered frames.

In summary, the BCI model has not only accounted for a vast number of perceptual and sensorimotor phenomena, but has also been used to quantitatively characterize and elucidate effects of modulations to processes such as attention, adaptation and learning.

3. Algorithmic level of analysis:

The studies discussed so far have collectively provided compelling evidence that BCI accounts for the computations carried out by the perceptual and sensorimotor systems in the human brain. These studies have shown that human (and other primates) behavior in a vast range of settings and tasks is consistent with the behaviors exhibited by Bayesian causal inference, both qualitatively and quantitatively. Importantly, BCI has been shown to be not overly flexible or powerful and has passed the various tests of parsimony, specificity, model comparison, and model prediction testing (see Box 1). Therefore, all-in-all, the evidence for the Bayesian causal inference as a governing computation carried out by the human nervous system in the wide range of tasks and domains discussed above has proved to be compelling.

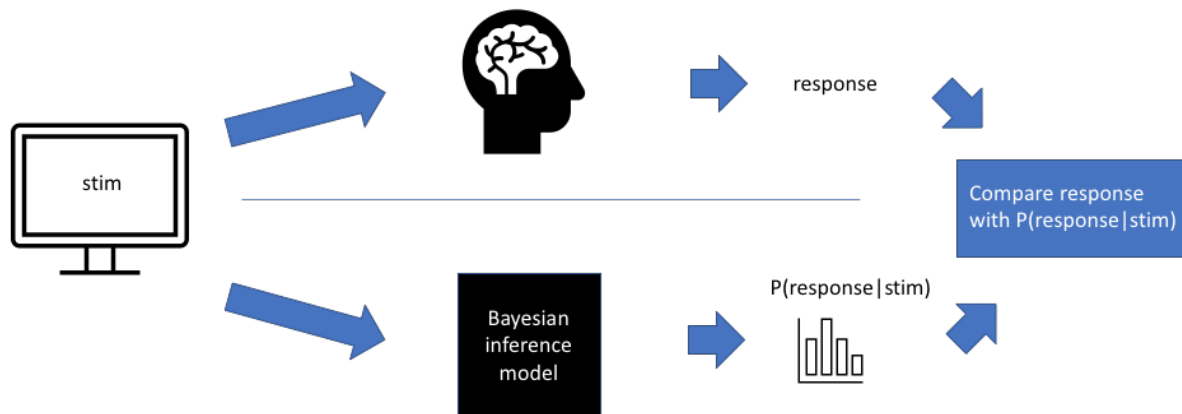


Figure 3. The general approach used to gain insight into the computations used by the nervous system to solve a problem (to perform a given task). This approach has been successful in shedding light on the computations and goals of different brain systems.

The approach used in these studies is to compare the human observers' behavior in a given task and for a given set of stimuli with that of a BCI model, and examine whether the responses qualitatively and/or quantitatively are consistent with each other. Therefore, this approach can be characterized as comparing two black "boxes" with each other (see Figure 3). When providing the same input to the boxes, they produce the same output, suggesting that the two boxes carry out the same computation. While this approach is informative about the computation involved, it does not shed light on how the computation is carried out. For example, if we feed number 2, 5, and 7 to the two black boxes and they both output 4, 10, and 14, respectively, we can infer that the computation that is carried out by the boxes is multiplication by 2, but this computation can be carried out in a lot of different ways, both at an algorithmic level, and at an implementation level. For example, one box may carry out the computation by using a look-up table, whereas the other one may carry it out by applying the multiplication by 2 operation, or by adding the input to itself. Also, one box may use a calculator, whereas the other box may use an abacus.

A recent example of how this approach can be used to examine the algorithms used to perform inference examined the role of perceptual uncertainty²⁵. A key aspect of Bayesian inference is the estimation of the uncertainty of each variable. The estimation could conceivably be done instantaneously by the nervous system for each sense, but instead seems to be estimated over several seconds, thus providing a smoother, less noise-prone estimation.

In order to shed light on the representations or algorithm used by the human nervous system in computing BCI, Beierholm et al.¹⁹ asked whether priors and likelihoods are represented independently in human auditory-visual spatial processing. To approach this question, the visual stimulus contrast was manipulated to lead to two different levels of reliability/precision, and hence two different likelihood functions. The study examined whether the estimate of the priors changed as a result of change in the stimulus precision (and in turn, change in likelihoods). The results suggested that priors remained the same despite a substantial change in likelihoods. Therefore, this study provided evidence against

representations and algorithms that would rely on lookup tables, and provided support for the idea that the nervous system achieves the optimal Bayesian causal inference by encoding likelihoods and prior distributions independently of each other and then combining them according to Bayes Rule (Fig. 3a).

