

A Bayesian model of distance perception from ocular convergence

Article

Published Version

Creative Commons: Attribution 4.0 (CC-BY)

Open Access

Scarfe, P. ORCID: <https://orcid.org/0000-0002-3587-6198> and Hibbard, P. B. (2025) A Bayesian model of distance perception from ocular convergence. PLoS Computational Biology, 21 (10). e1013506. ISSN 1553-734X doi: 10.1371/journal.pcbi.1013506 Available at <https://centaur.reading.ac.uk/125136/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

To link to this article DOI: <http://dx.doi.org/10.1371/journal.pcbi.1013506>

Publisher: Public Library of Science

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

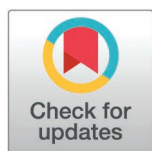
RESEARCH ARTICLE

A Bayesian model of distance perception from ocular convergence

Peter Scarfe^{1*}, Paul B. Hibbard²

1 Vision and Haptics Laboratory, School of Psychology and Clinical Language Sciences, University of Reading, Reading, United Kingdom, **2** Department of Psychology, University of Stirling, Stirling, Scotland, United Kingdom

* p.scarfe@reading.ac.uk



OPEN ACCESS

Citation: Scarfe P, Hibbard PB (2025) A Bayesian model of distance perception from ocular convergence. PLoS Comput Biol 21(10): e1013506. <https://doi.org/10.1371/journal.pcbi.1013506>

Editor: Christoph Strauch, Utrecht University: Universiteit Utrecht, NETHERLANDS, KINGDOM OF THE

Received: January 17, 2025

Accepted: September 9, 2025

Published: October 3, 2025

Copyright: © 2025 Scarfe, Hibbard. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data availability statement: All code written in support of this publication is publicly available at <https://github.com/scarfelab/BayesianVergence>.

Funding: PS was initially funded by an EPSRC PhD studentship. The funders had no role in study design, data collection and analysis,

Abstract

Ocular convergence is one of the critical cues from which to estimate the absolute distance to objects in the world, because unlike most other distance cues a one-to-one mapping exists between absolute distance and ocular convergence. However, even when accurately converging their eyes on an object, humans tend to underestimate its distance, particularly for more distant objects. This systematic bias in distance perception has yet to be explained and questions the utility of vergence as an absolute distance cue. Here we present a probabilistic geometric model that shows how distance underestimation can be explained by the visual system estimating the most likely distance in the world to have caused an accurate, but noisy, ocular convergence signal. Furthermore, we find that the noise in the vergence signal needed to account for human distance underestimation is comparable to that experimentally measured. Critically, our results depend on the formulation of a likelihood function that takes account of the generative function relating distance to ocular convergence.

Author summary

Since at least the time of Descartes the ocular convergence state of the eyes has been considered one of the most important cues to distance in the world. However, despite accurate fixation, humans underestimate distance from ocular convergence, particularly for more distant objects. This questions the utility of one of the primary cues from which distance could be estimated. Here we present a rigorous probabilistic analysis of how humans might estimate distance from ocular convergence. This analysis shows that the experimentally observed distance underestimation can be explained by observers estimating the most likely distance in the world to have caused the accurate but uncertain measurement of the convergence state of the eyes. The level of uncertainty needed to account for observed distance underestimation is consistent with that experimentally measured.

decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Introduction

Estimating distance from ocular convergence

To interact successfully with our environment, we need to infer distal properties of the world, such as distance, depth and shape, from proximal sources of sensory information. Proximal sensory data is generally considered to consist of quasi-independent sources of information termed ‘cues’, each of which gives rise to largely statistically separable perceptual estimates of a given distal world property [1,2]. For the visual modality, this involves making inferences from both retinal and extra-retinal cues [3]. Retinal cues are available from the two-dimensional images of the world projected to the back of each eye and include cues such as perspective [4], texture [5,6], shading [7] and binocular disparity [8]. Extra-retinal cues are not contained within these images and include cues such as the physical orientation of the eyes in their orbits [9,10], and information about the movement of the eyes [11,12] and body [13] over time.

Since at least the time of Descartes, the extra-retinal cue of ocular convergence (or more succinctly vergence) has been considered a particularly important cue to absolute distance [8]. Absolute distance is defined as the Euclidean distance between the cyclopean eye and a 3D coordinate in the world. In the present paper to be succinct we will use the term “distance” interchangeably with “absolute distance”. Where needed we will draw clear distinctions, e.g., between absolute and relative distance. Vergence is defined as the horizontal angular difference between the physical orientations of the two eyes in their orbits [14] and if the eyes are converged on a point in the world, for a particular version angle, there is a geometric one-to-one mapping between vergence and the absolute distance to that point. This is absent for virtually all other visual cues [8,15–17].

The geometry of vergence is shown in Fig 1a. Following standard convention, we have defined the interocular distance to be the distance between the centre of rotation of each eye and considered the eyes to be rigid, circular and to rotate about their centres (see S1 Text for discussion). With these assumptions, D_F is the absolute distance of the fixated F and D_R is the radial distance to that same point. I is the observer’s interocular distance and h half this value. The directions of gaze for the left and right eyes relative to straight ahead are β_L and β_R , where $\beta_L = \theta_L$ and $\beta_R = \theta_R$, defining right angled triangles given by

$$\theta_L = \tan^{-1} \left(\frac{h + B}{D_F} \right) \quad (1)$$

and

$$\theta_R = \tan^{-1} \left(\frac{h - B}{D_F} \right) \quad (2)$$

The vergence angle, θ_V , is given by

$$\theta_V = \theta_L + \theta_R \quad (3)$$

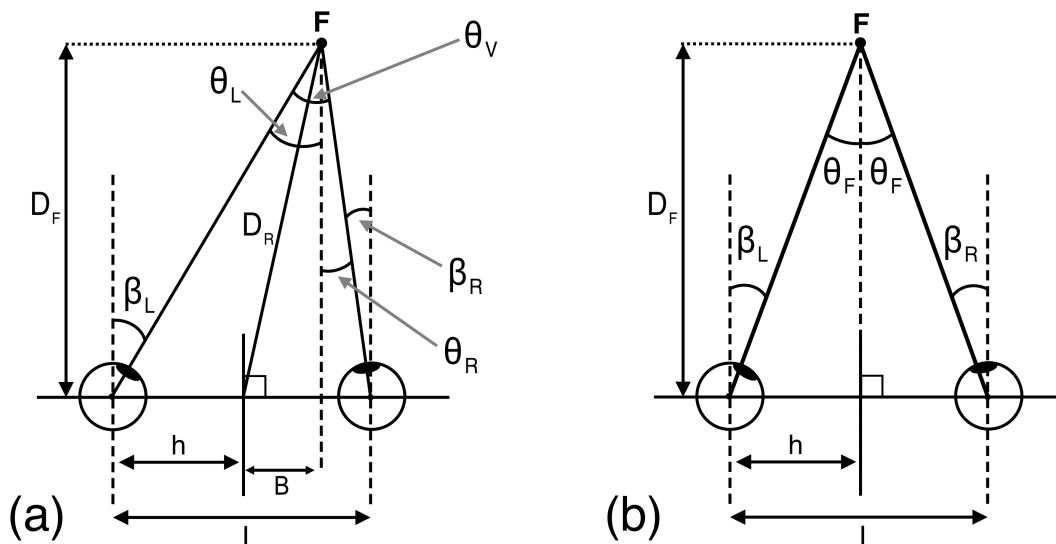


Fig 1. Diagram showing the geometry of estimating absolute distance from ocular convergence. (a) shows the geometry for the general case of a point of fixation off the median plane and (b) for the case in which the observer fixates a point on the median plane. See main text for details.

<https://doi.org/10.1371/journal.pcbi.1013506.g001>

Distance estimation from vergence has primarily been studied experimentally when observers are estimating the distance of objects placed directly in front of them on either (a) the median plane (Fig 1b) [10] or (b) directly in front of one eye [9]. Under both circumstances the geometry of vergence is simplified to a right-angled triangle with a base of length h or l respectively. In the case of a point on the median plane, $D_F = D_R$ and $B = 0$, and the half vergence angle for each eye, θ_F , is given by

$$\theta_F = \tan^{-1} \left(\frac{h}{D_F} \right) \quad (4)$$

Equation 4 defines the mathematical relationship relating the distal property of distance in the world (D_F) to the measured proximal vergence signal (θ_F). By rearranging to make D_F the subject, one can see that estimation of distance from vergence requires knowledge of the convergence state of the eyes and the interocular separation

$$D_F = \frac{h}{\tan(\theta_F)} \quad (5)$$

Equation 5 “inverts” the relationship between distance and vergence, geometrically allowing distance in the world to be estimated from vergence [9,18]. The strategy of inverting a generative model relating proximal cues to distal properties is termed “inverse optics” [19] and is central to contemporary models of perception, which formulate perception as a process of statistical inference [20–24]. Vergence is thought to primarily be of utility in judging near distances as its magnitude drops off rapidly with distance (Fig 2a).

Therefore, if a person can accurately estimate θ_F and h , and has an appropriate inverse generative model, they could accurately estimate distance, D_F . In the following we will use the “hat” symbol to represent the sensory system’s *estimate* of a property, be that distal or proximal. Thus, $\hat{\theta}_F$ represents the sensory system’s *estimate* of the *proximal* signal θ_F , and $\hat{D}_F$ the sensory system’s *estimate* of the *distal* signal D_F . As such, if a person does not have accurate information about θ_F

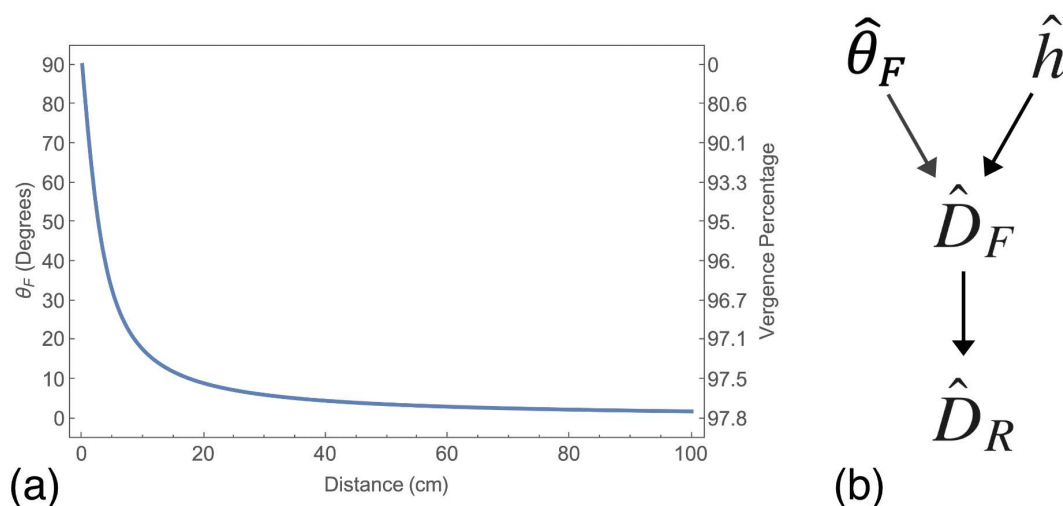


Fig 2. Vergence as a cue to absolute distance. (a) Plot shows the magnitude of the half vergence signal θ_F over distance for a person with a 6.5 cm interocular distance. This is also plotted as a percentage of the total vergence range. (b) Shows the steps in estimating distance from vergence. $\hat{\theta}_F$ represents the observers estimate of the vergence state θ_F and $\hat{h}$ their estimate of the half interocular separation h . With these estimates they can make a distance estimate $\hat{D}_F$. A final step relates this perceptual estimate to the report of this estimate as $\hat{D}_R$, e.g., by manual estimation or verbal report.

<https://doi.org/10.1371/journal.pcbi.1013506.g002>

and/or h ($\hat{\theta}_F \neq \theta_F$ and/or $\hat{h} \neq h$) they will misestimate distance, i.e., $\hat{D}_F \neq D_F$. Under these circumstances their perception would be said to be “non-veridical”, “biased” [25–28] or lacking “external accuracy” [29].

Given an internal estimate of distance $\hat{D}_F$, a further function f_R transforms this estimate into a perceptual report $f_R(\hat{D}_F) = \hat{D}_R$ (Fig 2b). If this is not an identity function a person will accurately estimate distance but incorrectly report that accurate estimate. This “response bias” could have both “perceptual” and “cognitive” components. An example of a “perceptual” component would be an incorrect mapping between the perceptual estimate and motor commands needed for a manual estimate. An example of a “cognitive” component would be a bias to report values toward the centre of the experimental stimulus range [30].

