MAGNITUDE ESTIMATION IN HUMANS

A Bayesian Approach to Characteristic Behavior in Path Integration

Frederike Hermi Petzschner

Dissertation at the

Graduate School of Systemic Neurosciences at the

Ludwig-Maximilians-Universität München

Dissertation at the

Graduate School of Systemic Neurosciences at the

Ludwig-Maximilians-Universität München

MAGNITUDE ESTIMATION IN HUMANS

A Bayesian Approach to Characteristic Behavior in Path Integration

Submitted by
Frederike Hermi Petzschner
Munich, 23nd of November 2012

Defense
Munich, 30st of January 2013
First Supervisor / Reviewer
Prof. Dr. Stefan Glasauer
Second Supervisor / Reviewer
Dr. Virgina Flanagin

Meiner Familie.

ABSTRACT

Anyone who has climbed a mountain before knows that the perceived distance walked depends on more than just its physical length. This intriguing relationship between physical and experienced magnitudes has fascinated researchers across various disciplines for more than 200 years. Part of the enthusiasm is driven by the fact that, although magnitudes, as well as the sensory organs with which we measure them, differ in so many ways, there are unifying principles in behavior common to all types of magnitudes estimated. In this thesis, the general characteristics of human magnitude estimation are studied in the case of visual path integration. The aim is to clarify the role of a-priori knowledge on the estimate of magnitude and to provide a unifying mathematical framework that explains the behavior. In particular, we investigated human linear and angular displacement estimation in different experimental situations with varying experience-dependent and abstract a-priori knowledge. We find systematic behavioral characteristics that are omnipresent in magnitude estimation studies, like the range effect, the regression effect or scalar variability. These characteristics are explained by a general model that combines a logarithmic scaling of magnitudes according to the Weber-Fechner law with the concept of Bayesian inference. The model incorporates a-priori knowledge about the stimulus and updates this knowledge on a trial-by-trial basis. The resulting iterative Bayesian estimation accounts for the aforementioned behavioral characteristics and provides a link between the two most well-known laws in psychophysics: the Weber-Fechner and Stevens' power-law. This work provides substantial evidence that magnitude estimation is not purely driven by sensation but underlies perceptual estimation processes that exploit and incorporate different types of information sources, in particular short-term prior experience. The proposed mathematical framework is likely applicable to magnitude estimation across different modalities and consequently contributes to a unifying account of the behavior.

OVERVIEW

The thesis is structured in four main chapters. In the introduction, I will review general findings in human magnitude estimation behavior, the psychophysical laws of Weber-Fechner and Stevens and characteristic deviations from these laws. I will relate these findings to a probabilistic understanding of perception as Bayesian inference.

The research is presented in form of two articles published in peer-reviewed journals, that are included in separate chapters. Each chapter starts with a short summary of the article's content, followed by the research article in its published format.

In the first article we aim to clarify the origins of characteristics in human visual path integration behavior. We test human linear and angular displacement estimation for varying underlying sample distributions and find general magnitude estimation behavior. This characteristic behavior is explained by an iterative Bayesian estimation model that incorporates short-term prior experience into the current estimate of displacement.

The second article addresses the question whether humans are also capable of incorporating abstract knowledge into their magnitude estimate. A symbolic verbal cue is introduced to the path integration paradigm containing additional information about the stimulus. The significant change in behavior is modeled with two iterative Bayesian estimation models. The models differ in their generative assumptions of how the different information sources (current sensory input, prior experience, and symbolic verbal cue) are combined into a single estimate of displacement.

The discussion relates the findings in human path integration to magnitude estimation in general. I will discuss the validity of our assumptions and how our findings can provide a general framework for modeling magnitude estimation.

CONTENTS

1 INTRODUCTION

Man is *not* the measure of all things.

Protagoras (slightly modified)

The safe in the International Bureau for Weights and Measures in Paris contains, protected by three bell jars, the International Prototype Kilogram, the base unit of mass[1]. It is one of the seven units of measure[2] defined by the International System of Units that determine the measurement of the physical quantities length, mass, time, electric current, temperature, luminous intensity, and mole. One reason why we need artificial units to measure these magnitudes is the discrepancy between physical magnitudes such as weight, luminance or distance and our corresponding subjectively experienced heaviness, brightness or displacement. The ringing of our cell phone, for example, is drowned out by a crowd of chatting people, while it is alarmingly loud when we are sitting in the audience of a lecture, although its true loudness remains the same. Our subjective experience of magnitude is not veridical. This is why we use watches to measure time, thermometers to measure temperature, scales to measure weight, and so forth. But where does the discrepancy come from?

To answer this question, one has to consider how information about physical quantities is mapped onto our senses and how the resulting sensation in turn leads to perception and behavior. Multiple domains of research, including philosophy, neuroscience, and psychology have focused on different aspects of this process.

In this thesis, I present a framework that explains human behav-

1 The International Prototype Kilogram might soon be outdated. It was originally built as provisional mass standard that has almost exactly the mass of one liter of water (water itself was inconvenient to use as a measure for trading). Consequently, the International Prototype Kilogram is the only remaining base unit that is based on a metallic artifact rather than a fundamental physical property that is easily reproducible. Consequently the General Conference on Weights and Measures agreed that the kilogram should be redefined based on fundamental constants in nature. The issue will be decided on the next conference in 2014.

2 These units of measure are meter, kilogram, second, ampere, kelvin, candela, and mole.

ior in a specific type of magnitude estimation[3], the estimation of linear and angular displacements from self-motion cues, termed path integration. The framework proposed is likely to provide a general account of magnitude estimation, as the behavioral characteristics explained are omnipresent across various domains, including loudness estimation, brightness estimation, or time estimation.

In the following, I will introduce three elements that are key to the approach: The *psychophysical laws* that describe the fundamental mathematical relationship between the experienced intensity of a magnitude and its physical size and therefore build the starting point for a model of magnitude estimation. The *characteristic effects* that demonstrate how the above relationship is dependent on contextual influences and which will build the motivation for a new extended modeling approach to magnitude estimation. Finally, *Bayesian inference* that can provide a general explanation for the characteristic effects and incorporate the psychophysical laws.