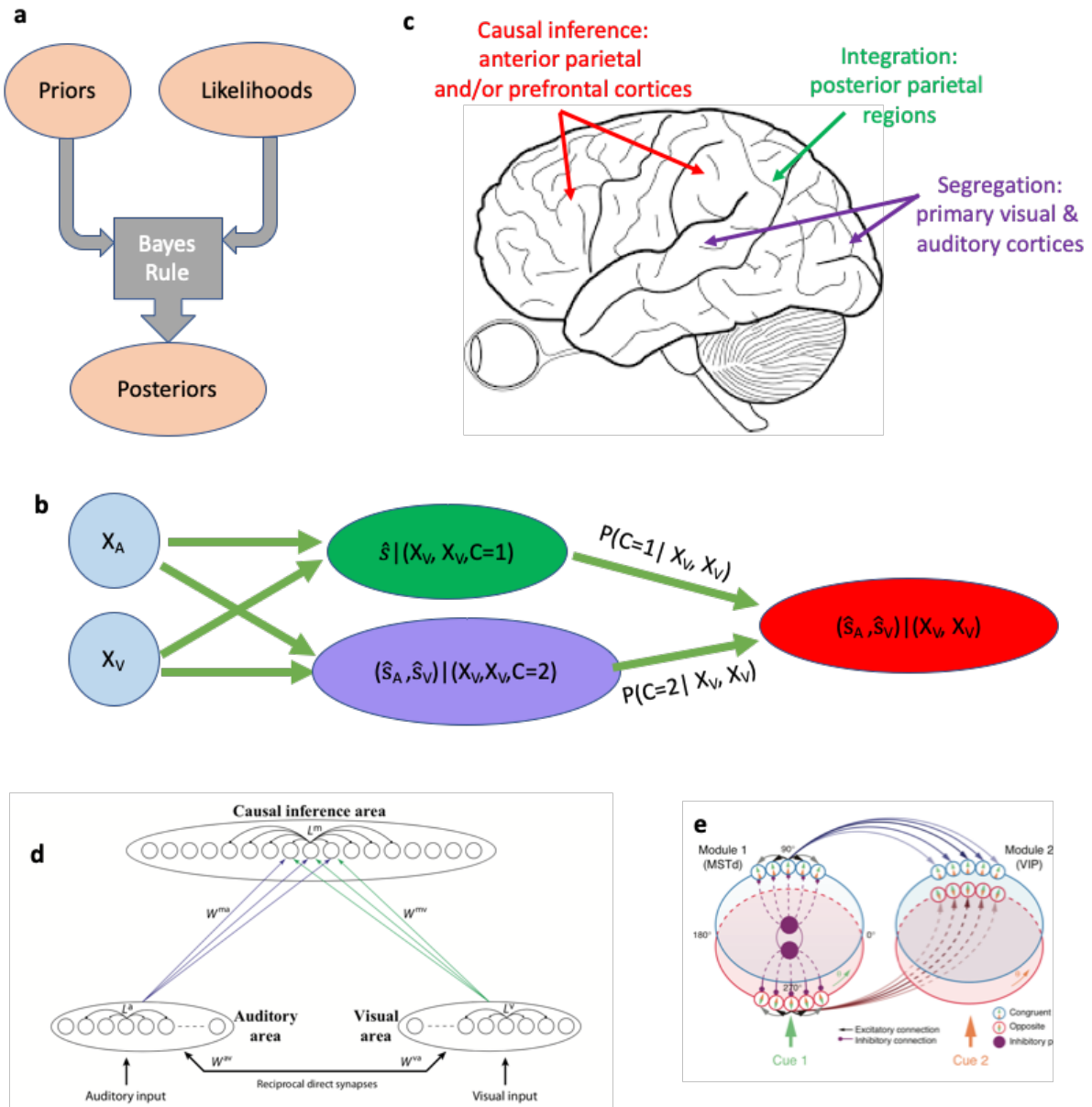


Figure 4. a) One algorithm for carrying out Bayesian inference in the brain. b) An algorithm for carrying out BCI computations. c) Possible neural architecture of BCI in sensory tasks in human brain^{40,89,90}. d) A proposed neural circuitry for BCI in auditory-visual spatial tasks, reproduced from⁹¹. e) A proposed neural circuitry for BCI in a vestibular visual heading task, reproduced from⁹².

More recently, multiple studies using a variety of neuroimaging techniques have probed the representations and processing of BCI in human observers. Despite the significant variety in tasks used, and the methods of study, all of these studies have suggested that the human brain

carries out BCI in a sequential fashion (Fig. 3b), by first representing the unisensory estimates (e.g., X_A and X_V), followed by computing and representing the reliability-weighted fusion of the unisensory estimates, followed by estimates of causal structure, followed by the combination of the fusion and segregation estimates according to the probability of their respective causal structure to produce the BCI estimates of the variable of interest (time, space, etc.)^{40,89,90,93}. Therefore, these studies confirmed what Beierholm et al.'s¹⁹ study had suggested which is BCI is not only a good computational model of perception, but also a good process model⁹⁴.

4. Implementational level of analysis:

The studies discussed above shed light on the representations and processes involved in BCI. The next important question is how this algorithm is carried out by the machinery of the brain, networks of neurons.

Some of the studies discussed above, in addition to shedding light on the representations and algorithm used by the human nervous system, have also examined the specific brain areas involved in each of these processes. These studies have used physiological measurements of brain activity together with the predictions of the BCI model to gain insight into brain areas involved in encoding the different components of BCI computation.

For example, using fMRI and a spatial localization task, Rohe et al.⁸⁹ investigated the role of different brain areas along the human visual pathway in representing various relevant variables and computations of Bayesian causal inference. They found a hierarchical and sequential evolution of computation starting with the representation of unisensory auditory and visual estimates of location in primary visual and auditory cortical areas, followed by the reliability weighted integration of the two signals in the anterior parietal region, followed by the final combination of location estimate from the two causal scenarios in the posterior parietal region. A similar approach⁹⁰, however using the temporal numerosity judgment task (counting number of flashes and beeps) and EEG measurements and analyzing the temporal dynamics, showed the same style of processing. Sensory cortical areas appeared to produce unisensory estimates of the number of flashes and beeps, followed by higher-up representation of multisensory estimates of numerosity under the common cause vs. independent causes assumption, followed by the final BCI estimate. The data also suggested that pre-stimulus oscillatory alpha power and phase may be the mechanism for encoding the prior bias for a common cause. Another study using EEG to investigate temporal dynamics of BCI, but utilizing a spatial localization task, reported the same hierarchical processing of information as found in the temporal numerosity task⁹³.

Furthermore, a study probing the neural mechanisms of BCI by investigating the temporal dynamics using MEG employed a somewhat different temporal task (temporal rate categorization) and yet reported the same pattern of hierarchical evolution of the computation

from sensory cortical areas advancing to higher areas of processing including parietal regions and prefrontal regions of the brain⁴⁰. While this study also showed parietal regions to be involved in processing multisensory estimates, the arbitration between the common cause vs. independent cause hypotheses, and the combination of the two estimates appeared to take place in the prefrontal cortex.

A recent study sought to illuminate the implementation of BCI at the level of individual neurons and populations of neurons within a brain area in the context of body ownership⁶⁶. The activity of neurons in monkey premotor cortex was recorded in unisensory and proprioceptive-visual conditions in a behavioral paradigm analogous to the Rubber-hand illusion. The study found neuronal activities (both at individual neuron level and population level) correlated with various components of BCI, namely segregation, integration of the two modalities, and posterior probability of a common cause.