Whilst acknowledged to exist, response bias is a controversial, complicated topic typically assumed to be absent in models of statistical inference. The reason for this is that in the extreme it is currently an intractable perceptual and philosophical issue to establish if a person has correctly estimated a distal world property, but then incorrectly reported that correct estimate [31]. In this regard a “response bias” could effectively be an unprovable redescription of the estimation bias one is wanting to explain. What we are interested in here is modelling how observers might estimate distance from ocular convergence, not how they might estimate distance but then respond in a way inconsistent with their estimate.

Estimating distance from ocular convergence: Experimental evidence

Despite over 150 years of research, there is still active debate as to whether observers can accurately estimate absolute distance from vergence (see [16] for a comprehensive review). Within the literature, the primary ways to measure perceived distance from vergence have been (1) verbal distance estimates, (2) manual distance estimates, (3) relative distance judgements and (4) judgements about other properties such as depth, motion or shape. Verbal distance estimates tend to be used less frequently and there is evidence that they can be more variable across observers (e.g., [32]), so we will primarily describe other response modes here as they are reflective of the literature.

In an early study, Swenson [33] had observers manually estimate the distance of a disc light source, the angular size and intensity of which were kept constant, and found highly accurate judgements within the 25–40 cm range. Gogel and Tietz [34] investigated the perception of further distances with an illusory motion parallax task (where incorrect distance estimation would result in a stationary light appearing to move either with or against lateral head movements) and concluded that observers overestimate near distances and underestimate far. They described this as being consistent with observers' estimates of distance contracting toward an intermediate default value, a phenomenon termed the 'specific distance tendency' [35].

Later studies using manual distance estimates found largely accurate estimates in near visual space (less than approximately 60 cm) and suggested that any residual misestimation of distance from vergence could be due to a cognitive 'contraction bias', whereby people report perceptual estimates contracted to the centre of an available experimental range [9,18,30,36,37]. This is an example where it is suggested that distance might be accurately estimated but inaccurately reported. A similar process has been suggested to distort the results of studies seeking to estimate the relative importance of binocular cues in the control of grasping [38].

In a widely cited study, Viguier et al. [10] used behavioural measures and eye tracking to examine distance perception from vergence. Behavioural measures included manual distance matching, manual half- and double- distance setting and verbal report. All measures showed progressive underestimation of distance as physical distance increased. Given the additional steps in half and double distance setting we will focus on manual distance matching.

Here, when asking observers to manually set the distance of a point light cursor to match that of another previously seen target (leaving 5 seconds between presentations to eliminate any disparity cue between the cursor and target), Viguier et al. found that observers accurately estimated near distances, but progressively underestimated further distances, despite them correctly fixated the very same targets (as shown through measuring the physical orientation of the eyes).

All of the above studies used physical light sources, avoiding the conflicting cues that are typically present when using computer generated stimuli [39]. However, the issue of conflicting cues is still relevant. To "isolate" the vergence cue many studies hold cues such as retinal size and object luminance constant. For example, in Viguier et al. [10] the target had a fixed retinal size and luminance, whereas the matched point light cursor's retinal size and luminance varied consistently with vergence-specified distance. Cue conflicts could therefore influence measured distance misestimation from vergence. This is an issue that affects virtually all studies of sensory cues (see [40] for a detailed exposition of problems inherent in isolating sensory cues).

Recently Linton [16] has suggested that all published studies claiming to show distance estimation from vergence are in fact the result of other uncontrolled cues to distance. Aiming to experimentally isolate vergence fully, Linton found that observers were insensitive to slow changes in vergence, and distance estimates did not track vergence when it was changed in this way. Puzzlingly, many observers made consistently non-random distance estimates suggesting that they were using some source of distance information or strategy and not simply guessing on each trial. Linton (p. 3187) acknowledges that "there are several alternative interpretations of these results that we cannot conclusively reject" but considered these all less plausible than observers not being able to judge distance from vergence.

Indirectly inferring distance perception from ocular convergence: Experimental evidence

In addition to having observers make direct distance estimates, tasks where depth and shape are estimated from retinal images have been widely used to infer vergence specified object distance and thus indirectly assess the accuracy of distance estimation from vergence (e.g., [41,42]). Consider an observer fixating at a distance D_F in the median plane, with another point on the median plane is placed at $D_P = D_F - d$ (Fig 3) and the observer is asked to perceptually judge the depth d .

The half retinal disparity produced by this viewing geometry is given by

$$\eta = \tan^{-1} \left(\frac{dh}{-dD_F + D_F^2 + h^2} \right) \quad (6)$$

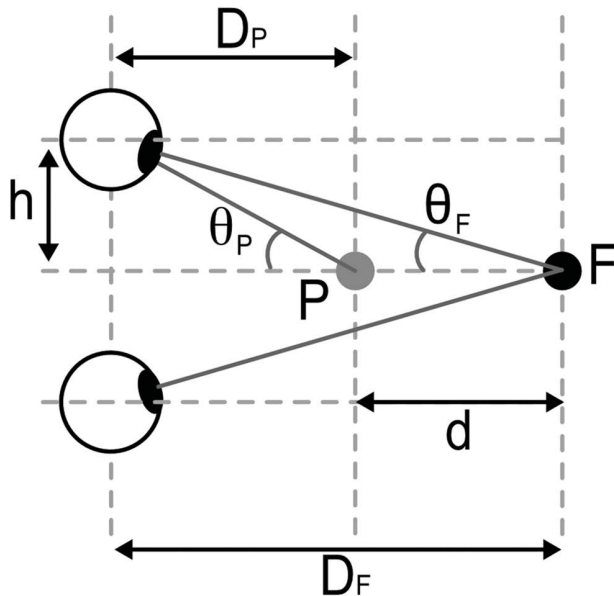


Fig 3. Diagram showing the geometry of estimating depth from retinal disparity. An observer fixates a point F which is at a distance D_F directly in front of the observer in the median plane (this produces the half vergence angle θ_F). A second point, P , on the median plane is placed at $D_P = D_F - d$ (which produces the half vergence angle θ_P). This produces a half retinal disparity of η , which is equal to $\theta_P - \theta_F$.

<https://doi.org/10.1371/journal.pcbi.1013506.g003>

Rearranging Equation 6, we can see that one way in which an observer could estimate d is to scale retinal disparity by an estimate of D_F provided by vergence together with knowledge of their interocular distance.

$$d = \frac{(D_F^2 + h^2) \tan(\eta)}{h + D_F \tan(\eta)} \quad (7)$$

This process is termed ‘disparity scaling’, as the same disparity can be produced by infinitely many depths and distances, so the measured disparity signal needs to be “scaled” in order for depth to be estimated [8].

Due to this geometric relationship, investigating the perception of depth from disparity, where vergence is the sole or primary distance cue, has been used to infer information about the accuracy of distance estimation from vergence. For example, if an observer makes a depth estimate, $\hat{d}$, given an object of depth d at a distance D_F , where vergence is the only distance cue, we can infer the vergence specified distance, $\hat{D}_F$, that they used to scale retinal disparity. This “scaling distance” [41,42] is given by:

$$\hat{D}_F = \frac{1}{2} \left(d + \sqrt{d^2 - 4h^2 + 4dh \cot(\eta)} \right) \quad (8)$$

One of the most influential studies in this area is that of Johnston [41]. In this study, observers judged whether disparity-defined horizontally orientated elliptical hemi-cylinders, viewed at a range of distances, had a circular cross section. It was found that for a hemi-cylinder to have a perceived circular cross-section it had to be physically ‘squashed’ in depth extent at near distances (approximately <80 cm) and physically ‘stretched’ in depth at far distances (approximately >80 cm). This pattern of results has been replicated in numerous studies with different types of stimuli, either viewed on a computer monitor [25–27,42–48], via a virtual reality headset [49,50], or with real-world physical stimuli [39,51].

These results are consistent with a progressive underestimation of distance as physical distance increases, but with an additional overestimation of near distances. Similar experiments have shown that vergence is lawfully combined with information from vertical disparity (a retinal cue to binocular viewing geometry) when estimating depth [43,52]. Here, with each cue in isolation the inferred scaling distance was an overestimate at near distances and a progressive underestimate as distance increased, but with an improvement in scaling when both cues varied consistently [43].

It is difficult to determine whether the pattern of results of depth, size and shape judgement tasks can be used to directly infer vergence specified distance. This inference would be based upon assuming that scaled retinal disparity is the sole cue to the distal property being estimated (e.g., depth/shape/size) and that vergence is the sole cue providing the distance estimate used to scale retinal disparity. These are significant assumptions given the difficulty of isolating sensory cues [40]. Additionally vergence responses can be driven independent of perceived depth [53] and depth can be perceived from diplopic images independent of vergence state [54]. Therefore, in what follows, we focus on direct distance estimation from vergence, rather than distance estimates inferred indirectly.

Summary of the present paper

We present a Bayesian model of how an observer might estimate the distal property of absolute distance from the proximal ocular vergence signal. Our focus is not on resolving the longstanding experimental debate as to whether observers can accurately estimate distance from vergence. Rather, our aim is to rigorously examine how an observer might do this based upon the proximal signal available to the sensory system. We find that the progressive underestimation of distance with increasing physical distance observed in experimental studies can be predicted to arise from observers trying to estimate the most likely distance in the world to have produced a measured but uncertain vergence signal. The amount of vergence uncertainty needed to account for the misestimation of distance is directly comparable to that which has been experimentally measured [55,56].

Methods and results

Probability of world distance from ocular convergence

Sensory signals are inherently stochastic [57]. As a result, for a given physically identical proximal cue observers can make repeated sensory estimates that differ from one another. One of the most widely used methods to model perceptual estimation from stochastic (noisy) signals is Bayesian inference. In this framework perception is seen as a process of inverting the generative function relating measured noisy proximal sensory cues to distal world properties. To do this, current sensory data is combined with prior knowledge that the observer holds about the probability of distal properties of the world. The estimate made from this probabilistic information is determined by the resulting posterior probability density function and the cost associated with making perceptual errors [20,23,58].

We adopt the Bayesian framework here to examine how observers might estimate absolute distance from vergence. Our approach is consistent with other studies that have examined how binocular viewing geometry and estimation from noisy sensory signals [59–64] contribute to our perception of distal world properties. All derivations and simulations reported in the paper were produced using Mathematica 14.1 in conjunction with the Mathstatistica plugin (version 2.73) [65] run on a M3 Pro MacBook Pro (macOS 14.6.1). Where possible, computations were spread over all cores using the parallel computing functionality in Mathematica and Mathstatistica.

We first define a probability density function $P(\hat{\theta})$ that specifies the likelihood of a given measured vergence state $\hat{\theta}$ with fixation on a point F at a distance D_F on the median plane (Fig 1b). For a given value of D_F there is a specific probability of observing any given vergence state $\hat{\theta}$. Let $\mathbf{D}$ be the domain of D_F , i.e., all possible distances in the world ($\mathbf{D} : 0 < D_F < \infty$) and let θ be the domain of $\hat{\theta}$, i.e., all possible vergence states given these distances ($\theta : 0 < \hat{\theta} < \frac{\pi}{2}$).

As is standard, we assume that an observer has an unbiased measurement of θ_F corrupted by zero-mean Gaussian noise with a standard deviation of σ .

$$\hat{\theta} \sim \mathcal{N}(\theta_F, \sigma^2) \quad (9)$$

Given the domain $\mathbf{D}$, this defines a truncated normal likelihood function, $p(\hat{\theta}|\theta_F)$, where θ_F is the true half-vergence angle corresponding to fixation on distance D_F and erf the standard error function $\text{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} dt$.

$$p(\hat{\theta}|\theta_F) = \frac{\sqrt{\frac{2}{\pi}}}{\sigma \left(\text{erf}\left(\frac{\theta_F}{\sqrt{2}\sigma}\right) - \text{erf}\left(\frac{\theta_F - \frac{\pi}{2}}{\sqrt{2}\sigma}\right) \right)} e^{-\frac{(\theta_F - \hat{\theta})^2}{2\sigma^2}} \quad (10)$$

Here, This likelihood function is defined as a function of vergence angle θ_F , not of distance D_F , which is the distance property of the world being estimated. As such we need to reformulate $p(\hat{\theta}|\theta_F)$ in terms of distance rather than vergence angle (Equation 5). To do this we recognise that for a random variable X , distributed according to $f_X(x)$, if $Y = g(X)$ is a differentiable monotonic transformation then Y is also a random variable [66].

$$f_Y(y) = f_X(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right| \quad (11)$$

Here $f_X(x)$ is the truncated normal probability density function representing the likelihood of observing a vergence angle $\hat{\theta}$ given that the true vergence angle is $x = \theta_F$. $f_Y(y)$ is transformed probability density function representing the likelihood of observing a vergence angle $\hat{\theta}$ given that the true fixation distance is $y = D_F$.