1.1 THE PSYCHOPHYSICAL LAWS

1.1.1 Weber-Fechner Law

From stimulus to subjective sensation. Envision science in the early 19th century when physics was already a rather mature science that used quantitative measures to relate phenomena in nature via well-defined mathematical relationships, later considered as the *laws of physics*. The idea that unifying principles in nature can be expressed in terms of mathematical relationships must have been, and still is, quite appealing to other sciences, including psychology. Thus, it does not come as a surprise that Gustav Theodor Fechner in 1860 attempted to establish a natural-science-inspired branch of psychology, called *psychophysics*. He was looking for a mathematical relationship between physical and psychological quantities, a *psychophysical law*. In his *Elements of Psychophysics*, Fechner states that *'psychophysics*

3 The term *magnitude estimation* in the context of classical psychophysics can have different meanings. If not explicitly stated, I will always refer to magnitude estimation as a general quantitative assessment of the relationship between stimulus magnitude and corresponding experienced sensation or perception independent of the particular response-method used (production-reproduction, cross-modality matching, etc.). Magnitude estimation can, however, also refer to a very specific method in classical psychophysics introduced by Stevens [1946, 1951], where subjects are asked to assign numbers in proportion to the magnitude of a given stimulus.

like physics should rest upon evidence and mathematical relations of empirical facts [...]' [Fechner, 1860][4]. He found a mathematical relationship between physical and psychological quantities in the work of Ernst Heinrich Weber [1850] who was one of the first to approach the study of the human response to a physical stimulus in a quantitative fashion. Weber asked subjects to discriminate between two stimulus magnitudes that differed only by a slight increment in order to determine their subjective discrimination threshold; a method that still is the gold standard to measure sensitivity in psychophysics. Weber found that this *just-noticeable difference* between two stimuli is approximately proportional to the absolute stimulus magnitude. The relation becomes intuitively accessible if we imagine lifting a weight of 10 g. If another 5 g were added on top of this one, the difference would immediately be noticeable. If, however, the first weight was 1 kg, a 5 g difference could not be detected anymore. The proportional relationship between the just-noticeable stimulus difference ΔI and the absolute stimulus magnitude I can be expressed by the following equation, later known as *Weber's law*,

$$\frac{\Delta I}{I} = k \tag{1.1}$$

where the constant factor k is referred to as the *Weber fraction* and depends on the type of stimulus tested.

In his search for a law of psychophysics, Fechner had the idea that Weber's just-noticeable difference between physical stimuli is proportional to a just-noticeable difference in subjective sensation[5] ΔS:

$$\Delta S \propto \frac{\Delta I}{I} \tag{1.2}$$

He thereby assumed that the factor k is constant and independent of the intensity I. Adding up just-noticeable differences across all stimulus magnitudes I in a specific sensation, formally achieved by an integration, yields Fechner's generalized psychophysical law: The psychophysical function that relates the subjective sensation S of magnitude to the corresponding physical magnitude I follows a logarithmic relationship:

$$S = c \cdot \ln \frac{I}{I_0} \tag{1.3}$$

Here, c and I_0 are constants, the latter is often referred to as the

4 Translation by the author, original quote: *'Als exakte Lehre hat die Psychophysik wie die Physik auf Erfahrung und mathematischer Verknüpfung erfahrungsmäßiger Tatsachen [...] zu fußen'* [Fechner, 1860]

5 Here, subjective sensation corresponds to an *'intermediary between sensory input and perception'* [Laming, 1999].

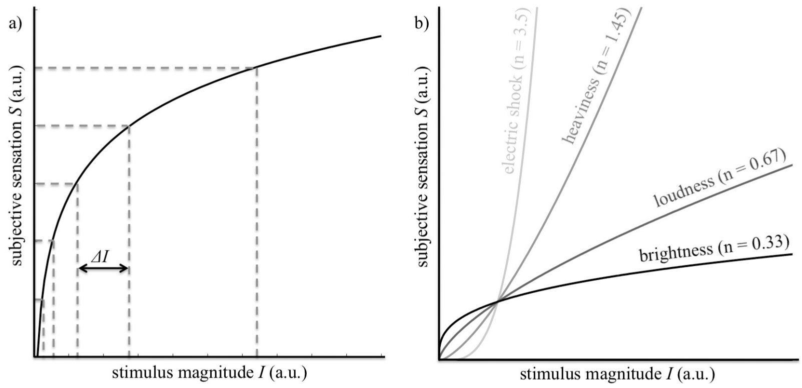

Figure 1: Relationship between stimulus magnitude and subjective sensation. a) Schematic plot of the logarithmic relationship between stimulus magnitude and reported subjective sensation as proposed by the Weber-Fechner law. ΔI corresponds to the just-noticeable difference. b) Schematic plot of the power-law relationship proposed by Stevens for four different perceptual continua (electric shock, heaviness, loudness, and brightness). The exponent n is modality-dependent. All plots are depicted on linear scales.

detection threshold of the respective stimulus magnitude. Equation 1.3, termed the *Weber-Fechner law*[6], states that a linear increase in stimulus magnitude causes a logarithmic increase in subjective sensation intensity (see figure 1a).

1.1.2 Stevens' Power-Law

Nearly one hundred years later, another psychophysicist, Stanley Smith Stevens, refuted Fechner's psychophysical law in his article *To honor Fechner and repeal his law* [Stevens, 1961]. Stevens suggested that a different fundamental relationship exists between subjective intensity of sensation and physical magnitude of a stimulus, namely a power function [Stevens, 1957]. He argued that Fechner's extension of Weber's law was too general and suggested that the subjective sensation is, in contrast to Fechner's original assumption, dependent on the Weber fraction [Stevens, 1960b]:

$$\frac{\Delta S}{S} \propto \frac{\Delta I}{I} \tag{1.4}$$

6 Fechner himself referred to this relation as *Weber's law*. But Weber was also not the first one to discover a logarithmic relationship. For a detailed discussion on this historical background please refer to Masin et al. [2009].