These studies provide compelling evidence that the human and primate nervous systems implement BCI for a variety of tasks across a variety of sensori-motor pathways. However, how exactly these brain areas and neurons accomplish these computations remains unclear. Previous work had shown how theoretically neurons could perform cue integration through Probabilistic Population Coding^{95,96} (PPC), while segregation relies only on potential parameter transformation. It has also been proposed that the dynamic properties of single neurons, single synapses, and sensory receptive fields in effect carry out the non-linear computations involved in causal inference⁹⁷. The challenging question that had not been probed until recently was how the posterior over the common cause would be computed by networks of neurons, and how to use this to non-linearly combine the integrated and segregated estimates. In the last several years, suggestions have been made to address these questions.

Ma & Rahmati⁹⁸ used the vocabulary provided by PPC to build a 'neural circuit' that would be able to perform the exact computation required for estimating the posterior probability of the common cause. While they were able to devise such a circuit, they acknowledged that the biological plausibility of this method was questionable due to the complexity of the circuit.

Yamashita and colleagues⁹⁹ developed a network, consisting of a single layer with lateral connections, that performed an implicit calculation of the probability of a single cause and was able to replicate the observed behavior. However, as a consequence of just using a single layer, access to unisensory information would be lost.

In contrast, Cuppini, Ursino and collaborators created a neural network based on two parallel layers representing two unisensory modalities as well as a crossmodal layer receiving inputs from the uni-sensory layers^{91,100,101}. Although not explicitly encoding probabilities, the connectivity within, and across layers, produced behavior very similar to the non-linear cue combination of causal inference, with cues being pulled together when close together, but unaffected when further apart.

Neurophysiological studies of heading direction in monkeys have reported neurons that appear to encode full integration of the signals and neurons that appear to be tuned to opposite directions (known as “opposite neurons”). Inspired by these findings Zhang et al.^{92,102} proposed a model consisting of integration neurons and opposite neurons that can estimate the probability of a common cause, and account for behavioral data on heading direction. A limitation to this theory is however that it only applies to circular variables (e.g. direction) as it relies on the properties of neurons with opposite direction tuning.

Using an alternative approach Yu and colleagues¹⁰⁰ used importance sampling as a proposed method of neural computations, an approximation that has previously been shown to be biologically feasible¹⁰¹. By combining importance sampling with principles from PPC they were able to recreate the posterior probability of common cause in the BCI model. An intriguing possibility explored in the paper is to what degree this method can be used to generalize to multiple stimuli (>2), a computationally difficult problem (see Box 1). In a later paper¹⁰³ these ideas were extended to allow the same neural circuit to perform both the causal inference estimation, as well as the integration of cues.

These models have used simple biological mechanisms and connectivity patterns (e.g., long lateral inhibition) that are known to exist in sensory cortices of primate brains, to either explicitly or implicitly encode the probabilities required for the algorithms in BCI. Overall, while there are still uncertainties about how the brain implements a causal inference strategy, there are several proposals of ways it could happen. Future experimental work, in collaboration with more modelling, is needed to narrow down these hypotheses.

5. Discussion

5.1 The emergence of BCI as a neuroscience theory

Up until early 2000s, the predominant model of multisensory processing and cue combination was maximum likelihood estimation that implicitly assumed the sensory signals stem from the same source, and therefore, should completely get fused in the estimation of the source. Research on the BCI model started by making the observation that sensory signals don't always get fused^{14,35}; sometimes they get fused (full integration leading to illusions if there is a discrepancy), sometimes they are segregated, and sometimes they interact partially³³. It was shown that a normative Bayesian model that allowed both integration and segregation could account for these observations^{1,33} and also explain counter-intuitive phenomena¹ such as negative bias¹⁴. Later, the same model was shown to account for other multisensory phenomena, while making novel and unexpected predictions in different tasks which in turn were also empirically tested and confirmed^{51,61}.

The BCI model involves competition amongst hypotheses, aka, priors. For this reason, this model can also be referred to as a “competitive prior” model. The BCI model that was proposed in the context of multisensory perception is mathematically equivalent to the

competitive prior model that had been proposed earlier in the context of visual perceptual tasks^{5,7}. As discussed in section 2, the BCI model or variants of it have successfully accounted for human (and other mammals) behavior in a number of different tasks ranging in processing domain (spatial, temporal, spatio-temporal, speech, etc.), ranging in combination of sensory modalities (unisensory visual, unisensory auditory, multisensory with various combinations of sensory modalities), perception of the world as well as perception of self (body perception and ownership), sensorimotor, and motor processing. It appears that evolution discovered this powerful computational mechanism and employed it in a variety of processes in the nervous system.