The inverse of the transform we wish to apply is

$$g^{-1}(y) = \theta_F = \tan^{-1}\left(\frac{h}{D_F}\right) \quad (12)$$

Whilst the tangent function is periodic over π , it is bijective and monotonic over the vergence domain $\mathbf{D}$. We can therefore use a *change of variables* to perform a closed form transform between probability density functions.

Differentiating [12] with respect to D_F gives

$$\frac{d}{dy} g^{-1}(y) = -\frac{h}{h^2 + D_F^2} \quad (13)$$

Substituting [10] and [13] into [11] gives

$$p(\hat{\theta}|D_F) = \frac{\sqrt{\frac{2}{\pi}}}{\sigma \left(\text{erf}\left(\frac{\theta_F}{\sqrt{2}\sigma}\right) - \text{erf}\left(\frac{\theta_F - \frac{\pi}{2}}{\sqrt{2}\sigma}\right) \right)} e^{-\frac{(\theta_F - \hat{\theta})^2}{2\sigma^2}} \cdot \left| \frac{h}{h^2 + D_F^2} \right| \quad (14)$$

With $y \in \mathbb{R}$, $D > 0$ and $h > 0$ this gives the transformed likelihood function

$$p(\hat{\theta}|D_F) = \frac{h\sqrt{\frac{2}{\pi}}}{(h^2 + D^2) \sigma \left(\text{erf}\left(\frac{\theta_F}{\sqrt{2}\sigma}\right) - \text{erf}\left(\frac{\theta_F - \frac{\pi}{2}}{\sqrt{2}\sigma}\right) \right)} e^{-\frac{(\theta_F - \hat{\theta})^2}{2\sigma^2}} \quad (15)$$

Here $\theta_F = \tan^{-1} \left(\frac{h}{D_F} \right)$ is the true vergence angle for the fixation distance D_F , and $\hat{\theta} = \tan^{-1} \left(\frac{h}{D_i} \right)$, i.e., the internal vergence angle measurement corresponding to a possible distance value D_i .

Therefore, for a given value of noise, σ , in the proximal vergence signal, the likelihood function $p(\hat{\theta}|D_F)$ represents current sensory information from vergence that can be used to estimate distance in the world.

In the Bayesian framework this likelihood function is combined with a prior over the distance ($p(D_F)$) using Bayes rule to obtain a posterior distribution of possible distances from vergence states.

$$p(D_F|\hat{\theta}) \propto p(\hat{\theta}|D_F) \cdot p(D_F) \quad (16)$$

The prior represents the *a priori* belief of encountering any possible distance in the world independent of current sensory information. For now, following numerous other studies [2,26,39,67–72], we make the assumption that the prior for distance is flat, or so broad as to be noninformative compared to the likelihood function. However, we examine this assumption below.

With this assumption, the posterior probability distribution has the same form as the likelihood function:

$$p(D_F|\hat{\theta}) = p(\hat{\theta}|D_F) = \frac{h\sqrt{\frac{2}{\pi}}}{(h^2 + D^2)\sigma \left(\operatorname{erf}\left(\frac{\theta_F}{\sqrt{2}\sigma}\right) - \operatorname{erf}\left(\frac{\theta_F - \frac{\pi}{2}}{\sqrt{2}\sigma}\right) \right)} e^{-\frac{(\theta_F - \hat{\theta})^2}{2\sigma^2}} \quad (17)$$

By inverting the generative model between distance and vergence we can express likelihood functions for vergence (Fig 4a) as corresponding posterior probability distributions for distance from vergence (Fig 4b). In Fig 4, for illustrative purposes, we use a value of 1 degree for σ [55,56]. We consider a range of values below.

The non-linear relationship between vergence and distance transforms the truncated Gaussian likelihood functions for vergence estimates from vergence $p(\hat{\theta}|\theta_F)$ into positively skewed likelihood functions for vergence estimates from distance $p(\hat{\theta}|D_F)$ and therefore (assuming a flat prior) distance estimates from vergence $p(D_F|\hat{\theta})$. This transformation

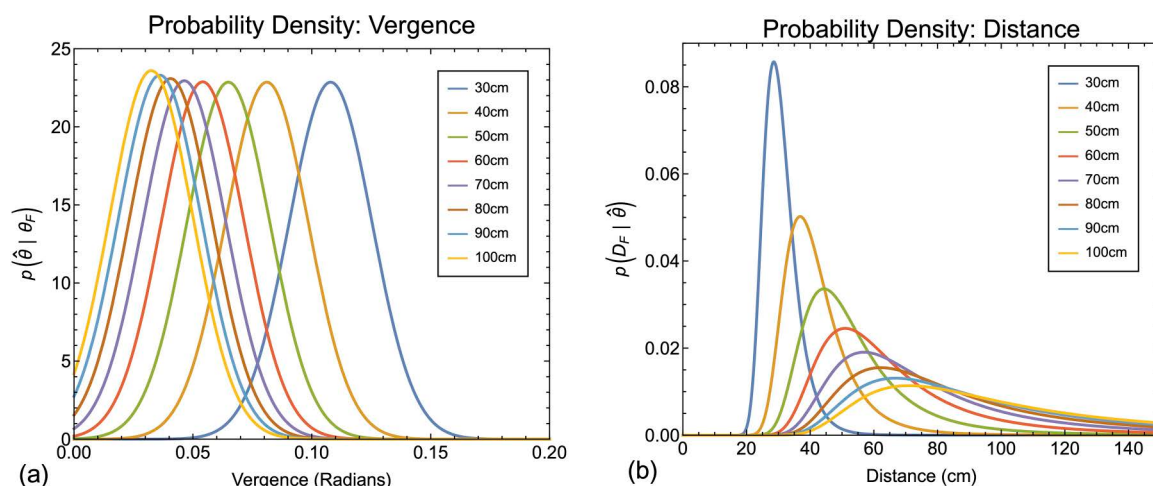


Fig 4. Probability density functions for vergence and distance. (a) Gaussian probability density functions for vergence (Equation 10) for distances between 30 cm and 100 cm. In all cases the sigma of the truncated Gaussian is 1 degree. (b) Corresponding probability density functions for distance from vergence (Equation 17). These are plotted for fixation distances at 10 cm intervals between 30 cm (roughly the nearest distance a person can comfortably fixate) and 140 cm (a distance at which more than 90% of the vergence range is used [8]).

<https://doi.org/10.1371/journal.pcbi.1013506.g004>

has this form because any given value of noise δ in the vergence signal θ_F corresponds to larger “overestimation” versus “underestimation” errors in distance [18]. That is

$$\left(\frac{h}{\tan(\theta_F - \delta)}\right) - D_F > D_F - \left(\frac{h}{\tan(\theta_F + \delta)}\right) \quad (18)$$

Making a perceptual estimate

The estimate of distance, $\hat{D}$, made from the posterior distributions shown in Fig 4b depends on the loss function defining the costs associated with misestimating the distal property of interest. The loss is itself a random variable $\mathcal{L}(\hat{D}_F | D_F)$, so the goal of the decision process is to minimise the expected loss associated with a given decision, commonly termed risk, $R_{\hat{D}_F}(D_F)$ [65].

$$R_{\hat{D}_F}(D_F) = E\left(\mathcal{L}(\hat{D}_F | D_F)\right) \quad (19)$$

The loss function is critical because it determines the perceptual estimate that is made from the posterior. This estimate is often termed an “optimal” estimate (e.g., Ernst and Banks [67]) due to it minimising the risk associated with a perceptual decision. However, the structure of the loss function is typically unknown and there are infinitely many possibilities. As a result, in most applications of Bayesian modelling the structure of the loss function is assumed (although see [73] for an attempt at an experimental estimate).

Here, to avoid assuming a loss function, but also to make our analysis tractable, following [74], we examine three of the most commonly used loss functions in Bayesian modelling. These are the quadratic loss:

$$\mathcal{L}(\hat{D}_F | D_F) = (D - D_F)^2 \quad (20)$$

The absolute error loss:

$$\mathcal{L}(\hat{D}_F | D_F) = |D - D_F| \quad (21)$$

And the zero-one loss:

$$\mathcal{L}(\hat{D}_F | D_F) = \begin{cases} 0 & \text{if } |D - D_F| \leq \varepsilon \\ 1 & \text{if } |D - D_F| > \varepsilon \end{cases} \quad (22)$$

For the zero-one loss, ε is the value of error ‘acceptable’ to the observer, which defines an interval within which all estimates accrue equal loss. Therefore, when $\varepsilon > 0$ there are a range of perceptual estimates which would accrue equal loss. When $\varepsilon \rightarrow 0$, the loss function approaches a Dirac Delta function, and the decision rule results in a single estimate. With a uniform prior and $\varepsilon = 0$ this is termed Maximum Likelihood Estimation (MLE) observers would be trying to estimate the most likely world property to have caused the measured noisy sensory cue [67].

For the quadratic loss function the estimate minimising risk is the mean of the posterior, for the absolute loss function it is the median and for the zero-one loss function it is the peak [74]. However, in many applications of Bayesian modelling the likelihood, prior and posterior distributions are assumed to be Gaussian over the domain of the units used for estimation [2,39,67–72,75,76]. In this case all three decision rules discussed above result in the same ‘optimal’ estimate being made. However, for many cues, the likelihood, prior and posterior will not be Gaussian, and these decision rules will provide different estimates. This makes assuming a loss function problematic.

Here, the peaks of $p(D_F|\hat{\theta})$ were estimated using Mathematica's *FindMaximum* function with the Conjugate Gradient method. It was not possible to find a simple closed form solution for the mean (expectation) of the $p(D_F|\hat{\theta})$, so we calculated this by numerical integration over the range 10 cm (estimated minimum possible vergence angle) to 6m (estimated upper limit on the utility of vergence as a cue to distance). One might argue that a larger domain should be used for integration, as observers are able to use vergence and stereoscopic cues at much larger distances than is commonly thought [77]. For the present paper this does not matter as increasing the domain of integration simply results in a greater progressive overestimation of distance (due to the tail of the skewed posterior distribution). Conversely, decreasing the domain decreases the overestimation but never results in underestimation.

To determine the median, we convert $p(D_F|\hat{\theta})$ into a cumulative density function:

$$CDF(p(D_F|\hat{\theta})) = \frac{\text{erf}\left(\frac{-2\theta_F + \pi}{2\sqrt{2}\sigma}\right) + \text{erf}\left(\frac{\theta_F - \hat{\theta}}{\sqrt{2}\sigma}\right)}{\text{erf}\left(\frac{\theta_F}{\sqrt{2}\sigma}\right) - \text{erf}\left(\frac{\theta_F - \frac{\pi}{2}}{\sqrt{2}\sigma}\right)} \quad (23)$$

We then solve this function at the 0.5 point. For a given $p(D_F|\hat{\theta})$ this represents the distance value at which 50% of the probability density has accumulated (Fig 5).

The image in Fig 6 plots the likelihood of perceived distance across a range of physical distances. Each column of pixels is a posterior probability density function as shown in Fig 4b. Overlaid on the image are the peak (blue line), expectation (green line) and median (red line) of the distributions. The three cost functions result in distinct predictions regarding the estimation of distance from vergence. The expectation predicts a progressive overestimation of distance as vergence-specified distance increases, which is the opposite pattern to that found in the literature. The median predicts virtually no perceptual bias at all, again, inconsistent with the experimental literature. By contrast, the peak results in a progressive underestimation of distance consistent with the experimental literature.

Given that the zero-one loss function was the only one to predict an underestimation of distance consistent with that found in the experimental literature, we further examined its properties to see whether it could quantitatively, as well as

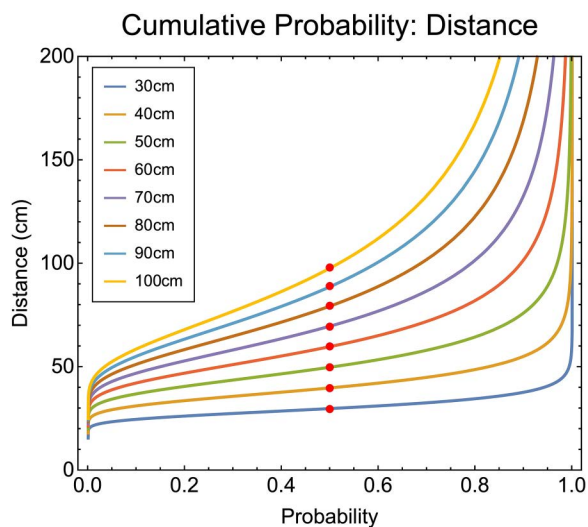


Fig 5. Inverse cumulative density functions for distance from vergence (corresponding probability density functions for distance shown in Fig 4b). The red points show the median of the functions. We show the inverse CDF's, rather than the CDF's, for ease of visualisation.