Consequently, this leads to a power-law dependence between subjective intensity of sensation and the physical magnitude of a stimulus, *Stevens' power-law* [Stevens, 1957]:

$$S = c \cdot I^{n} \tag{1.5}$$

Here, c refers to another constant and the exponent n depends on the respective modality tested (figure 1b). Cases where the fitted exponent is smaller than one ($n < 1$) are referred to as *compression*. Here, compression means that the increase in sensation intensity decreases with increasing stimulus magnitude [Stevens, 1960b]. In other words, subjects become less sensitive for larger stimulus magnitudes. A compression is observed for most sensory modalities (examples for brightness and loudness in figure 1b). For these modalities it is hard to distinguish between Stevens' power-law and Weber-Fechner's logarithmic law, because the shape of a power function with exponent smaller than one and the logarithmic function are hardly distinguishable. Stevens' power-law, however, also covers rare cases where the dependency between subjective sensation and stimulus is linear ($n = 1$) or reversed (*expansion*, $n > 1$). In the latter case, subjects become more sensitive with increasing stimulus magnitude. A vivid example of an expansion is the pain response to an electric shock (figure 1b).
Stevens' foray was based on numerous experiments across all sensory modalities that tested magnitude estimation based on the judgment of ratios [Aiba and Stevens, 1964; Stevens, 1966b, 1970; Stevens and Greenbaum, 1966; Stevens and Guirao, 1963, 1964; Stevens and Harris, 1962]. Subjects were, for instance, asked to assign a number to the magnitude of the perceived stimulus with respect to a previously experienced standard stimulus (*magnitude estimation*[7]) or to match the magnitude of the perceived stimulus with a magnitude in another modality, e.g. the brightness of a light source with the loudness of a tone (*cross-modality matching*). He found that psychophysical data acquired with these methods are best fit by a power function.

1.1.3 Comparison

Both Stevens and Fechner believed there is an accessible intermediate stage that reflects the intensity of a sensation before

7 Here, magnitude estimation refers to the specific method introduced by Stevens to measure the perceived intensity of a stimulus magnitude with a numerical response (see also footnote 3).

any higher cognitive processes take place [Laming, 1999]. More specifically, they both suggested that there is a scale transformation from the physical stimulus magnitude to this subjective sensation that can be determined via the response in a magnitude estimation task.

The important difference is that Stevens and Fechner had slightly different conceptions of *how* the response in magnitude estimation relates to the subjective sensation, that resulted in a logarithmic or a power-law. Fechner, on the one hand, assumed that there is no direct behavioral measure of subjective sensation. Thus, he proposed to measure just-noticeable increments in sensation. The sum of these increments then yields an indirect measure of subjective sensation. Stevens, on the other hand, believed that absolute magnitude estimation on a ratio scale already provides a direct measurement of subjective sensation[8] [Stevens, 1960b]. The different conceptualizations and concomitant methodological approaches make a direct comparison between the two psychophysical laws complicated [Laming, 1997]. It has been argued that one cannot discriminate between the two ideas of Stevens and Fechner without additional knowledge about the underlying mechanisms of magnitude estimation. Both laws might in fact be valid, but simply reflect different operational levels in the process of magnitude estimation. If our subjective sense of magnitude was logarithmic - including the subjective sense of numbers - and a numerical response method was used to determine our sensation intensity - as it is the case in Stevens' magnitude estimation and cross-modal matching - one would still find a power-law dependence between the physical magnitude and response of subjective sensation [MacKay, 1963] (for a detailed discussion, see section 4.2).

In this thesis, I present a combined understanding of Stevens and Weber-Fechner. By following up on these ideas it is proposed that magnitude estimation is not a pure measurement of sensation. Instead, the response reflects a perceptual estimate that involves different processing stages. These processing stages capture the different scale transformations proposed by Fechner and Stevens. The most striking evidence for such an estimation process comes from the characteristic effects in magnitude estimation, that will be reviewed in the following section.

8 Interestingly, Fechner somewhat anticipated Stevens approach when he noted that *'[...] a real measure of sensation would demand that we are able to call a given sensation twice, thrice or in general so-and-so many times as intense as another - but who could say such a thing?'* (Translation by Michell [2006], addendum by the author, original quote: *'[...] aber zum Maße der Empfindung gehörte, daß wir sagen könnten, diese Empfindung ist doppelt, dreimal, überhaupt so und so vielmal so stark als jene, und wer vermag dies bisher zu sagen.'* [Fechner, 1860]).

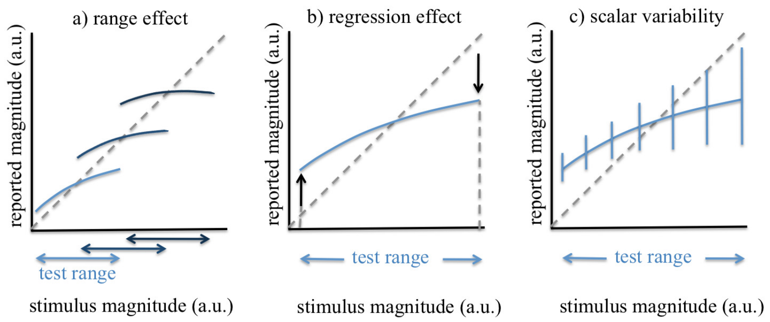

Figure 2: Schematic overview of selected effects in human magnitude estimation. a) Range effect: the slope of the relationship between reported magnitude and stimulus magnitude decreases with increasing stimulus range, leading to different power-law exponents dependent on the stimulus range. b) Regression effect: reported magnitudes show a tendency to be biased towards the center of the full range of stimulus magnitudes. As a consequence, large magnitudes are underestimated and small ones are overestimated. c) Scalar variability: the standard deviation of reported magnitudes increases linearly with the corresponding mean reported magnitude. All axes are plotted on linear scales.

1.2 CHARACTERISTIC EFFECTS IN MAGNITUDE ESTIMATION

From subjective sensation to behavior. When Stevens and colleagues empirically showed that the power-law holds for more than a dozen modalities (e.g. brightness, loudness, duration) and different testing methods, they originally interpreted the fitted exponents as an invariant measure of the rate of growth between perceived and physical magnitudes for each particular modality [Stevens, 1961, 1970]. This view was challenged when a number of studies discovered that the fitted exponent in a given modality is influenced by the observer, the history of stimuli and the response method used [Poulton, 1968; Robinson, 1976; Teghtsoonian and Teghtsoonian, 1971]. These determining factors were omnipresent across modalities and the variation in behavior associated with them became soon known as *biases*, or, as I will refer to them, *characteristic effects* in magnitude estimation. In the following, I will introduce the most common effects, the range effect, the regression effect, subject-specific, experience-dependent, and scalar variability.

1.2.1 Range Effect

When magnitude estimation behavior was compared across different experiments, one of the most striking findings was an almost perfect correlation between the range of stimuli tested and the reciprocal exponent of the power function for 24 different test modalities [Teghtsoonian and Teghtsoonian, 1971]. In figure 2a, this is depicted by a decreasing slope of the stimulus-response relationship for stimulus ranges that contain larger magnitudes [King, 1986]. As a consequence, Teghtsoonian and Teghtsoonian suggested that the exponent in Stevens' power-law is not a fixed entity but could reflect the subjective sensitivity range of observers [Teghtsoonian, 1973; Teghtsoonian and Teghtsoonian, 1978]. An alternative interpretation would be that the exponent in Stevens' power-law does not contain any information about the observers subjective sensation range, but purely reflects variability in the experimental stimulus range [Poulton, 1968].