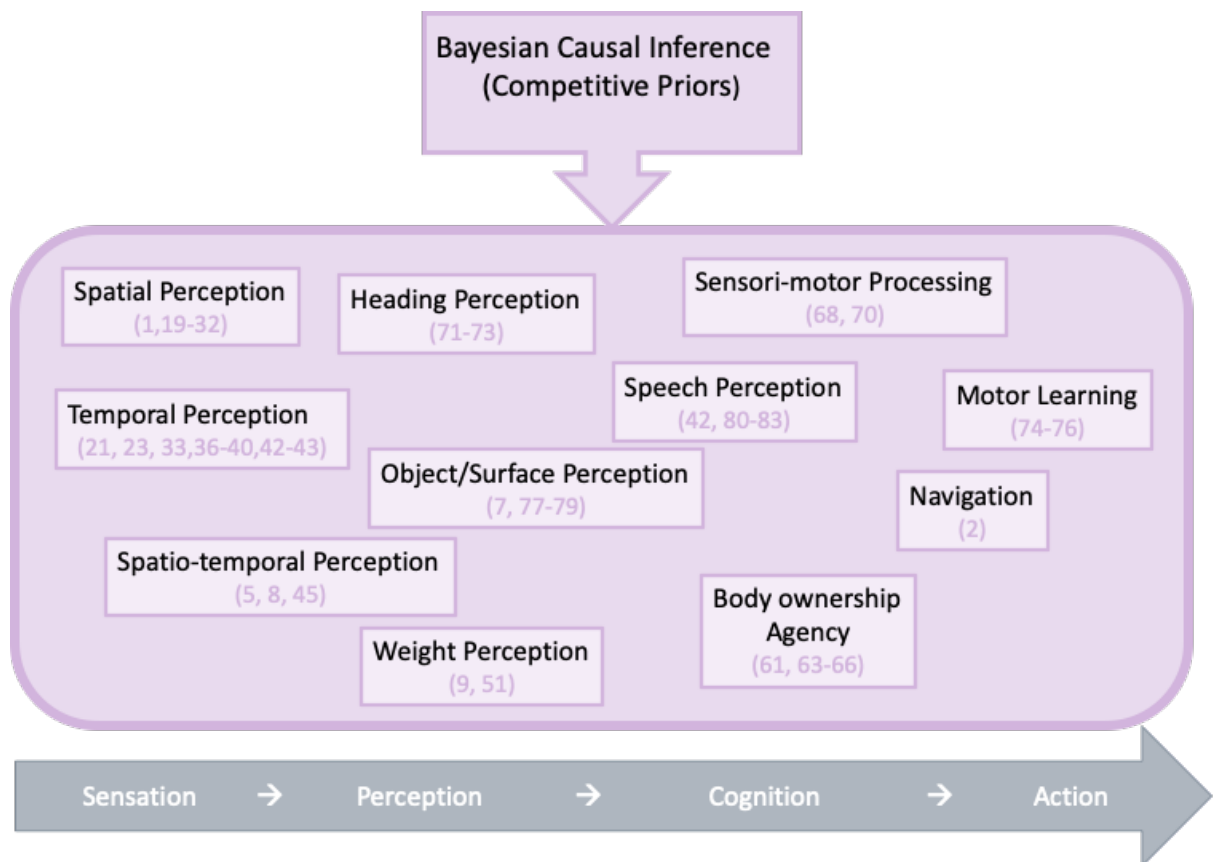


Figure 5. Bayesian Causal Inference is a unifying theory. Each box represents a domain of brain processing and the numbers in each box are references to the studies that have shown BCI accounts for a task in that domain. BCI accounts for empirical findings in a large number of studies probing diverse tasks ranging from sensation to action, and a vast number of phenomena that appear completely unrelated, or appear counter-intuitive.

5.2 The validity of BCI as a neuroscience theory

A strong neuroscience theory should be able to account for existing data, and make predictions that can be tested and verified. It should be simple, yet able to explain a diversity of phenomena¹⁰⁴. And lastly, it should be verifiable at all three levels of analysis as proposed by Marr³.

BCI is a simple and normative model (see Box 1) that has accounted for a wide variety of phenomena (see section 2). It is a parsimonious, unifying theory of perceptual and sensorimotor processing, governing the realms of multisensory perception, unisensory perception, body ownership perception, as well as sensorimotor processing (Fig 5).

The research summarized in section 2 is the computational level of analysis that was carried out in a distributed fashion by several research groups around the globe. Once BCI was established as a successful computational model, research groups started exploring the algorithmic/representational and implementation aspects of it by investigating neuronal architectures and circuitries that can carry out the computation in the nervous system¹⁰⁴.

As summarized in sections 3 and 4, a variety of recent studies employing sophisticated neuroscientific methods, for example, EEG, fMRI, MEG, and single-electrode recordings, and using machine learning methods of analysis, have reported a hierarchical architecture in the human (and primate) cortex that appears to implement BCI in multiple processing domains. Moreover, several studies by different research groups have proposed specific and biologically plausible neural circuits that could carry out BCI in the context of spatial or temporal processing tasks. Thus, BCI has been tested on all three of Marr's levels, emerging as a strong neuroscience theory.

We argue that this hierarchical and systematic approach to understanding brain function that is based on a sequence of observation of behavior, developing theoretical ideas, testing predictions, exploring generality, and investigating algorithms and potential neural implementation provides a template for a successful approach to understanding the brain function. It also demonstrates the effectiveness of the scientific method^{104,105} and scientific collaboration, and the power of multi-disciplinary approaches to brain research.

Box 1:

How meaningful is a good fit to the data?

Some have criticized Bayesian models of cognition by questioning their flexibility due to the potential use of ad hoc priors or a large number of free parameters. It is indeed possible for a model to be excessively flexible/powerful and account for any set of data and thus, not shed light on the true underlying computations at work. This warrants examination of the nature, structure, and flexibility of the model.

- **BCI is not an ad hoc model that was concocted to fit the data.** It is based on a generative model that reflects the causal structure of events in the environment (or the body) leading to the stimuli that are observed by the nervous system. BCI is a normative model and represents the optimal way of solving the problems of causal inference and source estimation, problems that the nervous system has to solve in a large number of processing domains from early sensory stages all the way to cognitive and motor stages.
- **BCI accounts for data with few free parameters.** BCI models typically have a small number of free parameters (usually 3-4) accounting for large sets of data (usually more than 500 data points). Moreover, fitting the parameters using a subset of the data (not used in the subsequent parameter-free model prediction testing)^{33,37} or showing that the model fails to account for scrambled data³³ provide additional evidence that the model is not overly flexible, and its account is selective to the observer data.
- **BCI models can account for patterns of behavior even with no free parameters.** BCI makes qualitative predictions about the pattern of behavior in a range of stimulus conditions that have been repeatedly confirmed by empirical data in a variety of tasks, populations, and processing domains. This includes complete fusion and a large bias by the more reliable/precise cue when the disparity is small, complete segregation of the cues when the discrepancy is large, or 'partial integration' (assuming model averaging) when the discrepancy is moderate. No other model had been able to account for this pattern of behavior.
- **BCI models outperform alternative models.** In some studies (e.g., ^{1,26,78}) BCI was compared with other models of comparable complexity (same or larger number of free parameters) and was shown to provide a better account for the data.
- **BCI's predictions--including some counter-intuitive predictions-- have been empirically confirmed^{51,61}.** This is generally considered the ultimate test of a theory, and so far BCI has successfully passed this test.

These facts and findings collectively provide compelling support for BCI as computation governing the perceptual and sensorimotor processing as discussed in section 2.