<https://doi.org/10.1371/journal.pcbi.1013506.g005>

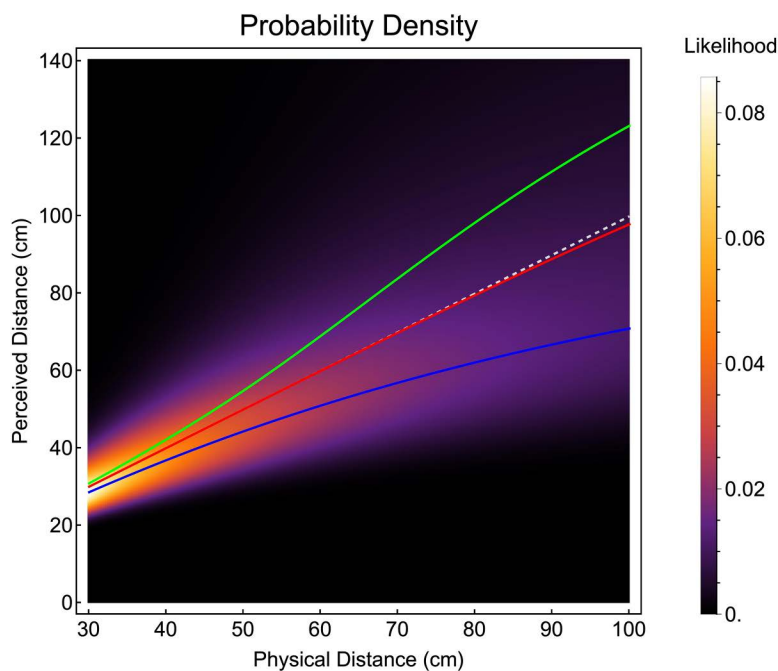


Fig 6. Probability density functions for perceived distance (vertical pixels), overlaid with distance estimates made by choosing the peak (blue line), expectation (green line) or median (red line) of the distributions. The diagonal dashed grey line represents veridical performance, i.e., accurate estimates of distance.

<https://doi.org/10.1371/journal.pcbi.1013506.g006>

qualitatively, predict distance underestimation from vergence. Fig 7 shows how the progressive underestimation of distance by choosing the peak lawfully varies for a range of noise levels in the vergence signal. As vergence noise increases, the underestimation of further distances also increases, with near distances being estimated most accurately. Overlaid for comparison is the data from Viguiet et al. [10].

Mon-Williams and Tresilian [18] recognised the asymmetric effect of noise in distance from vergence (Equation 18) and the effect that this would have for average distance estimates (they did not consider the nature of the posterior distribution), and proposed that “(i)n an attempt to compensate for this, the system could incorporate an underestimation bias for more distant targets” (p. 177). Here we show that with a different cost function this “necessarily speculative” (p. 177) suggestion is not needed. We next examine whether response distributions predicted by this model are consistent with data from the experimental literature.

Response distributions for distance from ocular convergence

The posterior probability density functions shown in Figs 4b and 6 represent the information available to the observer at a given instant in time, for example the point at which a perceptual estimate is made on a trial of an experiment. However, across trials the posterior probability density function will not be identical due to the stochastic nature of sensory noise. As such, despite often being assumed, posterior probability density functions do not necessarily show the same distribution as perceptual estimates across trials in an experiment (for discussion see [23]).

For a given physical distance D_F and vergence noise level σ_i , to estimate response distributions $p(\hat{D}_R | D_F)$ across trials in a hypothetical experiment, the following procedure was followed (Fig 8). On each trial i a sensory observation $\hat{\theta}_i$ was drawn from the distribution $p(\hat{\theta} | \theta_F)$ (Fig 8a). This represents the fact that for a constant value of D_F an observer's measured vergence angle will be different across trials due to sensory noise. This produces a unique trial-by-trial posterior

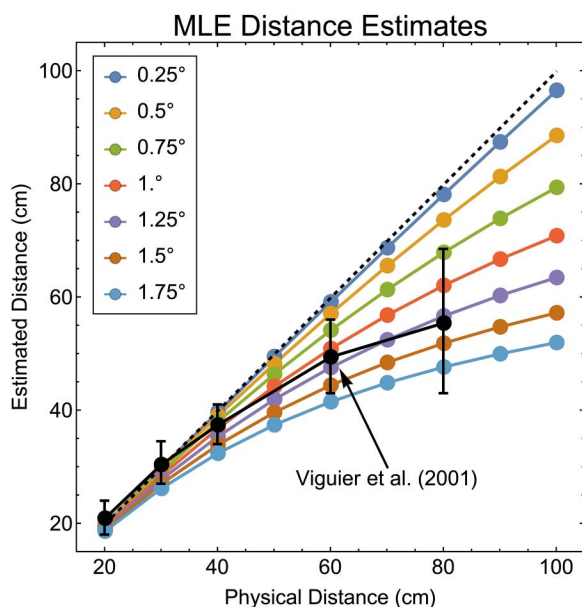


Fig 7. Maximum likelihood estimates of perceived distance (uniform steps of 10 cm between 20 cm and 100 cm) from ocular convergence for a range of noise values (uniform steps of 0.25° between 0.25° and 1.75°). The diagonal dashed grey line represents veridical performance, i.e., accurate estimates of distance. The black datapoints are manual distance estimates from ocular convergence from Viguiet et al. (2001) as described above (error bars show $\pm SD$). Note that the lines in the graph represent point estimates from the posterior distributions, not estimates over the course of an experiment. We compare the Viguiet et al. data to direct predictions from response distributions below (see main text).

<https://doi.org/10.1371/journal.pcbi.1013506.g007>

distribution of distance from vergence $p(D_F | \hat{\theta}_i)$ (Fig 8b). The peak of the posterior was taken as the observer's distance estimate on that trial. We simulated 100000 trials to build response distributions for a given physical distance and vergence noise level (Fig 8c).

Example response distributions for all distances with a vergence noise value of 1° are shown in Fig 9. Response distributions are skewed in a similar fashion to the per-trial posterior distributions. Overlaid on these histograms are best fitting parametric distributions found using Mathematica's *FindDistribution* function (this function uses the Bayesian information criterion together with priors over possible distribution types to select both the best distribution and the best parameters for that distribution). These functions represent a description of the data, not a computational model, thus, the best fitting parametric distribution (and its parameters) can differ across simulated data (identical conclusions are made throughout if we instead fit smooth kernel distributions). The peaks of the fitted parametric distributions (red dots in Fig 9) were taken as the most likely distance estimate an observer would give *across trials* for a given distance and vergence noise value.

Peaks of the response distributions for a range of vergence noise values (uniform steps of 0.25° between 0.25° and 1.75°) and distances (uniform steps of 10 cm between 20 cm and 100 cm) are shown in Fig 10 (main plot). Given these data we can infer the amount of noise in the vergence signal that would be required to account for the distance underestimation observed in Viguiet et al. [10]. To do this the data were fit by least squares with a quadratic surface (x , y , x^2 , y^2 , and xy terms). A second-order polynomial was chosen not as a model of the data but as a simple way of describing the data. The fit of the surface to the data was excellent, $R^2 = 0.99$.

Next we minimised the sum of the squared differences between the Viguiet et al. [10] data and the fitted surface using Mathematica's *Minimize* function. This suggested that given our Bayesian model of distance estimation, a vergence noise of $\sigma = 0.79^\circ$ would be needed to produce the experimentally observed distance misestimation in Viguiet et al. [10]. This value is directly comparable with experimental estimates of the magnitude of vergence noise [55,56]. The minimised

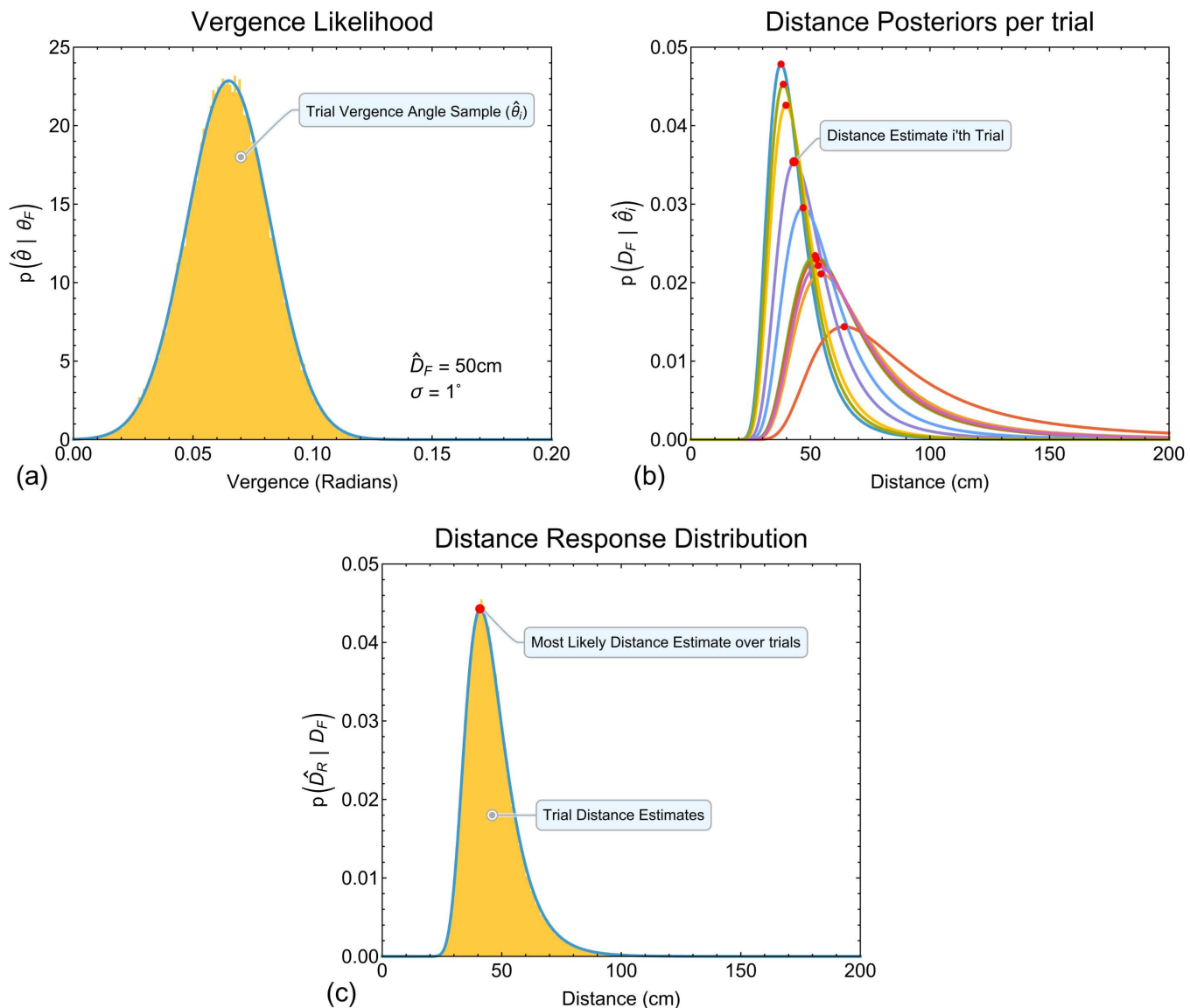


Fig 8. Illustrative example of how response distributions were calculated. (a) For a given distance and vergence noise level (here $D_F = 50\text{cm}$ and $\sigma_i = 1^\circ$), for each simulated experimental trial i , a random sample $\hat{\theta}_i$ (across trials collectively illustrated by the orange histogram) was drawn from the likelihood function for vergence $p(\hat{\theta} | \theta_F)$ (blue distribution curve). Each sample $\hat{\theta}_i$ represents the observer's measured vergence state on that trial. (b) This produces a unique trial-by-trial posterior $p(D_F | \hat{\theta}_i)$, here shown for ten examples. The peaks of the posteriors (red dots) were taken as the observer's distance estimates. (c) The response distribution (orange histogram) is the distribution of posterior peaks across trials. The blue distribution curve is the best fitting parametric distribution (see main text for details), the peak of which was taken as the most likely distance estimate an observer would make across trials for that specific distance and vergence noise level.

<https://doi.org/10.1371/journal.pcbi.1013506.g008>

fit of the vergence model to the Viguier et al. data is shown inset in Fig 10. The fit captures the experimental data very well $R^2 = 0.92$. Possibly the experimental data more rapidly shifts to distance underestimation compared to the model, however, this is difficult to definitively quantify given the error bounds around the experimental data (especially at larger distances).