1.2.2 Regression Effect

The *regression effect*, first termed the *central tendency of judgment* [Hollingworth, 1910], can be observed within a single test range. It describes the tendency of estimates of time, weight, force, area, or size of angle *'to gravitate toward a mean magnitude'* [Hollingworth, 1910], resulting in a systematic underestimation of magnitudes above the mean, an overestimation of magnitudes below the mean and a correct estimation of the mean magnitude in the tested distribution (see figure 2b). The strength of the regression varied for the same stimulus modality depending on the response methods used (e.g. using numbers instead of brightness to match loudness) [Stevens, 1960b]. Stevens consequently explained the regression effect by a tendency of the subjects to compress the range of whichever *'variable'* they have to adjust [Stevens, 1960a; Stevens and Greenbaum, 1966]. In other words, the range of responses is smaller than the original range of tested magnitudes. Stevens' description, as well as later attempts [Baird, 1997; Poulton, 1989] to categorize the regression effect, provided, however, no mechanistic explanation of *why* subjects would compress the stimulus range.

1.2.3 Subject-specific Variability

When the same subjects were tested in different modalities, such as estimating weight, taste, and numbers, the fitted Stevens' power-law exponent for each of these modalities was different.

However, the same subjects that showed a large exponent in one modality also showed larger exponents in the other modalities. That is, the variation in the fitted exponents correlated for individual subjects. To account for this subject-dependent component in the behavior, the authors proposed a modification of Stevens' power-law. They suggested adding a variable in the power-law exponent that takes an individual value for each subject [Jones and Marcus, 1961]. A similar correlation in the fitted power-law exponents for individual subjects was also found for the estimation of circle size, numerousness, and line length. Here, the authors simply ascribed their findings to the fact that subjects might consistently use the same number scale across modalities [Robinson, 1976; Rule, 1966]. Finally, Zwislocki suggested a method to cancel out the subject-dependent component to obtain *unbiased* Stevens' power-law exponents. He divided the fitted exponents obtained from one experiment (e.g. loudness) by the exponent obtained from the estimation of line length for each participant. The resulting reduction in inter-subject variability was interpreted as an indication that the original variation indeed stemmed from individual response characteristics [Gescheider, 1988; Zwislocki, 1983].

1.2.4 Experience-dependent Variability

Responses of the same subjects were also significantly correlated across consecutive experimental sessions for the same [Gescheider, 1988; Gescheider and Collins, 1989] as well as for different test modalities [Ekman et al., 1968; Jones and Woskow, 1962]. The same holds true for responses on subsequent trials [Cross, 1973; Garner, 1953]. Here, responses to stimuli that were presented after a large stimulus tended to be larger, whereas responses to stimuli with a small preceding stimulus tended to be smaller. This correlation was referred to as the *sequential effect*. This effect was suggested to be the basis of the stimulus regression effect, as trials with very high stimulus magnitudes are more likely preceded by a smaller stimulus magnitude in a fixed test range and vice versa. Thus, estimates should show a tendency to be biased towards the center of the full stimulus range [Cross, 1973]. Both experience-dependent correlations suggest that there might be a memory component in the behavior as previous stimuli influence the estimation of the current stimulus. There is also evidence that the session-to-session correlations are only influenced by short-term experience, as they are not persistent over long time scales [Teghtsoonian and Teghtsoonian, 1971].

1.2.5 Scalar Variability

Finally, the variability of a repeated response to the same stimulus is a function of the stimulus magnitude. As Weber noted, responses of subjects became more variable with increasing stimulus magnitude. This linear relationship between the standard deviation and mean of a given reported magnitude was either referred to as a constant *coefficient-of-variation*, or as *scalar variability* (see figure 2c) [Gibbon, 1977][9]. There are several possible explanations for this relationship: either subjects become increasingly more noisy in their response to larger magnitudes (i.e. there is a true scalar increase in variability), or magnitudes are represented on a logarithmic scale and this compression of the scale for larger magnitudes causes the increase in variability [Cantlon et al., 2009; Dehaene et al., 2008]. It is still unresolved which of the two explanations is correct (for a detailed discussion, see section 4.1.1).

1.2.6 Interpretation of the Effects

The discovery of the characteristic effects in magnitude estimation across various sensory modalities challenged the interpretation of the classical psychophysical laws.
On the one hand, Stevens and some of his colleagues adhered to the statement that the power-law holds and that the exponents are determined by the nature of the *'sensory tansducers'* of a given modality [Stevens, 1960b]. Consequently, variations in the exponent could be *'no more than the inevitable noise that characterizes these complex processes'* [Stevens, 1960b, 1966a].
On the the other hand, one of his former students, Poulton [1968], was one of the first to argue against this view. In a paper that he originally attempted to entitle *'The New Pseudophysics'*[10], he noted that *'there does not appear to be a stable relationship between an observer's use of numbers or similar measuring techniques and his perception of physical magnitudes. Insofar as Stevens has given the opposite impression in his published work, it has been by using trained selected observers and restricted experimental conditions.'* [Poulton, 1968]. Poulton believed that the effects in magnitude estimation are de facto properties of the process itself rather than artifacts caused by individuals. Consequently, he proposed that *'central*

9 Gibbon originally referred to this relationship as *scalar variance*. However, this term was misleading and *scalar variability* established instead.

10 The editor of Psychological Bulletin refused to accept this title, so that the final version was entitled *'The new psychophysics: six models for magnitude estimation'* [Poulton, 1981].

processes involving judgment cannot be understood entirely in terms of the behavior of the sense organs. The mechanisms of response learning and of response bias must be included in any adequate description.' [Poulton, 1968]. Poulton heralded the start of a new understanding of magnitude estimation that rejects pure sensation-driven explanations in favor of higher perceptual processes. These processes can only be understood if all of the characteristic effects are considered.

The framework presented in this thesis ties in with Poultons ideas. I suggest that the characteristics observed in magnitude estimation yield an insight to its underlying processes. In particular, that the behavior is dependent on the experience, the context, or the subject argues against a purely sensation-driven understanding of magnitude estimation. Instead, it seems like the behavior is also based on an incorporation of a-priori knowledge in the broadest sense. This a-priori knowledge needs to be combined with sensory information to yield an *estimate* of the magnitude. Such a process can be described in a probabilistic framework for perception that is rooted in Bayesian inference.