Box 2:

Limitations and outstanding questions:

Computational level of analysis

- **How are BCI computations carried out in natural environments?** The nervous system typically is faced with a large number of sensory (or sensorimotor) signals at any given moment. These situations would require the nervous system to choose from a large number of causal structures, and hence the problem of combinatorial explosion. As the number of parameters increases faster than exponential, approximations are needed when multiple cues are present³⁷ or more clever generative models are also possible¹⁰⁶ that allow for the number of parameters to grow with the stimulus size using non-parametric Bayesian inference. It is also possible that instead of computing exact probabilities, the nervous system can overcome the computational intractability by instead computing approximations of the probabilities¹⁰⁷. Are there strategies, constraints, heuristics, or approximations utilized by the nervous system to make the probabilistic inference involved in BCI more computationally efficient and tractable?
- **Is there a difference in the way high-level priors vs. low-level priors influence the computation?** Our current knowledge of the priors involved in BCI is limited. It is clear that biases can influence the inference in the various tasks explored at multiple levels of processing. For example, in a spatial localization task, the prior expectation that the visual and auditory signals have a common cause plays an important role in the perception of location, however, this prior bias can be induced by high-level cognitive factors such as experimenter's instructions, as well as low-level biases that are due to the statistics of the environment and likely encoded at early sensory stages of processing.
- **What is the nature of the priors?** While some studies have started investigating the properties of Bayesian priors^{19,21,30}, there are still some basic questions about the nature, plasticity, and other characteristics of priors that are not explored and understood. For example, in a multisensory setting, are there different priors for unisensory vs. multisensory conditions? In a similar vein, in situations where sensory information is available from multiple senses, does BCI in each sensory modality (e.g., the various visual cues for depth, or object identification) take place prior to inference across the senses (e.g., auditory and visual depth cues, or cues to object identity), or do they occur in parallel?

- **What determines the loss function?** There is some evidence that at least in some basic spatial and temporal tasks the loss function is not uniform across individuals, and may not be uniform even within an individual across time^{20,21}. What are the factors that determine the loss function for any given task, individual, situation?

Box 3:

Limitations and outstanding questions:

Representational and implementation levels of analysis

- **To what extent the inferred causal structure accessible to consciousness?** While the probability of each causal structure is computed in BCI and contributes to the estimation of the variables of interest (e.g., spatial location, time, speech, etc.), it is not clear whether this inference is accessible to consciousness, and whether the nervous system commits to a given causal structure at any stage of processing. If we do not explicitly ask participants to report their perceived causal structure (e.g., common cause or independent causes), would the nervous system bother with an explicit/conscious encoding of these probabilities?
- **Does the nervous system encode probability distributions or some summary statistics of the distributions?** For example, it is possible that the system only encodes the mean (or max) and/or variance of the probability distributions. For a complicated probability distribution it becomes impossible for the brain to explicitly encode it. In such cases does the brain rely on other approximations, such as sampling?¹⁰⁷
- **Is BCI hardwired or learned by the human brain?** It is possible that it is hardwired for some basic tasks and learned for others. If it is learned, what are the mechanisms for learning, at what age do they emerge, and how fast are they learned?
- **What are the neural mechanisms of BCI in the human brain?** Very little is known about the implementation of Bayesian inference in the brain in general. It is not clear whether and how probability distributions are represented in the nervous system, how prior knowledge is combined with likelihood functions (or sensory input). The studies using fMRI, EEG, MEG and single electrode recordings discussed in sections 3 and 4 have shed light on the architecture and brain areas where some of the BCI computations may take place, however how the computations are carried out at the level of circuitry and network of neurons is still unclear. The neural network models discussed in section 4 provide a proof of concept that such implementation is possible, however which mechanism is employed by the biological neural networks remains a topic of future research.

Interestingly, gaining insight into the neural mechanisms of BCI may shed light on the computational strategies of BCI for realistic complex situations with several sensory inputs, and the problem of combinatorial explosion discussed above. It is likely that evolution has found a trick to make this computationally intractable task feasible, and investigating the neural implementations of BCI may reveal simple approximations or heuristics that make this problem computationally tractable.

Acknowledgements

We thank Uta Noppeney and Megan Peters for their helpful comments on the first draft of the manuscript. We thank Leon Shams for his assistance with the figures.

References:

1. Körding, K. P. *et al.* Causal inference in multisensory perception. *Plos One* **2**, e943 (2007).
2. Shams, L. & Beierholm, U. R. Causal inference in perception. *Trends Cogn. Sci.* **14**, 1–8 (2010).
3. Marr, D. *Vision*. (W. H. Freeman, 1982).
4. Yuille, A. L. & Bulthoff, H. H. Bayesian decision theory and psychophysics. in *Perception as Bayesian Inference* (eds. Knill, D. C. & Richards, W.) (Cambridge University Press, 1996).
5. Yuille, A. L. & Clark, J. J. Bayesian models, deformable templates and competitive priors. in *Proceedings of the 1991 York conference on Spatial vision in humans and robots* 333–347 (Cambridge University Press, 1994).
6. Colas, F., Diard, J. & Bessi re, P. Common Bayesian Models for Common Cognitive Issues. *Acta Biotheor.* **58**, 191–216 (2010).
7. Knill, D. C. Robust cue integration: A Bayesian model and evidence from cue-con ict studies with stereoscopic and gure cues to slant. *J. Vis.* **7**, 1–24 (2007).
8. Stocker, A. A. & Simoncelli, E. P. Noise characteristics and prior expectations in human visual speed perception. *Nat. Neurosci.* **9**, 578–585 (2006).
9. Peters, M. A. K., Ma, W. J. & Shams, L. The Size-Weight Illusion is not anti-Bayesian after all: a unifying Bayesian account. *PeerJ* **4**, e2124 (2016).
10. Thurlow, W. R. & Jack, C. E. Certain determinants of the ventriloquism effect. *Percept Mot Ski.* **36**, 1171–1184 (1973).
11. Warren, D. H., Welch, R. B. & McCarthy, T. J. The role of visual-auditory ‘compellingness’ in the ventriloquism effect: implications for transitivity among the spatial senses. *Percept Psychophys* **30**, 557–564 (1981).
12. Alais, D. & Burr, D. The Ventriloquist Effect Results from Near-Optimal Bimodal