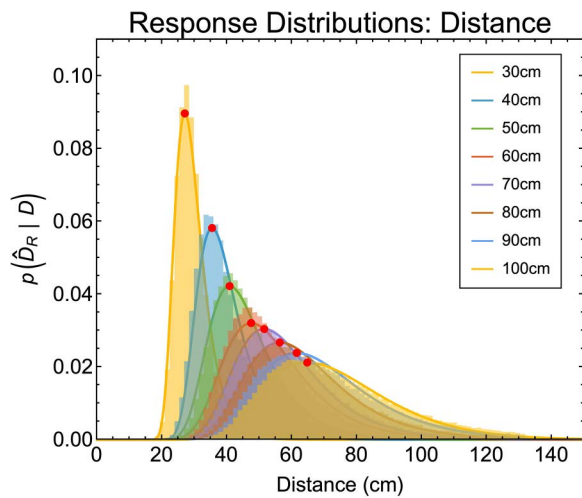


Fig 9. Shows example response distribution histograms for distance estimates across trials for a range of distances and a vergence noise value of 1°. Solid lines show the best fitting parametric distribution. Red points show the peaks of these parametric distributions.

<https://doi.org/10.1371/journal.pcbi.1013506.g009>

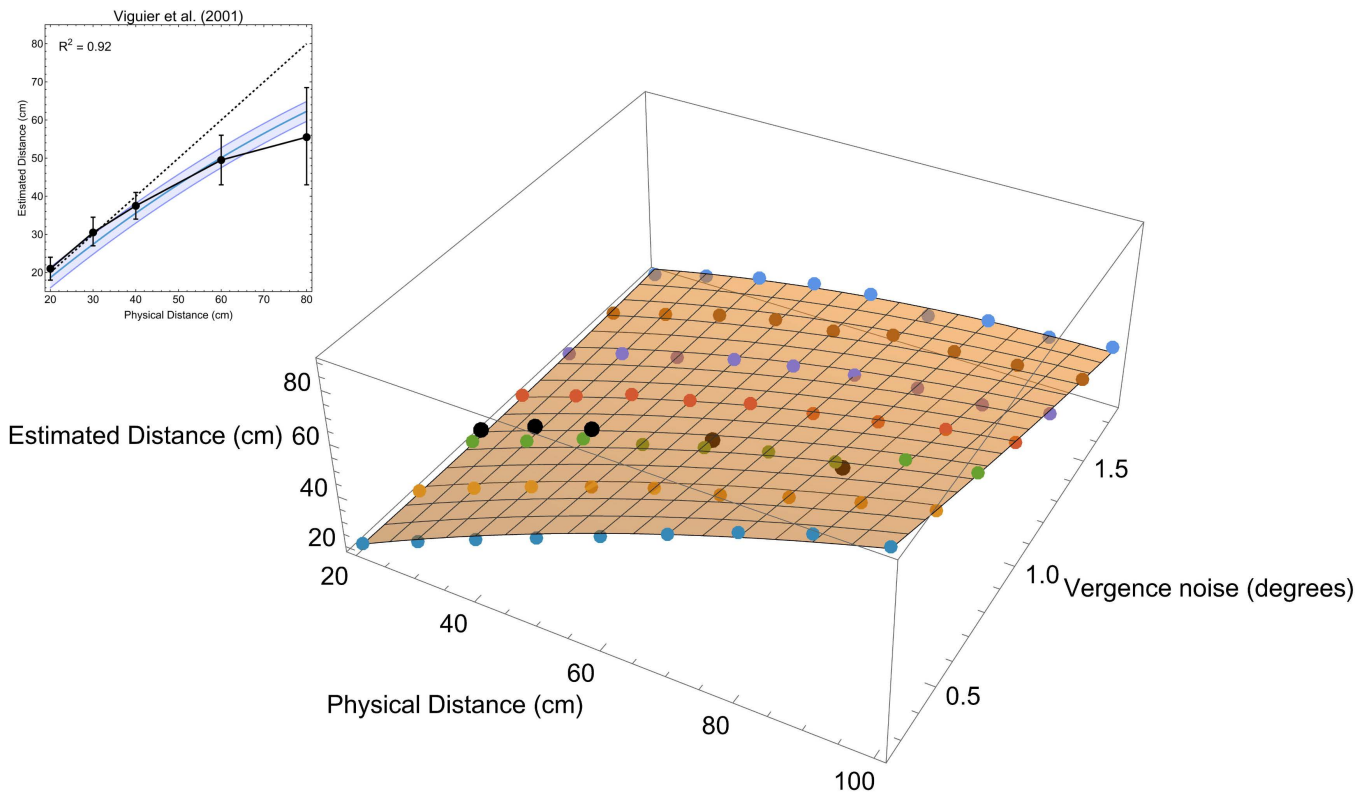


Fig 10. The main image shows peaks of response distributions for estimating distance from a noisy vergence signal (purple points) with the best fitting quadratic surface (semi-opaque orange surface). The black points show the Viguer et al. [10] (error bars show $\pm SD$). The data from Viguer et al. are positioned on the vergence noise axis such that the distance between the data and the fitted surface are minimised (minimum of the sum of squared differences). Inset shows the Viguer et al. data and the slice through the quadratic surface at this point. The fit to the data is excellent $R^2 = 0.92$. The shaded region in the inset shows 95% confidence intervals for single predictions.

<https://doi.org/10.1371/journal.pcbi.1013506.g010>

Bayesian priors for distance

Here we assumed that the prior for distance was flat or so broad as to be noninformative compared to the likelihood function, but we know that distances in the world are not uniformly distributed [78–80]. The choice of prior was made for two reasons. First, Bayesian models have been criticised for the proposition of priors to explain patterns of experimental data with little justification, i.e., the prior is just a redescription of the pattern in the data the experimenter wants to account for [81]. Second, whilst substantial progress has been made in directly measuring the statistical structure of the world (for an overview see the discussion section of [82]), there remain significant difficulties, rarely acknowledged, in interpreting measured frequency distributions of distal world properties as internal Bayesian priors.

A typical approach to measuring a distance prior is to place a scanner at positions in the environment and measure a frequency count of radial distance. The experimenter must make choices such as where in the environment to sample and normally positions the scanner at a minimum distance from objects so that they do not occlude the scanners full field of view. These choices affect the measured distributions. Scanners also have a minimum and maximum measurable distance, meaning that distances outside its range, by definition, cannot exist in the database. Minimum distances in current databases are around 2m (e.g., 3m in [78], 1m in [80], and approximately 2m in [83]).

We spend a significant amount of time interacting with objects within arm's reach, so databases such as these do not contain most of the distances we interact with. This is particularly problematic for modelling cues such as vergence, whose utility drops off rapidly with distance (Fig 2a). Scanners also sample the scene uniformly across their field of view, which is not true for humans who sample the scene based upon factors such as interest, saliency and task [84–86].

A key tenet of Bayesian modelling and Bayesian statistics in general [74] is that priors, likelihoods and posteriors *do not* simply represent frequency distributions, rather they represent distributions of *belief*; in our case, our belief about distances in the world. This allows Bayesian models to be derived to reason about properties that cannot be measured directly, which is not possible from a frequentist perspective. This is also why we ensured that the choice of a truncated normal distribution for the noisy vergence signal, rather than a normal distribution, did not affect our results. Belief about physically impossible distances could exist for a true Bayesian observer.

A final critical assumption is that observers accurately estimate the values of the distal property in the database, otherwise the measured frequency distribution will not represent the internal perceptual prior. One could argue that we have access to a rich array of sensory cues, coupled with the ability to calibrate cues based on sensory feedback [87,88], thus distance should be accurately estimated. However, this is inconsistent with a wide body of literature showing that observers misestimate absolute distance (and most other distal properties) in both simulated and real environments, with rich availability of visual cues [89–93]. Furthermore, integration of sensory cues to make a perceptual estimate does not necessarily result in accurate calibration of those same cues [26,94].

Internalisation of Bayesian priors

Given the above, we are not in the position to infer a distance prior in order to examine its effects on the perceptual estimation of distance from vergence. We can however examine the effect that biased perceptual estimation of distance from vergence could have on internalising a distance prior. To do this we simulated 200 semi-natural scenes in Matlab 2024b using 3D meshes of scanned natural objects (apple, pumpkin, sweetcorn, pomegranate, lemon, pinecone, nectarine, tomato, ginger, kiwi, garlic, carrot, sprout, and avocado) (see S2 Text). The high-resolution 3D meshes were loaded into Matlab using the gptoolbox [95].

Meshes were randomly positioned on a 35cm square planar surface area such that bounding circles around the objects, in the x-z plane, did not intersect (Fig 11). Meshes were positioned in the y dimension such that their minimum y coordinate touched the floor plane. Placement was sequential and if this resulted in all objects not being able to be placed

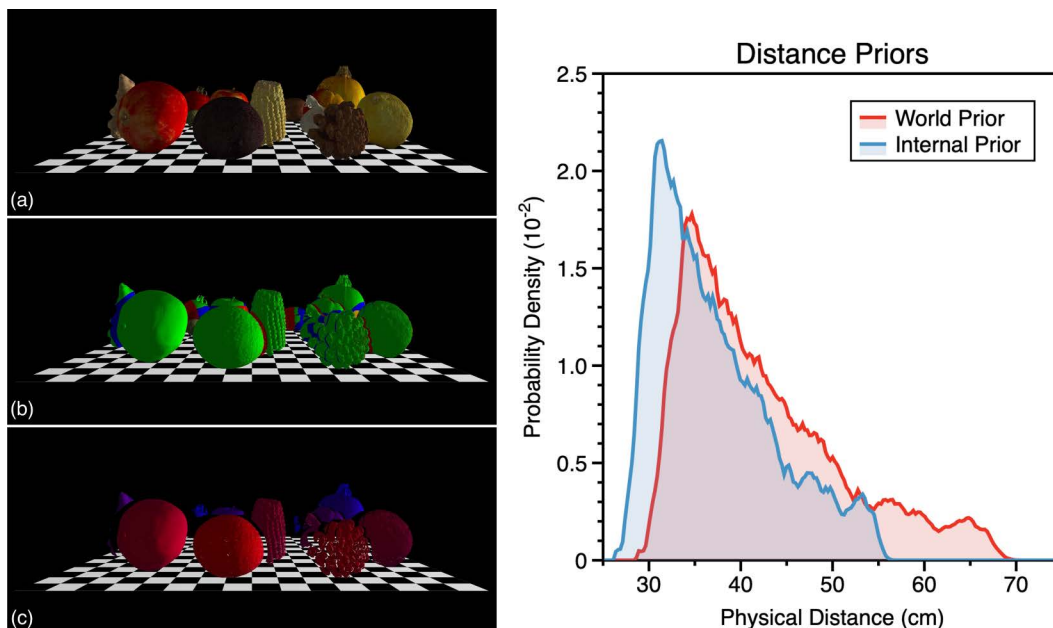


Fig 11. Estimating a distance prior. (a) Example scene created from scanned objects, the scene is rendered on a checkboard surface for visualisation, (b) same scene colour coded to show the portions of the objects viewable from each eye (green = both eyes, red = right eye only, blue = left eye only), (c) same scene now showing only the vertices viewable from both eyes, colour coded for radial distance from the cyclopean eye (red = near, blue = far), (d) distance priors calculated over all 200 scenes, red shows the distal prior present in the world, blue shows the prior which gets estimated from the vergence signal. Note: that in (a-c) the checkerboard surface is shown for ease of visualisation and was not part of the raytraced scenes. Additionally, the simulated viewpoint has been set to aid visualisation of the raytracing.

<https://doi.org/10.1371/journal.pcbi.1013506.g011>

in the scene without their bounding circles intersecting the scene was discarded. This process was repeated to create the 200 semi-natural scenes. In total there were just over 8 million mesh vertices per scene.

Each scene was then raytraced to determine the visibility of every mesh vertex to each eye. To do this we ray-traced each scene as viewed by an observer with an interocular distance of 6.5 cm positioned at $x=0$ cm, $y=8$ cm and $z=60$ cm relative to the centre of the 35 cm square planar area. Raytracing was performed using parallel processing with Embree (<https://www.embree.org>) accessed via the libigl C++ geometry processing library [96], called through the gptoolbox [95].

An example scene is shown in Fig 11a. This is shown again artificially coloured for visibility from each eye in Fig 11b and for radial distance to each vertex (from the cyclopean eye) viewable from both eyes in Fig 11c. Radial distances across scenes together constituted our world distance “prior” (Fig 11d). Note, that for the reasons stated, we are not claiming that this is a true estimate of the world prior for distance. Indeed, the structure of the prior is completely irrelevant for our reasoning.