1.3 A BAYESIAN APPROACH TO PERCEPTION

1.3.1 Perception as Inference

From stimulus to perception. When Fechner started the study of absolute thresholds and just-noticeable differences, he assumed that if a sensation is sufficiently strong, then there is a transition from an unconscious to a conscious mental state [Rouder and Morey, 2009]. This view of the threshold as a discrete step function changed in the 1950's. The insight came from a combination of *statistical decision theory* [Grenander, 1950; Neyman and Pearson, 1933; Wald, 1939] and *communication systems theory* [Rice, 1944; Shannon, 1948]. Experiments in these fields proved that individuals' thresholds could vary to a large extent depending on the level of background noise, the strength of the stimulus or the subjective motivation. As a consequence, the detectability of a signal was no longer viewed as a direct reflection of sensitivity, but rather as the result of a decision process. This decision process is, on the one hand, corrupted by external and internal noise and, on the other hand, influenced by the personal decision criterion, e.g. responding very accurately or very quickly. Consequently, there is a smooth transition between stimuli that can be detected and stimuli that cannot. Furthermore, instead of an on-off switch there is a certain probability of detecting a sig-

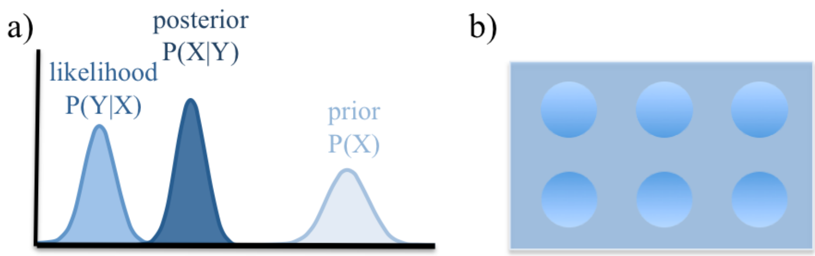

Figure 3: a) Combining prior and likelihood based on their reliability (width of the distribution) according to Bayes theorem. The prior represents the a-priori probability of a given state-of-the-world. The likelihood represents the probability of measuring the new observation (e.g. sensory input). The posterior is the probability of the state-of-the-world given the new observation. b) Example stimulus used to describe the interaction between prior and sensory input. Depending on the orientation the colored disc appears to be convex (upper row) or when flipped appears to be concave (lower row). The interpretation is caused by an a-priori assumption that light comes from above.

nal. The classical psychophysical methods, in contrast, did not discern between the true sensitivity of subjects and the personal decision criteria.
The first comprehensive mathematical analysis to explain how stimuli are detected in the presence of background noise was the *signal detection theory* [Green and Swets, 1966; Peterson et al., 1954]. The principal concept is simple: If we have to discriminate whether a signal is present or not in a noisy environment, and we respond that we detected the signal, there will be a certain rate of cases where this response is correct (when the signal was present, *hit rate*) or incorrect (when we detected a signal that was not present, *false alarm rate*). Signal detection theory states that our sensitivity is given by the difference between the probability of a hit and the probability of a false alarm. That way, signal detection theory allowed for the measurement of sensitivity in the presence of noise independent of the personal decision criterion.

Perception as unconscious inference. Uncertainty is present in all perceptual processes. Consider for instance the perception of color from reflected light. If our eyes received blue light reflected from a wall, then it is either because the wall is painted blue or it is white and illuminated by a blue light source. In most cases we can tell the difference. Somehow our perceptual system must have found a way to solve most of these *ill-posed* problems. One extremely influential idea of *how* this could be achieved came from Hermann von Helmholtz. He proposed that perception is more than a mere reflection of the sum of sensory inputs; it is a process that constructs a hypothesis about what-the-world-is-like based on our sensory inputs *and* our implicit assumptions. By incorporating these implicit assumptions, the brain is capable of drawing conclusions from the noisy and ambiguous observations of the world, a process that von Helmholtz termed *unconscious inference*. [von Helmholtz, 1867, 1878; Warren and Warren, 1986; Westheimer, 2008].

Let us assume we would like to make an inference of a given state-of-the-world X based on our sensory input Y. The simplest version of an inference would be to use some kind of look-up table where each sensation corresponds to a single state-of-the-world. By observing sensation Y, we could, thus, immediately determine the corresponding state X. Normally, a sensation does not, however, correspond to a single, but to a number of possible states X, of which some are more, and others less likely. Therefore, we would need to take all possible states and their respective probability into account which results in a so-called *conditional probability* of observing the sensory input Y, given state X: $P(Y|X)$. This is called the *likelihood*. To make an inference, we could take the most probable value of this likelihood function, its maximum, as our estimate of the state-of-the-world. This strategy is, therefore, known as *maximum likelihood estimation*[11] and is the simplest version of a *statistical* inference. There might be cases where it is unfavorable to rely on the sensory input alone. Consider our example with the blue wall: There are at least two possible explanations for the same sensory input (the wall appears blue, because it was painted blue or because blue light is reflected on it).

This can be resolved by another strategy which not only considers the likelihood of our observation, but also our a-priori assumptions before we made that observation, otherwise called our *prior* $P(X)$ (e.g. *'I have seen this wall a hundred times and it was always blue independent of the light.'*). This is where Bayes' theorem becomes important; it states how our prior and the new observation of the world, in this case the sensory input, can be combined in a statistically optimal manner to yield a so-called *pos-*

11 In the context of Bayes theory, the term *maximum likelihood estimation* is also often referred to as a Bayesian estimate with a uniform prior distribution. A prominent example in the literature is Ernst and Banks [2002].

terior probability[12] distribution $P(X|Y)$ [Colombo and Series, 2012; Doya et al., 2007]. That is:

$$P(X|Y) = \frac{P(Y|X) \cdot P(X)}{P(Y)} \tag{1.6}$$

where $P(Y)$ represents the probability of the new observation. The posterior probability is thus proportional to the product of both, the likelihood and the prior (see figure 3a)[13]. Consequently, one way to make an inference would now be to take the most probable value of this posterior function as an estimate of the state-of-the-world, which is called *maximum a-posteriori estimation*.

But this strategy also has a drawback. It does not provide any information about the shape of the posterior and thus about how reliable our estimate is. Imagine for instance if the function was very broad, various estimates would be almost equally likely. To remedy this, we could make a *Bayesian estimate* that takes the uncertainty of the posterior distribution[14] into account [Doya et al., 2007; Körding and Wolpert, 2004].