- Integration. *Curr. Biol.* **14**, 257–262 (2004).
13. Slutsky, D. A. & Recanzone, G. H. Temporal and spatial dependency of the ventriloquism effect. *Neuroreport* **12**, 7–10 (2001).
 14. Wallace, M. T. *et al.* Unifying multisensory signals across time and space. *Exp. Brain Res.* **158**, (2004).
 15. Jack, C. E. & Thurlow, W. R. Effects of degree of visual association and angle of displacement on the ‘ventriloquism’ effect. *Percept Mot Ski.* **37**, 967–979 (1973).
 16. Recanzone, G. H. Auditory influences on visual temporal rate perception. *J Neurophysiol* **89**, 1078–1093 (2003).
 17. Choe, C. S., Welch, R. B., Gilford, R. M. & Juola, J. F. The ‘ventriloquist effect’: visual dominance or response bias. *Percept Psychophys* **18**, 55–60 (1975).
 18. Bertelson, P., Vroomen, J., De Gelder, B. & Driver, J. The ventriloquist effect does not depend on the direction of deliberate visual attention. *Percept. Psychophys.* **62**, 321–332 (2000).
 19. Beierholm, U., Quartz, S. & Shams, L. Bayesian priors are encoded independently from likelihoods in human multisensory perception. *J. Vis.* **9**, 1–9 (2009).
 20. Wozny, D. R., Beierholm, U. R. & Shams, L. Probability Matching as a Computational Strategy Used in Perception. *PLoS Comput. Biol.* **6**, e1000871 (2010).
 21. Odegaard, B. & Shams, L. The Brain’s Tendency to Bind Audiovisual Signals Is Stable but Not General. *Psychol. Sci.* **27**, 583–591 (2016).
 22. Odegaard, B., Wozny, D. R. & Shams, L. Biases in Visual, Auditory, and Audiovisual Perception of Space. *PLOS Comput. Biol.* **11**, e1004649 (2015).
 23. Odegaard, B., Wozny, D. R. & Shams, L. The effects of selective and divided attention on sensory precision and integration. *Neurosci. Lett.* **614**, 24–28 (2016).
 24. Rohe, T. & Noppeney, U. Sensory reliability shapes perceptual inference via two mechanisms. *J. Vis.* **15**, 22 (2015).
 25. Beierholm, U., Rohe, T., Ferrari, A., Stegle, O. & Noppeney, U. Using the past to estimate sensory uncertainty. *eLife* **9**, e54172 (2020).

26. de Winkel, K. N., Katliar, M., Diers, D. & Bülthoff, H. H. Causal Inference in the Perception of Verticality. *Sci. Rep.* **8**, 5483 (2018).
27. Sato, Y., Toyoizumi, T. & Aihara, K. Bayesian Inference Explains Perception of Unity and Ventriloquism Aftereffect: Identification of Common Sources of Audiovisual Stimuli. *Neural Comput.* **19**, 3335–3355 (2007).
28. Hospedales, T. M. & Vijayakumar, S. Structure Inference for Bayesian Multisensory Scene Understanding. *IEEE Trans. Pattern Anal. Mach. Intell.* **30**, 2140–2157 (2008).
29. Mohl, J. T., Pearson, J. M. & Groh, J. M. Monkeys and humans implement causal inference to simultaneously localize auditory and visual stimuli. *J. Neurophysiol.* **124**, 715–727 (2020).
30. Odegaard, B., Beierholm, U. R., Carpenter, J. & Shams, L. Prior expectation of objects in space is dependent on the direction of gaze. *Cognition* **182**, (2019).
31. Jones, S. A., Beierholm, U., Meijer, D. & Noppeney, U. Older adults sacrifice response speed to preserve multisensory integration performance. *Neurobiol. Aging* **84**, 148–157 (2019).
32. Wozny, D. R. & Shams, L. Computational Characterization of Visually Induced Auditory Spatial Adaptation. *Front. Integr. Neurosci.* **5**, (2011).
33. Shams, L., Ma, W. J. W. J. & Beierholm, U. Sound-induced flash illusion as an optimal percept. *Neuroreport* **16**, 1923–7 (2005).
34. Shams, L., Kamitani, Y. & Shimojo, S. What you see is what you hear. *Nature* **408**, 788–788 (2000).
35. Shams, L., Kamitani, Y. & Shimojo, S. Visual illusion induced by sound. *Cogn. Brain Res.* **14**, 147–152 (2002).
36. Beierholm, U. R. Bayesian Modeling of Sensory Cue Combinations. (California Institute of Technology, 2007). doi:10.7907/K89B-XW75.
37. Wozny, D. R., Beierholm, U. R. & Shams, L. Human trimodal perception follows optimal statistical inference. *J. Vis.* **8**, 24 (2008).
38. Bresciani, J.-P. P., Dammeier, F. & Ernst, M. O. Vision and touch are automatically

- integrated for the perception of sequences of events. *J Vis* **6**, 554–564 (2006).
39. Roach, N. W., Heron, J. & McGraw, P. V. Resolving multisensory conflict: a strategy for balancing the costs and benefits of audio-visual integration. *Proc. R. Soc. B Biol. Sci.* **273**, 2159–2168 (2006).
 40. Cao, Y., Summerfield, C., Park, H., Giordano, B. L. & Kayser, C. Causal Inference in the Multisensory Brain. *Neuron* **102**, 1076-1087.e8 (2019).
 41. Ito, Y. *et al.* Auditory-induced visual illusions in rodents measured by spontaneous behavioural response. *Sci. Rep.* **9**, 19211 (2019).
 42. Magnotti, J. F., Ma, W. J. & Beauchamp, M. S. Causal inference of asynchronous audiovisual speech. *Front. Psychol.* **4**, (2013).
 43. Sawai, K., Sato, Y. & Aihara, K. Auditory Time-Interval Perception as Causal Inference on Sound Sources. *Front. Psychol.* **3**, (2012).
 44. Clark, J. J. & Yuille, A. L. *Data fusion for sensory information processing systems*. (Boston: Kluwer Academic Publishers, 1990).
 45. Lu, H., Lin, T., Lee, A., Vese, L. & Yuille, A. Recovering the functional form of the slow-and-smooth prior in global motion perception. *J. Vis.* **10**, 819–819 (2010).
 46. Charpentier, A. Analyse experimentale: De quelques elements de la sensation de poids [Experimental study of some aspects of weight perception]. *Arch Physiol Norm Pathol* **3**, 122–135 (1891).
 47. Koseleff, P. Studies in the perception of heaviness. I.1.2: Some relevant facts concerning the size-weight-effect (SWE). *Acta Psychol. (Amst.)* **13**, 242–252 (1957).
 48. Ernst, M. O. Perceptual Learning: Inverting the Size–Weight Illusion. *Curr. Biol.* **19**, R23–R25 (2009).
 49. Braynov, J. B. & Smith, M. A. Bayesian and “Anti-Bayesian” Biases in Sensory Integration for Action and Perception in the Size–Weight Illusion. *J. Neurophysiol.* **103**, 1518–1531 (2010).
 50. Harshfield, S. P. & DeHardt, D. C. Weight judgment as a function of apparent density of objects. *Psychon. Sci.* **20**, 365–366 (1970).