Distances in the prior were then passed through the transform relating physical to perceived distance from vergence for the best fitting vergence noise level (Fig 10). Passing the world prior through this transform results in a number of interesting properties. Some far distances present in the physical scene are absent in the internalised prior and the probability of the remaining distances is contracted towards near distances. Additionally, some distances that do not exist in the world prior exist in the internalised prior. This highlights the critical importance of understanding how properties of the world are estimated when interpreting statistical priors derived directly from environmental measurements.

At what stage(s) of processing do priors act?

Existing studies generally implement priors at the level of the distal property being modelled, for example, distance [78] or slant/tilt [83]. However, it is possible that a prior could act at an earlier stage of processing. In the case of distance estimation from ocular convergence, one possibility is that rather than, or in addition to, a prior for distance (the distal property being estimated) one could have a prior for the convergence state of the eyes (the measured proximal sensory signal). In some senses priors over proximal signals could be considered more realistic as the sensory system has direct access to the signal. This would remove assumptions related to the distal property being accuracy estimated from the proximal signal for it to be instantiated in a prior.

In Fig 12 we show the “vergence prior” that would be produced by an observer fixating the distances in the world prior shown in Fig 11. As is clear, the prior is not flat, as we have assumed for our modelling.

Indeed, we can use the same change of variables technique as above to derive the distribution of world distances $p(D_F)$ that would result in a flat vergence prior (S3 Text).

$$p(D_F) = \frac{h}{(-b+a)(h^2 + D^2)} \quad (24)$$

In Fig 13 we plot $p(D_F)$ with $a = 0$, $b = \frac{\pi}{2}$ (i.e., the full vergence domain θ), and $h = 3.25\text{cm}$ (half the average interocular distance). As can be seen, the distance prior needed to produce a flat vergence likelihood has a greater probability of near distances compared to far, with a peak at 0. The preponderance of near distances would be consistent with far distances being less likely due to occlusion and perspective projection [97], so in this sense a flat vergence prior is consistent with far distances being less likely to be encountered. However, it is not clear that a peak is realistic or physically meaningful [86] or that the true world prior for distance would decrease in this way. This again highlights the problems inherent in interpreting priors (and using them to account for perceptual bias).

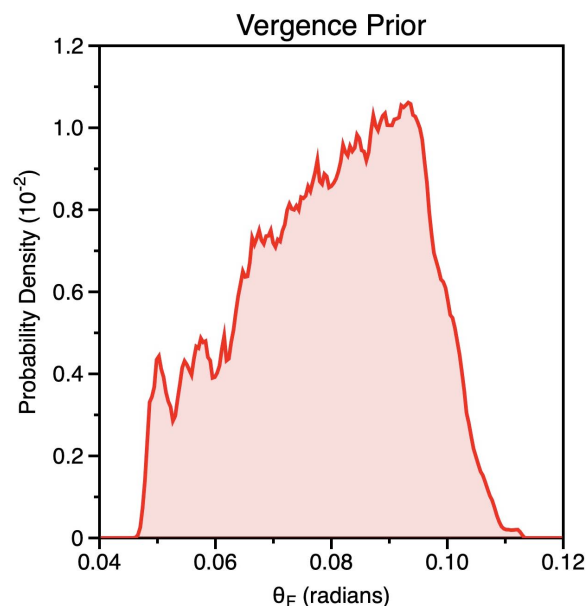


Fig 12. Vergence prior, calculated by passing the radial distances in the world prior in Fig 11b through the transform relating distance to ocular convergence (Equation 4) with an assumed interocular distance of 6.5 cm.

<https://doi.org/10.1371/journal.pcbi.1013506.g012>

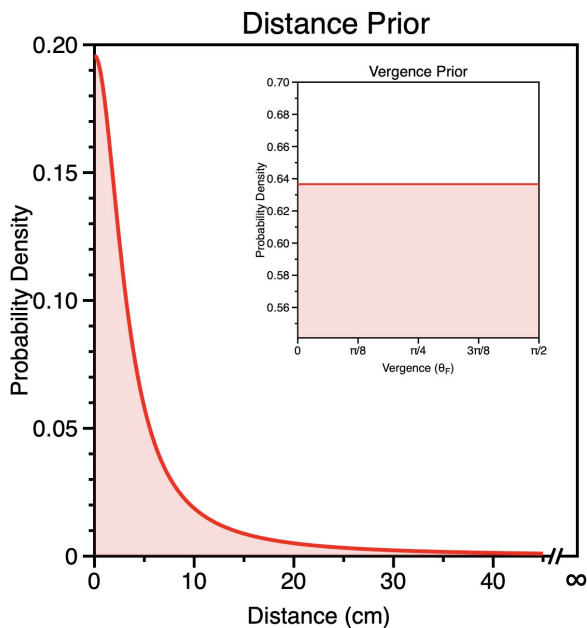


Fig 13. Distance prior which would be consistent with a flat vergence prior (inset). The domain of vergence is $\theta : 0 < \hat{\theta} < \frac{\pi}{2}$, as such the probability density function is a constant $\frac{2}{\pi}$. The corresponding domain for distance is $D : 0 < D_F < \infty$.

<https://doi.org/10.1371/journal.pcbi.1013506.g013>

This leads to the question of how such a “vergence prior” could be instantiated. In darkness the eyes are known to adopt a tonic state (dark vergence) of approximately 90 cm to 1 m [98]. This could be interpreted as a “prior” if it influenced distance estimation in the low light typical of experimental settings. It would likely have to operate after the stage at which vergence state (and its uncertainty) has been measured, as mis-convergence could cause diplopia. This is consistent with distance from vergence being underestimated despite correct physical convergence [10]. The control of convergence is also closely coupled to the accommodative system, which works in concert to maintain clear single binocular vision. The accommodative system has a tonic state (dark focus) around the same value as dark vergence.

If interpreted as a prior, the tonic state of the vergence and accommodation systems is not consistent with the progressive underestimation near distances. Instead, it would predict near distance overestimation as distance estimates would be pulled towards the centre of mass of the prior ~ 1 m. Whilst we are not averse to considering priors operating at multiple processing levels, we are very wary of overinterpreting any prior for the reasons previously stated. We have therefore focused on examining how distance might be (mis)estimated from a noisy ocular convergence signal without resorting to a prior, which is normally the route taken to explain perceptual bias.

Discussion

Summary

We have presented a Bayesian model of the perception of distance from ocular convergence. For a vergence noise value consistent with that measured experimentally [55,56] this model can account for the progressive underestimation of distance, despite accurate convergence [10]. This is achieved by observers trying to estimate the most likely distance to have produced a measured noisy vergence signal. Helmholtz ([99] p.318) recognised the importance of

the accuracy and precision of the vergence signal in making correct inferences about properties of the world when stating that:

“Owing to the uncertainty of our judgements as to the degree of convergence of the eyes, we are liable to have illusions also about the forms of things in space as seen binocularly. The interpretation of the visual phenomena would be correct if the amount of convergence were different, but it is not correct of the convergence actually used.”

It is not clear that our Bayesian model was exactly what Helmholtz had in mind in this quote, but it confirms that there are circumstances where the “degree of convergence” can be correct, but “uncertainty” in estimating this state can result in the misperception of absolute distance.

At all points we have aimed to be consistent with previous literature and be as transparent as possible in the modelling choices made. We are aware that Bayesian models offer a high degree of flexibility and can be criticised on this basis [81]. We have emphasised the critical importance of considering the nature of the transform function relating distal world properties to proximal sensory data, and the nature of the loss function chosen. Many examples of Bayesian modelling disregard the former and make assumptions about the latter. We have also highlighted the significant assumptions made in using Bayesian priors to account for perceptual bias, whether these priors are assumed, inferred or experimentally measured.

Our aim was not to experimentally resolve the debate around the experimental evidence of human use of vergence to estimate distance [3,10,16,43,55,100–103]. Rather, it was to rigorously analyse how the distal property of absolute distance might be estimated from the proximal vergence signal. Given the flexibility in scope of Bayesian models, we too made some modelling assumptions and simplifications, which we detailed. These include Gaussian noise in the vergence signal (here of a fixed value), an uninformative distance prior, and examination of three of the loss functions (despite their being infinitely many potentials). Future work could clearly examine the landscape of possible models further.

Experimentally measuring vergence noise

We would also like to highlight the difficulty one faces in experimentally inferring the level of noise in the vergence signal. One could measure the physiological state of the eyes using an eye tracker, however the inferred value could be affected by many factors related to the eye tracker, e.g., recording method, onboard filtering and sampling rate, recording duration, task and the specific metrics derived from the data [104]. Physiological measurements also do not take into account any upstream processes beyond the physical orientation of the eyes, so will likely underestimate the noise in the vergence signal. Alternatively, one could get observers to make absolute distance estimates from vergence and based on the variability of these infer the level of noise in the vergence signal. This is also problematic as it is extremely difficult to truly isolate the vergence cue from other cues to distance [16,40].

Vergence noise has been measured psychophysically using a nonius line method, in which an observer judges the horizontal distance between two vertical lines, one presented to each eye with a vertical offset between the two (i.e., a relative judgement). When estimated using this technique, the standard deviation of vergence decreases with distance, from 9.5 arc min to 4.0 arc min between 40 cm and 100 cm [105] and 3.75 arc min at 210 cm (Chopin et al., 2016). These values are much smaller than those used in our model and decrease with distance. In contrast, Brenner and van Damme [55] estimated a standard deviation for vergence noise when used as a cue to distance of 50–60 arc min. This latter value, which reflects the uncertainty in our ability to access vergence information as an absolute cue to distance, is much closer to the value required in our model to account for perceptual biases.

While the reason for this discrepancy is not clear, it is an example of the ‘absolute disparity anomaly’ [106] that describes how our ability to access vergence or disparity information to make absolute distance judgements is much impaired in comparison to its use in relative judgments. Chopin et al. [106] argue that, while relative depth judgments

are based on absolute disparity information, the latter is not directly accessible for absolute depth judgements, at least not with the same precision as they are used in relative judgements. As a result, absolute depth thresholds have been reported as anything between three [107,108] and thirty [109] times higher than relative depth.

Similarly, the uncertainty of distance judgements from vergence is a factor of around 10–20 times poorer than the level of vergence noise measured using a nonius task [105]. Brenner and van Damme [55] estimated the standard deviation for changes in distance from changes in vergence to be much lower (around 10 arc min) than for absolute judgements, in line with measures of vergence noise from a nonius task [55,105,106,110,111].

Variation in estimates will also be affected by numerous simple, but critical, aspects of the experimental procedure. For example, the time interval between presenting an object whose distance is to be estimated, and the object used to manually estimate that distance [10,55]. A time interval is needed to eliminate a disparity cue but introduces a memory component which will affect the inferred noise level. Tasks used to indirectly infer vergence specified distance suffer from the same problems and introduce numerous additional assumptions (discussed above).

As such, our focus has been on examining how an observer might, in principle, estimate distance from vergence within a normative Bayesian model, examining a range of noise values and common cost functions. We hope this provides guidance for experimentally testing models of distance estimation from vergence going forward and emphasises the importance of considering the generative function relating the measured proximal signal to distal world property being estimated.

Supporting information

S1 Text. Explanation of nodal point geometry.

(DOCX)

S2 Text. Description of how objects were 3D scanned.

(DOCX)

S3 Text. Description and equations for the flat vergence prior.

(DOCX)

Acknowledgments

The core idea of the model presented here was developed by the authors 20 years ago in the supervision by Hibbard of Scarfe's PhD thesis. Preliminary work was presented by poster at the 2004 European Conference on Visual Perception and the 2017 Vision Sciences Society (VSS) Conference. We thank Johannes Burge, Robin Held, Jenny Read, Wendy Adams, Andrew Glennerster and Mike Landy for their comments on the VSS poster. Additionally, we thank Eli Brenner for insightful comments when examining Scarfe's PhD thesis in 2007, of which the preliminaries of this work were present.

Author contributions

Conceptualization: Peter Scarfe, Paul B. Hibbard.

Data curation: Peter Scarfe, Paul B. Hibbard.

Formal analysis: Peter Scarfe, Paul B. Hibbard.

Investigation: Peter Scarfe, Paul B. Hibbard.

Methodology: Peter Scarfe, Paul B. Hibbard.

Project administration: Peter Scarfe, Paul B. Hibbard.

Resources: Peter Scarfe, Paul B. Hibbard.

Software: Peter Scarfe.

Supervision: Paul B. Hibbard.

Validation: Peter Scarfe, Paul B. Hibbard.

Visualization: Peter Scarfe.

Writing – original draft: Peter Scarfe.

Writing – review & editing: Peter Scarfe, Paul B. Hibbard.