So far we have attempted to estimate a given state-of-the-world at a given point in time. To maintain an optimal estimate over time one would need a continuous update of belief. One version of such an online Bayesian inference is the *Kalman filter* [Kalman, 1960]. Analogous to the above description, it combines a-priori assumptions of the state-of-the world with the new incoming evidence based on the respective uncertainty to yield an improved a-posteriori measurement (*measurement update*). This posterior will then be used to predict the prior of the future as well as its uncertainty based on an assumption about the dynamics of the

---

12 Probability in the Bayesian sense may be distinguished from probability in the classical (*frequentist*) view. Here, likelihood, prior, and posterior probability are all *subjective probabilities* in that they describe a degree of belief rather than the true probability of an event to occur. I can be sure that an event is going to occur, that is my subjective probability is 100%, although its true probability might be less than 50%. Bayesian inference is, therefore, sometimes also circumscribed as belief update [Doya et al., 2007].

13 The Bayesian estimate takes the uncertainty of prior and likelihood into account. It can be regarded as an elegant, mathematical elaboration on von Helmholtz' view (for a detailed review on whether von Helmholtz was a true Bayesian, see Westheimer [2008]). When the likelihood and prior are Gaussian and their noise sources are statistically independent, the most likely posterior estimate is given by the weighted sum of the most likely value of the prior and the likelihood distribution (a special case used in this thesis). The respective weights are thereby inverse proportional to the uncertainty (the variances) of prior and likelihood.

14 It does so by using for instance the full posterior probability distribution as an estimate of the state-of-the-world, or a cost function that minimizes a specific type of error.

world (*time update*) [Berniker and Körding, 2011]. Every update of the measurement is thus followed by a time update, and vice versa (see Welch and Bishop [2002] for a detailed mathematical description of the Kalman filter).

1.3.2 Modeling Perception as Bayesian Inference

Using statistical inference to model perception. The previous section focussed on how an observer would achieve an optimal statistical inference of the world based on the sensory measurement and a-priori assumptions. Such an observer is called an *ideal observer*. Ideal observers can now be used to evaluate the behavior of *real* observers [Kersten and Mamassian, 2009].

In the case of Bayesian theories of perception and behavior, ideal observers are applied as a metric of performance. The basic idea is to use a simplified descriptive model of the relationship between the input and the state-of-the-world, called a *generative model*. Consequently the ideal observer does not necessarily show the best possible performance in a task, but one that is best according to a respective generative model[15]. The performance of a real observer can then be compared to the predicted ideal performance for a specific generative model[16] [Kersten and Mamassian, 2009].

A popular application of *Bayesian ideal observers* is in the context of ill-posed problems, where the nervous systems works in the presence of uncertainty due to noise or ambiguity [Poggio et al., 1985]. Examples are the perception of a 3D shape from a 2D image, visual illusions that allow more than one interpretation of the same image, or the combination of different sensory modalities [Ernst and Banks, 2002; Kersten et al., 2004; Knill and Richards, 1996; Landy et al., 1995; Maloney, 2002]. In such situations, sensory information alone often cannot resolve the ambiguity. However, if we assume additional information sources, e.g. in the form of prior experience or context, the problem becomes tractable. The following example illuminates this, consider figure 3b: The retinal illumination of the objects in the upper row is the same if it was concave and light came from be-

15 *Best* or *optimal* does not mean free of errors. A Bayesian estimate, given a cost function, reduces the overall uncertainty (see figure 3a). This does not, however, mean that it can not cause local systematic errors.

16 Another possibility to use ideal observers is to use them as benchmark of what can be maximally achieved in a task, given the full information provided in an experiment. When the real performance deviates from that of the ideal observer this is an indication for a suboptimal processing of the information provided (for applications, see Barlow [1962]).

low or it was convex and light came from above (the opposite is true for the lower row). However, we usually perceive the upper row as convex and the lower row as concave. The behavior can be explained by a generative model incorporating an a-priori assumption that light comes from above [Mamassian et al., 2002].

1.3.3 Magnitude Estimation as Bayesian Inference

Modeling effects in magnitude estimation. Apart from the obvious situations, where additional information is required to resolve ambiguity, the question arises whether a-priori assumptions are also incorporated in magnitude estimation. A few promising attempts have been made to apply models of Bayesian inference to behavior in human magnitude estimation. The great challenge is that most magnitudes are experienced across a relatively broad range. It thus seems likely that our expectations adapt to the natural continuum of the task (e.g. estimating a duration in seconds or in hours) rather than being fundamentally biased towards a single a-priori assumption (like *'light comes from above'*). A model of magnitude estimation would, therefore, need to account for the adaptive nature of magnitude estimation, e.g. by an incorporation of a-priori assumptions in the form of changing prior experience.
One of the first mathematical implementations of this idea was proposed by Laming [1999]. He suggested a model that explained the regression effect in cross-modality matching by the use of prior expectations for the current stimulus estimate [Laming, 1999]. These prior expectations were built based on magnitudes experienced in previous trails and then incorporated into the estimate in an statistical optimal manner. Lamings approach, although extremely promising, did not receive much attention[17]. Only a few years later, Bayesian models were successfully applied to a range of perceptual data, including force estimation, coincidence timing or motion perception [Körding et al., 2004; Miyazaki, 2005; Weiss et al., 2002]. Jazayeri and Shadlen [2010] tested time estimation for three different time ranges (short, intermediate, and long). They found that the behavior differed substantially depending on the underlying sample range, observing the regression effect, the range effect and scalar variability (see section 1.2). They showed that a Bayesian observer model that uses the mean of each underlying sample distribu-

17 To this day Lamings article has only been cited twice according the citation database Scopus, including the article presented in chapter 2.

tion as prior expectation can account for these behavioral characteristics.
Jazayeri and Shadlen's work reveals two interesting facts. First, characteristics in magnitude estimation behavior might be the result of an incorporation of prior experience. Second, a Bayes observer model can account for this behavior. What is missing, however, is how a-priori knowledge is acquired during the experiment, and whether the model can be generalized to other effects such as subject-dependent variability and trial-to-trial correlations. In this thesis, it will be shown that these issues can be solved by a Bayesian estimation model that includes the psychophysical laws and an online learning of a-priori knowledge.