51. Peters, M. A. K., Zhang, L.-Q. & Shams, L. The material-weight illusion is a Bayes-optimal percept under competing density priors. *PeerJ* **6**, e5760 (2018).
52. Botvinick, M. & Cohen, J. Rubber hands 'feel' touch that eyes see. *Nature* **391**, 756–756 (1998).
53. Petkova, V. I. & Ehrsson, H. H. If I Were You: Perceptual Illusion of Body Swapping. *PLOS ONE* **3**, e3832 (2008).
54. Ehrsson, H. H. The experimental induction of out-of-body experiences. *Science* **317**, 1048 (2007).
55. Lenggenhager, B., Tadi, T., Metzinger, T. & Blanke, O. Video ergo sum: manipulating bodily self-consciousness. *Science* **317**, 1096–1099 (2007).
56. Hoort, B. van der, Guterstam, A. & Ehrsson, H. H. Being Barbie: The Size of One's Own Body Determines the Perceived Size of the World. *PLOS ONE* **6**, e20195 (2011).
57. Tsakiris, M. & Haggard, P. The rubber hand illusion revisited: visuotactile integration and self-attribution. *J. Exp. Psychol. Hum. Percept. Perform.* **31**, 80–91 (2005).
58. Blanke, O., Ortigue, S., Landis, T. & Seeck, M. Stimulating illusory own-body perceptions. *Nature* **419**, 269–270 (2002).
59. Blanke, O. Multisensory brain mechanisms of bodily self-consciousness. *Nat. Rev. Neurosci.* **13**, 556–571 (2012).
60. Ehrsson, H. H., Spence, C. & Passingham, R. E. That's my hand! Activity in premotor cortex reflects feeling of ownership of a limb. *Science* **305**, 875–877 (2004).
61. Samad, M., Chung, A. J. & Shams, L. Perception of Body Ownership Is Driven by Bayesian Sensory Inference. *PLOS ONE* **10**, e0117178 (2015).
62. Lloyd, D. M. Spatial limits on referred touch to an alien limb may reflect boundaries of visuo-tactile peripersonal space surrounding the hand. *Brain Cogn.* **64**, 104–109 (2007).
63. Schürmann, T., Vogt, J., Christ, O. & Beckerle, P. The Bayesian causal inference model benefits from an informed prior to predict proprioceptive drift in the rubber foot illusion. *Cogn. Process.* **20**, 447–457 (2019).
64. Kiltner, K., Maselli, A., Kording, K. P. & Slater, M. Over my fake body: body

- ownership illusions for studying the multisensory basis of own-body perception. *Front. Hum. Neurosci.* **9**, (2015).
65. Legaspi, R. & Toyoizumi, T. A Bayesian psychophysics model of sense of agency. *Nat. Commun.* **10**, 4250 (2019).
 66. Fang, W. *et al.* Statistical inference of body representation in the macaque brain. *Proc. Natl. Acad. Sci.* **116**, 20151–20157 (2019).
 67. Dokka, K., Kenyon, R. V., Keshner, E. A. & Kording, K. P. Self versus Environment Motion in Postural Control. *PLoS Comput. Biol.* **6**, e1000680 (2010).
 68. Atsma, J., Maij, F., Koppen, M., Irwin, D. E. & Medendorp, W. P. Causal Inference for Spatial Constancy across Saccades. *PLOS Comput. Biol.* **12**, e1004766 (2016).
 69. Jeannerod, M. Action Monitoring and Forward Control of Movements. in *The Handbook of Brain Theory and Neural Networks* 83–85 (MIT Press Cambridge, MA, 2003).
 70. Perdreau, F., Cooke, J. R. H., Koppen, M. & Medendorp, W. P. Causal inference for spatial constancy across whole body motion. *J. Neurophysiol.* **121**, 269–284 (2018).
 71. de Winkel, K. N., Katliar, M. & Bühlhoff, H. H. Causal Inference in Multisensory Heading Estimation. *PLOS ONE* **12**, e0169676 (2017).
 72. Acerbi, L., Dokka, K., Angelaki, D. E. & Ma, W. J. Bayesian comparison of explicit and implicit causal inference strategies in multisensory heading perception. *PLOS Comput. Biol.* **14**, e1006110 (2018).
 73. Dokka, K., Park, H., Jansen, M., DeAngelis, G. C. & Angelaki, D. E. Causal inference accounts for heading perception in the presence of object motion. *Proc. Natl. Acad. Sci.* **116**, 9060–9065 (2019).
 74. Wei, K. & Körding, K. Relevance of error: what drives motor adaptation? *J. Neurophysiol.* **101**, 655–64 (2009).
 75. Berniker, M. & Kording, K. Estimating the sources of motor errors for adaptation and generalization. *Nat. Neurosci.* **11**, 1454–1461 (2008).
 76. Wei, K. & Kording, K. Causal Inference in Sensorimotor Learning and Control. *Sens.*

Cue Integr. (2012) doi:10.1093/acprof:oso/9780195387247.003.0002.