References

- Ernst MO, Bühlhoff HH. Merging the senses into a robust percept. *Trends Cogn Sci*. 2004;8(4):162–9. <https://doi.org/10.1016/j.tics.2004.02.002> PMID: [15050512](#)
- Oruç I, Maloney LT, Landy MS. Weighted linear cue combination with possibly correlated error. *Vision Res*. 2003;43(23):2451–68. [https://doi.org/10.1016/s0042-6989\(03\)00435-8](https://doi.org/10.1016/s0042-6989(03)00435-8) PMID: [12972395](#)
- Brenner E, Smeets JB. Comparing extra-retinal information about distance and direction. *Vision Res*. 2000;40(13):1649–51. [https://doi.org/10.1016/s0042-6989\(00\)00062-6](https://doi.org/10.1016/s0042-6989(00)00062-6) PMID: [10814753](#)
- Wexler M, Panerai F, Lamouret I, Droulez J. Self-motion and the perception of stationary objects. *Nature*. 2001;409(6816):85–8. <https://doi.org/10.1038/35051081> PMID: [11343118](#)
- Knill DC. Discrimination of planar surface slant from texture: human and ideal observers compared. *Vision Res*. 1998;38(11):1683–711. [https://doi.org/10.1016/s0042-6989\(97\)00325-8](https://doi.org/10.1016/s0042-6989(97)00325-8) PMID: [9747503](#)
- Knill DC. Surface orientation from texture: ideal observers, generic observers and the information content of texture cues. *Vision Res*. 1998;38(11):1655–82. [https://doi.org/10.1016/s0042-6989\(97\)00324-6](https://doi.org/10.1016/s0042-6989(97)00324-6) PMID: [9747502](#)
- Lovell PG, Bloj M, Harris JM. Optimal integration of shading and binocular disparity for depth perception. *J Vis*. 2012;12(1):1. <https://doi.org/10.1167/12.1.1> PMID: [22214563](#)
- Howard IP, Rogers BJ. *Seeing in Depth: Depth Perception*. Toronto: I Porteous; 2002.
- Tresilian JR, Mon-Williams M, Kelly BM. Increasing confidence in vergence as a cue to distance. *Proc Biol Sci*. 1999;266(1414):39–44. <https://doi.org/10.1098/rspb.1999.0601> PMID: [10081157](#)
- Viguier A, Clément G, Trotter Y. Distance perception within near visual space. *Perception*. 2001;30(1):115–24. <https://doi.org/10.1068/p3119> PMID: [11257974](#)
- Backus BT, Matza-Brown D. The contribution of vergence change to the measurement of relative disparity. *J Vis*. 2003;3(11):737–50. <https://doi.org/10.1167/3.11.8> PMID: [14765957](#)
- Howard IP. Vergence modulation as a cue to movement in depth. *Spat Vis*. 2008;21(6):581–92. <https://doi.org/10.1163/156856808786451417> PMID: [19017484](#)
- Wexler M, van Boxtel JJA. Depth perception by the active observer. *Trends Cogn Sci*. 2005;9(9):431–8. <https://doi.org/10.1016/j.tics.2005.06.018> PMID: [16099197](#)
- Cormack R, Fox R. The computation of retinal disparity. *Percept Psychophys*. 1985;37(2):176–8. <https://doi.org/10.3758/bf03202855> PMID: [4011374](#)
- Hershenson MH. *Visual space perception: a primer*. London: MIT Press; 1999.
- Linton P. Does vision extract absolute distance from vergence? *Atten Percept Psychophys*. 2020;82(6):3176–95. <https://doi.org/10.3758/s13414-020-02006-1> PMID: [32406005](#)
- Cutting JE, Vishton PM. Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth. In: Epstein W, Rogers S, editors. *Perception of Space and Motion*. San Diego: Academic Press; 1995.
- Mon-Williams M, Tresilian JR. Some recent studies on the extraretinal contribution to distance perception. *Perception*. 1999;28(2):167–81. <https://doi.org/10.1068/p2737> PMID: [10615458](#)
- Pizlo Z. Perception viewed as an inverse problem. *Vision Res*. 2001;41(24):3145–61. [https://doi.org/10.1016/s0042-6989\(01\)00173-0](https://doi.org/10.1016/s0042-6989(01)00173-0) PMID: [11711140](#)
- Knill DC, Richards W. *Perception as Bayesian Inference*. Cambridge: Cambridge University Press; 1996.
- Mamassian P, Landy MS, Maloney LT. Bayesian modelling of visual perception. In: Rao RPN, Olshausen BA, Lewicki MS, editors. *Probabilistic models of the brain: perception and neural function*. 2002. p. 13–36.
- Berger JO. *Statistical decision theory and bayesian analysis*. New York: Springer-Verlag; 1985.
- Ma WJ. Bayesian decision models: a primer. *Neuron*. 2019;104(1):164–75.
- Trommershauser J, Körding KP, Landy MS. *Sensory Cue Integration*. Oxford: Oxford University Press; 2011.

25. Scarfe P, Hibbard PB. Reverse correlation reveals how observers sample visual information when estimating three-dimensional shape. *Vision Res.* 2013;86:115–27. <https://doi.org/10.1016/j.visres.2013.04.016> PMID: [23665429](#)
26. Scarfe P, Hibbard PB. Statistically optimal integration of biased sensory estimates. *J Vis.* 2011;11(7).
27. Scarfe P, Hibbard PB. Disparity-defined objects moving in depth do not elicit three-dimensional shape constancy. *Vision Res.* 2006;46(10):1599–610. <https://doi.org/10.1016/j.visres.2005.11.002> PMID: [16364392](#)
28. Macpherson F, Batty C. Redefining Illusion and Hallucination in Light of New Cases. *Philosophical Issues.* 2016;26(1):263–96. <https://doi.org/10.1111/phis.12086>
29. Burge J, Girshick AR, Banks MS. Visual-haptic adaptation is determined by relative reliability. *J Neurosci.* 2010;30(22):7714–21. <https://doi.org/10.1523/JNEUROSCI.6427-09.2010> PMID: [20519546](#)
30. Poulton EC. *Bias in Quantifying Judgments.* London: Lawrence Erlbaum Associates. 1989.
31. Dennett DC. *Consciousness explained.* Boston: Little, Brown and Co; 1991. xiii, 511 p.: ill. p.
32. Durgin FH, Leonard-Solis K, Masters O, Schmelz B, Li Z. Expert performance by athletes in the verbal estimation of spatial extents does not alter their perceptual metric of space. *Iperception.* 2012;3(5):357–67. <https://doi.org/10.1068/i0498> PMID: [22833782](#)
33. Swenson HA. The Relative Influence of Accommodation and Convergence in the Judgment of Distance. *The Journal of General Psychology.* 1932;7(2):360–80. <https://doi.org/10.1080/00221309.1932.9918473>
34. Gogel WC, Tietz JD. Absolute motion parallax and the specific distance tendency. *Perception & Psychophysics.* 1973;13(2):284–92. <https://doi.org/10.3758/bf03214141>
35. Gogel WC. The sensing of retinal size. *Vision Res.* 1969;9(9):1079–94. [https://doi.org/10.1016/0042-6989\(69\)90049-2](https://doi.org/10.1016/0042-6989(69)90049-2) PMID: [5350376](#)
36. Tresilian JR, Mon-Williams M. Getting the measure of vergence weight in nearness perception. *Exp Brain Res.* 2000;132(3):362–8. <https://doi.org/10.1007/s002210000333> PMID: [10883384](#)
37. Mon-Williams M, Tresilian JR, Roberts A. Vergence provides veridical depth perception from horizontal retinal image disparities. *Exp Brain Res.* 2000;133(3):407–13. <https://doi.org/10.1007/s002210000410> PMID: [10958531](#)
38. Keefe BD, Watt SJ. The role of binocular vision in grasping: a small stimulus-set distorts results. *Exp Brain Res.* 2009;194(3):435–44. <https://doi.org/10.1007/s00221-009-1718-4> PMID: [19198815](#)
39. Watt SJ, Akeley K, Ernst MO, Banks MS. Focus cues affect perceived depth. *J Vis.* 2005;5(10):834–62. <https://doi.org/10.1167/5.10.7> PMID: [16441189](#)
40. Zabulis X, Backus BT. Starry night: a texture devoid of depth cues. *J Opt Soc Am A Opt Image Sci Vis.* 2004;21(11):2049–60. <https://doi.org/10.1364/josaa.21.002049> PMID: [15535362](#)
41. Johnston EB. Systematic distortions of shape from stereopsis. *Vision Res.* 1991;31(7–8):1351–60. [https://doi.org/10.1016/0042-6989\(91\)90056-b](https://doi.org/10.1016/0042-6989(91)90056-b) PMID: [1891823](#)
42. Glennerster A, Rogers BJ, Bradshaw MF. Stereoscopic depth constancy depends on the subject's task. *Vision Res.* 1996;36(21):3441–56. [https://doi.org/10.1016/0042-6989\(96\)00090-9](https://doi.org/10.1016/0042-6989(96)00090-9) PMID: [8977011](#)
43. Bradshaw MF, Glennerster A, Rogers BJ. The effect of display size on disparity scaling from differential perspective and vergence cues. *Vision Res.* 1996;36(9):1255–64. [https://doi.org/10.1016/0042-6989\(95\)00190-5](https://doi.org/10.1016/0042-6989(95)00190-5) PMID: [8711905](#)
44. Glennerster A, Rogers BJ, Bradshaw MF. Cues to viewing distance for stereoscopic depth constancy. *Perception.* 1998;27(11):1357–65. <https://doi.org/10.1068/p271357> PMID: [10505180](#)
45. Todd JT. The visual perception of 3D shape. *Trends Cogn Sci.* 2004;8(3):115–21. <https://doi.org/10.1016/j.tics.2004.01.006> PMID: [15301751](#)
46. Todd JT, Chen L, Norman JF. On the relative salience of Euclidean, affine, and topological structure for 3-D form discrimination. *Perception.* 1998;27(3):273–82. <https://doi.org/10.1068/p270273> PMID: [9775311](#)
47. Todd JT, Norman JF. The visual perception of 3-D shape from multiple cues: are observers capable of perceiving metric structure? *Percept Psychophys.* 2003;65(1):31–47. <https://doi.org/10.3758/bf03194781> PMID: [12699307](#)
48. Todd JT, Tittle JS, Norman JF. Distortions of three-dimensional space in the perceptual analysis of motion and stereo. *Perception.* 1995;24(1):75–86. <https://doi.org/10.1068/p240075> PMID: [7617420](#)
49. Hornsey RL, Hibbard PB, Scarfe P. Size and shape constancy in consumer virtual reality. *Behav Res Methods.* 2020;52(4):1587–98. <https://doi.org/10.3758/s13428-019-01336-9> PMID: [32399659](#)
50. Hornsey RL, Hibbard PB. Contributions of pictorial and binocular cues to the perception of distance in virtual reality. *Virtual Reality.* 2021;25(4):1087–103. <https://doi.org/10.1007/s10055-021-00500-x>
51. Bradshaw MF, Parton AD, Glennerster A. The task-dependent use of binocular disparity and motion parallax information. *Vision Res.* 2000;40(27):3725–34. [https://doi.org/10.1016/s0042-6989\(00\)00214-5](https://doi.org/10.1016/s0042-6989(00)00214-5) PMID: [11090665](#)
52. Rogers BJ, Bradshaw MF. Vertical disparities, differential perspective and binocular stereopsis. *Nature.* 1993;361(6409):253–5. <https://doi.org/10.1038/361253a0> PMID: [8423851](#)
53. Masson GS, Busetini C, Miles FA. Vergence eye movements in response to binocular disparity without depth perception. *Nature.* 1997;389(6648):283–6. <https://doi.org/10.1038/38496> PMID: [9305842](#)