1.4 THE CASE OF HUMAN PATH INTEGRATION

Path integration is the ability to estimate ones self-displacement solely based on self-motion cues such as proprioception or vision. A simple example for a path integration task is a *homing* experiment. Homing means, subjects are asked to return to their original *home* position after experiencing a certain displacement (e.g. by walking or turning blindfolded). Although these tasks appear relatively simple, subjects make systematic errors in path integration. That is, they systematically over- or underestimate certain displacements. Originally, these errors were ascribed to subjects' poor navigational abilities. Diverse models were developed to explain the misestimation of displacements by an accumulation of small errors in the integration of either space or time, termed as *leaky integration* [Glasauer et al., 2007; Lappe et al., 2011]. This leaky integration predicts behavior that is independent of experience or context. This is, however, not the case. Instead, if we directly compare two different studies on path integration then the same walked distance is estimated correctly in one study [Klatzky et al., 1990] or underestimated by 2m in the other [Schwartz, 1999]. The main difference between these studies was the test range of stimuli.
Consequently, an alternative explanation for the systematic errors exists; namely, that path integration is subject to the same characteristic effects (the regression or range effect) as many other types of magnitude estimation. If this is the case, then a model that attempts to explain these effects in path integration should provide a general account for the behavior in magnitude estimation.

1.5 AIM OF THIS THESIS

As we have seen, magnitude estimation exhibits a number of interesting features. On the one hand, there seems to be a fundamental mathematical relation between physical and perceived magnitudes, that is captured by psychophysical laws. On the other hand, this relation is strongly influenced by the context and the individual experience, resulting in systematic behavioral characteristics. These characteristic effects were at first not considered as fundamental features of magnitude estimation behavior itself. In some cases, such as human path integration, they are, therefore, still regarded as specific to the task at hand.

In this thesis, the focus is on these systematic behavioral characteristics observed in human path integration. First of all, it is investigated whether the systematic errors in path integration can be characterized as effects of human magnitude estimation behavior. If this is the case, a valid explanation for these characteristics should be based on a framework that can capture not only the behavior in path integration but generally in magnitude estimation. I propose that Bayesian inference is a promising candidate for such a framework. Evidence already exists that characteristic effects in magnitude estimation can, similar to effects in higher cognitive processes, be caused by an incorporation of a-priori assumptions. These assumptions are likely learned from experience or based on other contextual information sources. Bayesian inference could provide the mathematical basis to combine these a-priori assumption with the sensory representation of magnitudes.

The approach to these ideas was bipartite.

First, we assessed whether the systematic errors reported in human path integration can be understood as characteristic effects in human magnitude estimation and whether the use of a-priori knowledge can provide an explanation for this characteristic behavior. We therefore designed two studies on human path integration that aimed to change the a-priori assumptions of participants by either changing their immediate prior experience or by providing them with additional abstract information on the stimulus magnitude. The first study tested human visual linear and angular displacement estimation in a production-reproduction task for three different prior experience conditions by changing the underlying sample distributions. The second study introduced an additional symbolic cue, containing information about the upcoming stimulus, to the distance estimation task and tested how this additional information influenced path integration.

Second, a mathematical framework, based on Bayesian inference, is provided to explain the observed characteristics (regression effect, range effect, individual variability). In particular, three different mechanistic Bayesian estimation models were developed and tested. The models differ from previous accounts in that they are (1) based on Fechner's psychophysical law and in that they are (2) capable of learning and incorporating prior experience as well as additional abstract information into the current estimate of magnitude. This learning is modeled as a trial-by-trial update by a Kalman filter. It allows a flexible adaption to the present range of stimuli and accounts for sequential dependencies in the behavior.

The two studies and the corresponding Bayesian estimation models are presented in two sections. The influence of prior experience on human path integration and the developed basic Bayesian estimation model that explains the resulting regression effect, range effect and other types of variability are described in chapter 2. The influence of abstract contextual information on human path integration and two different extensions to the basic Bayesian estimation model that account for this context are described in chapter 3.

2 EXPERIENCE EXPLAINS RANGE AND REGRESSION EFFECT

> Experience is the name we give to our mistakes.
>
> Oscar Wilde.

2.1 SUMMARY

Experience matters. In the presented article, we tested the influence of changing prior experience on human visual path integration. We therefore tested human linear and angular displacement estimation in a self-motion production-reproduction task under three conditions ('short', 'intermediate', long'), that differ in the choice of the respective underlying sample distributions. We find that (1) subjects' mean displacement estimation behavior was biased towards the center of the underlying sample distribution (regression effect), (2) the amount of bias increased with increasingly higher sample ranges (range effect), and (3) the standard deviation for all conditions was linearly dependent on the mean reproduced displacements (scalar variability).
An iterative Bayesian estimation model on logarithmic scales is proposed that explains all observed behavioral characteristics by the fusion of an experience-dependent prior expectation with the current noisy displacement measurement. The iterative update is modeled by the formulation of a discrete Kalman filter and provides a model for sensorimotor learning without external feedback. The model yields a direct link between Weber-Fechner and Stevens' power-law and should consequently prove applicable to a wide range of magnitude estimation data.

2.2 REFERENCE

This work was carried out under the supervision of Stefan Glasauer;
F.H.P. and S.G. designed research; F.H.P. and S.G. performed

research; F.H.P. conducted the experiments; S.G. contributed unpublished analytic tools; F.H.P. analyzed data; F.H.P. and S.G. wrote the paper.

The paper was published in *The Journal of Neuroscience* under the following reference:
FH Petzschner and S Glasauer: Iterative Bayesian Estimation as an Explanation for Range and Regression Effects: A Study on Human Path Integration. *The Journal of Neuroscience*; 31(47):17220-17229, 2011.

Behavioral/Systems/Cognitive

Iterative Bayesian Estimation as an Explanation for Range and Regression Effects: A Study on Human Path Integration

Frederike H. Petzschner[1,2,3] **and Stefan Glasauer**[1,2,3]
[1]Institute for Clinical Neurosciences, Ludwig-Maximilians-Universität, 81377 Munich, Germany, [2]Bernstein Center for Computational Neuroscience, 82152 Planegg-Martinried, Germany, and [3]Graduate School of Systemic Neurosciences, Ludwig-Maximilians-Universität, 81377 Munich, Germany

Systematic errors in human path integration were previously associated with processing deficits in the integration of space and time. In the present work, we hypothesized that these errors are de facto the result of a system that aims to optimize its performance by incorporating knowledge about prior experience into the current estimate of displacement. We tested human linear and angular displacement estimation behavior in a production–reproduction task under three different prior experience conditions where samples were drawn from different overlapping sample distributions. We found that (1) behavior was biased toward the center of the underlying sample distribution, (2) the amount of bias increased with increasing sample range, and (3) the standard deviation for all conditions was linearly dependent on the mean reproduced displacements. We propose a model of Bayesian estimation on logarithmic scales that explains the observed behavior by optimal fusion of an experience-dependent prior expectation with the current noisy displacement measurement. The iterative update of prior experience is modeled by the formulation of a discrete Kalman filter. The model provides a direct link between Weber–Fechner and Stevens' power law, providing a mechanistic explanation for universal psychophysical effects in human magnitude estimation such as the regression to the mean and the range effect.