77. Knill, D. C. C. Mixture models and the probabilistic structure of depth cues. *Vis. Res* **43**, 831–854 (2003).
78. Girshick, A. R. & Banks, M. S. Probabilistic combination of slant information: Weighted averaging and robustness as optimal percepts. *J. Vis.* **9**, 8–8 (2009).
79. Hospedales, T. & Vijayakumar, S. Multisensory Oddity Detection as Bayesian Inference. *PLoS ONE* **4**, e4205 (2009).
80. Magnotti, J. F. & Beauchamp, M. S. A Causal Inference Model Explains Perception of the McGurk Effect and Other Incongruent Audiovisual Speech. *PLOS Comput. Biol.* **13**, e1005229 (2017).
81. Magnotti, J. F. *et al.* A causal inference explanation for enhancement of multisensory integration by co-articulation. *Sci. Rep.* **8**, 18032 (2018).
82. Magnotti, J. F., Dzeda, K. B., Wegner-Clemens, K., Rennig, J. & Beauchamp, M. S. Weak observer-level correlation and strong stimulus-level correlation between the McGurk effect and audiovisual speech-in-noise: A causal inference explanation. *Cortex* **133**, 371–383 (2020).
83. Kleinschmidt, D. F. & Jaeger, T. F. Robust speech perception: Recognize the familiar, generalize to the similar, and adapt to the novel. *Psychol. Rev.* **122**, 148–203 (2015).
84. Shettleworth, S. J. & Sutton, J. E. Multiple systems for spatial learning: dead reckoning and beacon homing in rats. *J. Exp. Psychol. Anim. Behav. Process.* **31**, 125–141 (2005).
85. Etienne, A. S., Teroni, E., Hurni, C. & Portenier, V. The effect of a single light cue on homing behaviour of the golden hamster. *Anim. Behav.* **39**, 17–41 (1990).
86. Badde, S., Navarro, K. T. & Landy, M. S. Modality-specific attention attenuates visual-tactile integration and recalibration effects by reducing prior expectations of a common source for vision and touch. *Cognition* **197**, 104170 (2020).
87. Odegaard, B., Wozny, D. R. & Shams, L. A simple and efficient method to enhance

- audiovisual binding tendencies. *PeerJ* **5**, e3143 (2017).
88. McGovern, D. P., Roudaia, E., Newell, F. N. & Roach, N. W. Perceptual learning shapes multisensory causal inference via two distinct mechanisms. *Sci. Rep.* **6**, 24673 (2016).
 89. Rohe, T. & Noppeney, U. Cortical Hierarchies Perform Bayesian Causal Inference in Multisensory Perception. *PLOS Biol.* **13**, e1002073 (2015).
 90. Rohe, T., Ehrlis, A.-C. & Noppeney, U. The neural dynamics of hierarchical Bayesian causal inference in multisensory perception. *Nat. Commun.* **10**, 1907 (2019).
 91. Cuppini, C., Shams, L., Magosso, E. & Ursino, M. A biologically inspired neurocomputational model for audiovisual integration and causal inference. *Eur. J. Neurosci.* **46**, 2481–2498 (2017).
 92. Zhang, W.-H. *et al.* Complementary congruent and opposite neurons achieve concurrent multisensory integration and segregation. *eLife* **8**, e43753 (2019).
 93. Aller, M. & Noppeney, U. To integrate or not to integrate: Temporal dynamics of hierarchical Bayesian causal inference. *PLOS Biol.* **17**, e3000210 (2019).
 94. Maloney, L. T. & Mamassian, P. Bayesian decision theory as a model of human visual perception: Testing Bayesian transfer. *Vis. Neurosci.* **26**, 147–155 (2009).
 95. Beck, J. M. *et al.* Probabilistic Population Codes for Bayesian Decision Making. *Neuron* **60**, 1142–1152 (2008).
 96. Ma, W. J., Beck, J. M., Latham, P. E. & Pouget, A. Bayesian inference with probabilistic population codes. *Nat. Neurosci.* **9**, 1432–1438 (2006).
 97. Lochmann, T. & Deneve, S. Neural processing as causal inference. *Curr. Opin. Neurobiol.* **21**, 774–781 (2011).
 98. Ma, W. J. & Rahmati, M. Towards a Neural Implementation of Causal Inference in Cue Combination. *Multisensory Res.* **26**, 159–176 (2013).
 99. Yamashita, I., Katahira, K., Igarashi, Y., Okanoya, K. & Okada, M. Recurrent network for multisensory integration-identification of common sources of audiovisual stimuli. *Front. Comput. Neurosci.* **7**, (2013).

100. Yu, Z., Chen, F., Dong, J. & Dai, Q. Sampling-based causal inference in cue combination and its neural implementation. *Neurocomputing* **175**, 155–165 (2016).
101. Shi, L. & Griffiths, T. Neural implementation of hierarchical Bayesian inference by importance sampling. *Adv. Neural Inf. Process. Syst.* **22** 1669–1677 (2009).
102. Zhang, W., Wu, S., Doiron, B. & Lee, T. S. A Normative Theory for Causal Inference and Bayes Factor Computation in Neural Circuits. *Adv. Neural Inf. Process. Syst.* **32**, 3804–3813 (2019).
103. Fang, Y., Yu, Z., Liu, J. K. & Chen, F. A unified neural circuit of causal inference and multisensory integration. *Neurocomputing* **358**, 355–368 (2019).
104. Gorini, R. Al-Haytham the man of experience: First Steps in the Science of Vision. *J Inter Soc Hist. Islam. Med.* **2**, 53–55 (2003).
105. Toomer, G. J. Review of Ibn al-Haythams Weg zur Physik. *Isis J. Hist. Sci. Soc.* **55**, 463–465 (1964).
106. Yates, T., Larigaldie, N. & Beierholm, U. A non-parametric Bayesian prior for causal inference of auditory streaming. in *Proceedings of the 39th Annual Conference of the Cognitive Science Society* 1381–1386 (2017).
107. Sanborn, A. N. & Chater, N. Bayesian Brains without Probabilities. *Trends Cogn. Sci.* **20**, 883–893 (2016).