54. Lugtigheid AJ, Wilcox LM, Allison RS, Howard IP. Vergence eye movements are not essential for stereoscopic depth. *Proc Biol Sci*. 2013;281(1776):20132118. <https://doi.org/10.1098/rspb.2013.2118> PMID: [24352941](#)
55. Brenner E, van Damme WJ. Judging distance from ocular convergence. *Vision Res*. 1998;38(4):493–8. [https://doi.org/10.1016/s0042-6989\(97\)00236-8](https://doi.org/10.1016/s0042-6989(97)00236-8) PMID: [9536373](#)
56. Richards W, Miller JF. Convergence as a cue to depth. *Perception & Psychophysics*. 1969;5(5):317–20. <https://doi.org/10.3758/bf03209573>
57. Hubel DH, Wiesel TN. Ferrier lecture. Functional architecture of macaque monkey visual cortex. *Proc R Soc Lond B Biol Sci*. 1977;198(1130):1–59. <https://doi.org/10.1098/rspb.1977.0085> PMID: [20635](#)
58. Kersten D, Mamassian P, Yuille A. Object perception as Bayesian inference. *Annu Rev Psychol*. 2004;55:271–304. <https://doi.org/10.1146/annurev.psych.55.090902.142005> PMID: [14744217](#)
59. Lages M. Bayesian models of binocular 3-D motion perception. *J Vis*. 2006;6(4):508–22. <https://doi.org/10.1167/6.4.14> PMID: [16889483](#)
60. Backus BT, Banks MS, van Ee R, Crowell JA. Horizontal and vertical disparity, eye position, and stereoscopic slant perception. *Vision Res*. 1999;39(6):1143–70. [https://doi.org/10.1016/s0042-6989\(98\)00139-4](https://doi.org/10.1016/s0042-6989(98)00139-4) PMID: [10343832](#)
61. Welchman AE, Lam JM, Bühlhoff HH. Bayesian motion estimation accounts for a surprising bias in 3D vision. *Proc Natl Acad Sci U S A*. 2008;105(33):12087–92. <https://doi.org/10.1073/pnas.0804378105> PMID: [18697948](#)
62. Ji H, Fermüller C. Noise causes slant underestimation in stereo and motion. *Vision Res*. 2006;46(19):3105–20. <https://doi.org/10.1016/j.vis-res.2006.04.010> PMID: [16750551](#)
63. Gepshtein S, Banks MS. Viewing geometry determines how vision and haptics combine in size perception. *Curr Biol*. 2003;13(6):483–8. [https://doi.org/10.1016/s0960-9822\(03\)00133-7](https://doi.org/10.1016/s0960-9822(03)00133-7) PMID: [12646130](#)
64. Bonnen K, Czuba TB, Whritner JA, Kohn A, Huk AC, Cormack LK. Binocular viewing geometry shapes the neural representation of the dynamic three-dimensional environment. *Nat Neurosci*. 2020;23(1):113–21. <https://doi.org/10.1038/s41593-019-0544-7> PMID: [31792466](#)
65. Rose C, Smith MD. *Mathematical Statistics with Mathematica*. 2013.
66. Blitzstein JK, Hwang J. *Introduction to Probability*. London: CRC Press; 2015.
67. Ernst MO, Banks MS. Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*. 2002;415(6870):429–33. <https://doi.org/10.1038/415429a> PMID: [11807554](#)
68. Rosas P, Wagemans J, Ernst MO, Wichmann FA. Texture and haptic cues in slant discrimination: reliability-based cue weighting without statistically optimal cue combination. *J Opt Soc Am A Opt Image Sci Vis*. 2005;22(5):801–9. <https://doi.org/10.1364/josaa.22.000801> PMID: [15898539](#)
69. Helbig HB, Ernst MO. Optimal integration of shape information from vision and touch. *Exp Brain Res*. 2007;179(4):595–606. <https://doi.org/10.1007/s00221-006-0814-y> PMID: [17225091](#)
70. Rohde M, van Dam LCJ, Ernst M. Statistically optimal multisensory cue integration: A practical tutorial. *Multisens Res*. 2016;29(4–5):279–317.
71. Hillis JM, Watt SJ, Landy MS, Banks MS. Slant from texture and disparity cues: optimal cue combination. *J Vis*. 2004;4(12):967–92. <https://doi.org/10.1167/4.12.1> PMID: [15669906](#)
72. Young MJ, Landy MS, Maloney LT. A perturbation analysis of depth perception from combinations of texture and motion cues. *Vision Res*. 1993;33(18):2685–96. [https://doi.org/10.1016/0042-6989\(93\)90228-o](https://doi.org/10.1016/0042-6989(93)90228-o) PMID: [8296465](#)
73. Körding KP, Wolpert DM. The loss function of sensorimotor learning. *Proc Natl Acad Sci U S A*. 2004;101(26):9839–42. <https://doi.org/10.1073/pnas.0308394101> PMID: [15210973](#)
74. Lee PM. *Bayesian Statistics: An Introduction*. 4th ed. Chichester: John Wiley and Sons (Ltd); 2012.
75. Hillis JM, Ernst MO, Banks MS, Landy MS. Combining sensory information: mandatory fusion within, but not between, senses. *Science*. 2002;298(5598):1627–30. <https://doi.org/10.1126/science.1075396> PMID: [12446912](#)
76. Ernst MO. Learning to integrate arbitrary signals from vision and touch. *J Vis*. 2007;7(5):7.1–14. <https://doi.org/10.1167/7.5.7> PMID: [18217847](#)
77. Palmisano S, Gillam B, Govan DG, Allison RS, Harris JM. Stereoscopic perception of real depths at large distances. *J Vis*. 2010;10(6):19. <https://doi.org/10.1167/10.6.19> PMID: [20884568](#)
78. Yang Z, Purves D. A statistical explanation of visual space. *Nat Neurosci*. 2003;6(6):632–40. <https://doi.org/10.1038/nn1059> PMID: [12754512](#)
79. Burge J, Geisler WS. Optimal disparity estimation in natural stereo images. *J Vis*. 2014;14(2):1. <https://doi.org/10.1167/14.2.1> PMID: [24492596](#)
80. Adams WJ, Elder JH, Graf EW, Leyland J, Lugtigheid AJ, Murry A. The Southampton-York Natural Scenes (SYNS) dataset: Statistics of surface attitude. *Sci Rep*. 2016;6:35805. <https://doi.org/10.1038/srep35805> PMID: [27782103](#)
81. Bowers JS, Davis CJ. Bayesian just-so stories in psychology and neuroscience. *Psychol Bull*. 2012;138(3):389–414. <https://doi.org/10.1037/a0026450> PMID: [22545686](#)
82. Iyer AV, Burge J. Depth variation and stereo processing tasks in natural scenes. *J Vis*. 2018;18(6):4. <https://doi.org/10.1167/18.6.4> PMID: [30029214](#)
83. Burge J, McCann BC, Geisler WS. Estimating 3D tilt from local image cues in natural scenes. *J Vis*. 2016;16(13):2. <https://doi.org/10.1167/16.13.2> PMID: [27738702](#)
84. Gibaldi A, Banks MS. Binocular Eye Movements Are Adapted to the Natural Environment. *J Neurosci*. 2019;39(15):2877–88. <https://doi.org/10.1523/JNEUROSCI.2591-18.2018> PMID: [30733219](#)

85. Gibaldi A, Canessa A, Sabatini SP. The Active Side of Stereopsis: Fixation Strategy and Adaptation to Natural Environments. *Sci Rep*. 2017;7:44800. <https://doi.org/10.1038/srep44800> PMID: [28317909](#)
86. Sprague WW, Cooper EA, Tošić I, Banks MS. Stereopsis is adaptive for the natural environment. *Sci Adv*. 2015;1(4):e1400254. <https://doi.org/10.1126/sciadv.1400254> PMID: [26207262](#)
87. Atkins JE, Fiser J, Jacobs RA. Experience-dependent visual cue integration based on consistencies between visual and haptic percepts. *Vision Res*. 2001;41(4):449–61. [https://doi.org/10.1016/S0042-6989\(00\)00254-6](https://doi.org/10.1016/S0042-6989(00)00254-6) PMID: [11166048](#)
88. Atkins JE, Jacobs RA, Knill DC. Experience-dependent visual cue recalibration based on discrepancies between visual and haptic percepts. *Vision Res*. 2003;43(25):2603–13. [https://doi.org/10.1016/S0042-6989\(03\)00470-X](https://doi.org/10.1016/S0042-6989(03)00470-X) PMID: [14552802](#)
89. Cuijpers RH, Kappers AM, Koenderink JJ. Large systematic deviations in visual parallelism. *Perception*. 2000;29(12):1467–82. <https://doi.org/10.1068/p3041> PMID: [11257970](#)
90. Hecht H, van Doorn AJ, Koenderink JJ. Compression of visual space in natural scenes and in their photographic counterparts. *Percept Psychophys*. 1999;61(7):1269–86. <https://doi.org/10.3758/bf03206179> PMID: [10572457](#)
91. Koenderink JJ, van Doorn AJ, Kappers AML, Lappin JS. Large-scale visual frontoparallels under full-cue conditions. *Perception*. 2002;31(12):1467–75. <https://doi.org/10.1068/p3295> PMID: [12916671](#)
92. Koenderink JJ, van Doorn AJ, Kappers AML, Todd JT. Pappus in optical space. *Percept Psychophys*. 2002;64(3):380–91. <https://doi.org/10.3758/bf03194711> PMID: [12049279](#)
93. Koenderink JJ, van Doorn AJ, Lappin JS. Direct measurement of the curvature of visual space. *Perception*. 2000;29(1):69–79. <https://doi.org/10.1068/p2921> PMID: [10820592](#)
94. Smeets JBJ, van den Dobbelaars JJ, de Grave DDJ, van Beers RJ, Brenner E. Sensory integration does not lead to sensory calibration. *Proc Natl Acad Sci U S A*. 2006;103(49):18781–6. <https://doi.org/10.1073/pnas.0607687103> PMID: [17130453](#)
95. Jacobson A. Gptoolbox: Geometry Processing Toolbox. 2024.
96. Jacobson A, Panozzo D, others. libigl: A simple C++ geometry processing library. 2018.
97. Hibbard PB. A statistical model of binocular disparity. *Visual Cognition*. 2007;15(2):149–65. <https://doi.org/10.1080/13506280600648018>
98. Jaschinski W, Jainta S, Hoormann J, Walper N. Objective vs subjective measurements of dark vergence. *Ophthalmic Physiol Opt*. 2007;27(1):85–92.
99. Helmholtz H. Helmholtz's Treatise on Physiological Optics. Southhall JPC, editor. Thoemmes Press; 1925.
100. Linton P. Minimal theory of 3D vision: new approach to visual scale and visual shape. *Philos Trans R Soc Lond B Biol Sci*. 2023;378(1869):20210455. <https://doi.org/10.1098/rstb.2021.0455> PMID: [36511406](#)
101. Linton P. Does vergence affect perceived size? *Vision*. 2021;5(3).
102. van Damme W, Brenner E. The distance used for scaling disparities is the same as the one used for scaling retinal size. *Vision Res*. 1997;37(6):757–64. [https://doi.org/10.1016/S0042-6989\(96\)00213-1](https://doi.org/10.1016/S0042-6989(96)00213-1) PMID: [9156220](#)
103. Bradshaw MF, Elliott KM, Watt SJ, Hibbard PB, Davies IRL, Simpson PJ. Binocular cues and the control of prehension. *Spat Vis*. 2004;17(1–2):95–110. <https://doi.org/10.1163/156856804322778288> PMID: [15078014](#)
104. Niehorster DC, Zemblys R, Beelders T, Holmqvist K. Characterizing gaze position signals and synthesizing noise during fixations in eye-tracking data. *Behav Res Methods*. 2020;52(6):2515–34. <https://doi.org/10.3758/s13428-020-01400-9> PMID: [32472501](#)
105. Ranson RE, Scarfe P, van Dam LCJ, Hibbard PB. Depth constancy and the absolute vergence anomaly. *Vision Res*. 2025;226:108501. <https://doi.org/10.1016/j.visres.2024.108501> PMID: [39488862](#)
106. Chopin A, Levi D, Knill D, Bavelier D. The absolute disparity anomaly and the mechanism of relative disparities. *J Vis*. 2016;16(8):2. <https://doi.org/10.1167/16.8.2> PMID: [27248566](#)
107. McKee SP, Welch L, Taylor DG, Bowne SF. Finding the common bond: stereoacuity and the other hyperacuties. *Vision Res*. 1990;30(6):879–91. [https://doi.org/10.1016/0042-6989\(90\)90056-q](https://doi.org/10.1016/0042-6989(90)90056-q) PMID: [2385928](#)
108. Westheimer G. Scaling of visual acuity measurements. *Arch Ophthalmol*. 1979;97(2):327–30. <https://doi.org/10.1001/archophth.1979.01020010173020> PMID: [550809](#)
109. Cottareau BR, McKee SP, Ales JM, Norcia AM. Disparity-specific spatial interactions: evidence from EEG source imaging. *J Neurosci*. 2012;32(3):826–40. <https://doi.org/10.1523/JNEUROSCI.2709-11.2012> PMID: [22262881](#)
110. Saunders JA, Knill DC. Humans use continuous visual feedback from the hand to control both the direction and distance of pointing movements. *Exp Brain Res*. 2005;162(4):458–73. <https://doi.org/10.1007/s00221-004-2064-1> PMID: [15754182](#)
111. Westheimer G. Cooperative neural processes involved in stereoscopic acuity. *Exp Brain Res*. 1979;36(3):585–97. <https://doi.org/10.1007/BF00238525> PMID: [477784](#)