Introduction

Path integration, that is, the ability to keep track of changes in orientation and position using self-motion cues, constitutes an essential component of spatial navigation (Mittelstaedt and Mittelstaedt, 1980; Etienne and Jeffery, 2004). Yet human path integration performance exhibits systematic errors. Characteristic overestimation and underestimation of traveled distances and turning angles and thus a tendency to bias toward certain displacements have been reported for path integration tasks in real and virtual environments (Loomis et al., 1993; Jürgens et al., 1999; Riecke et al., 2002; Seemungal et al., 2007; Glasauer et al., 2009b). Furthermore, systematic errors differ between studies: while participants correctly reproduced a 10 m distance in one study (Klatzky et al., 1990), they underestimated the same distance by 2 m in another one (Schwartz, 1999). The main difference between the two studies was the range of distances tested. In the context of magnitude estimation, these systematic errors can be interpreted as regression and range effects (Stevens and Greenbaum, 1966; Teghtsoonian and Teghtsoonian, 1978).

One account of a bias in path integration posits processing deficits that accumulate during integration over space or time (Mittelstaedt and Glasauer, 1991; Fujita et al., 1993; Glasauer et al., 2007; Lappe et al., 2007; Mossio et al., 2008; Bergmann et al., 2011). However, research in related domains has shown that a bias is not necessarily a result of deficient processing, but can also represent the optimal solution of a system that incorporates prior knowledge about the world to maximize its use of information provided by sensory cues (Knill and Pouget, 2004; Burge et al., 2008; Fetsch et al., 2009). A probabilistic interpretation of this statement is the model of an optimal Bayesian estimator that combines a current noisy measurement with an a priori estimate that depends on previous experience (Körding et al., 2004; Miyazaki et al., 2005; Jazayeri and Shadlen, 2010) or reflects a general intrinsic tendency (Jürgens and Becker, 2006; Stocker and Simoncelli, 2006).

We hypothesize that such an estimation process could provide a potential explanation for systematic biases in human path integration. In particular, we speculate that an experience-dependent prior could cause the posterior estimate to adapt to the range of stimuli presented and show a regression toward the expectancy value of the underlying distribution. Thus, we tested human linear and angular displacement estimation separately in three different prior-experience conditions. In a virtual environment, participants were asked to produce and subsequently reproduce distances and turning angles that were drawn from three partially overlapping sample ranges. If participants incorporated knowledge about prior experience into their current estimate of displacement, their behavior should depend significantly on the underlying sample distribution. In a second step, we developed and tested two variants of a Bayesian estimator model where the reproduced displacement was determined by fusion of the current measurement and an a priori expectation. The prior was either modeled as a fixed value that approximated the statistical

Received April 20, 2011; revised Sept. 22, 2011; accepted Oct. 7, 2011.
Author contributions: F.H.P. and S.G. designed research; F.H.P. and S.G. performed research; S.G. contributed unpublished reagents/analytic tools; F.H.P. analyzed data; F.H.P. and S.G. wrote the paper.
This work was supported by the Federal Ministry of Education and Research Grant BCCN 01GQ0440. We thank Virginia Flanagin and Paul MacNeilage for helpful comments and copy-editing.
Correspondence should be addressed to Frederike H. Petzschner, Institute for Clinical Neurosciences, Ludwig-Maximilians-Universität, Marchioninistrasse 23, 81377 München, Germany. E-mail: fpetzschner@lrz.uni-muenchen.de.
DOI:10.1523/JNEUROSCI.2028-11.2011
 0270-6474/11/3117220-10$15.00/0

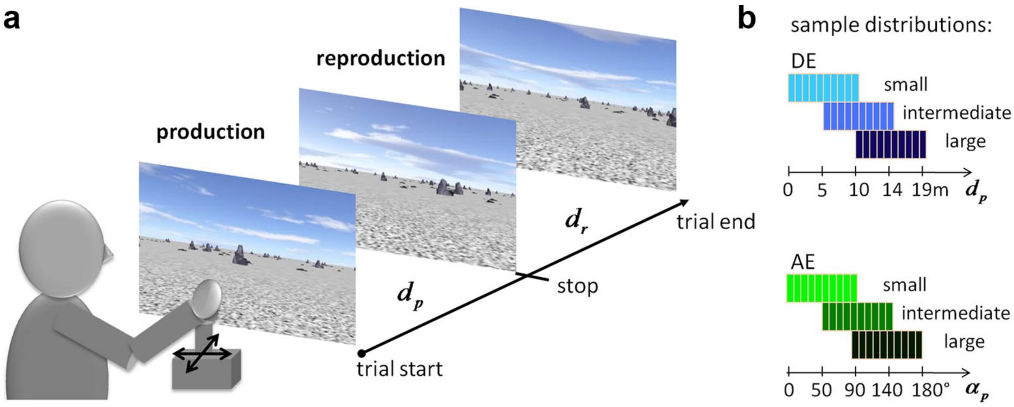

Figure 1. The production–reproduction task. ***a***, Temporal sequence of events in each test trial. Participants had to produce and subsequently reproduce a certain displacement in the VR by using the joystick. Depending on the session, participants walked on a linear path (DE) or turned on the spot (AE) until they were automatically stopped after a certain displacement, d_p or α_p, respectively. Next, participants were instructed to reproduce the same amount of perceived displacement while keeping the direction of movement constant. Participants indicated that they reached their final position by a button press. This estimated displacement is referred to as d_r in DE sessions and α_r in AE sessions. ***b***, DE and AE were tested separately under three different conditions that differed only in the underlying uniform sample distribution (small displacements, intermediate displacements, and large displacements range, for DE: turquoise, light blue, and dark blue, respectively; for AE: bright green, olive, dark green, respectively) from which the production displacements d_p and α_p were drawn.

properties of the underlying sample distribution or as an iterative estimate updated in each trial, which represented the immediate prior experience.

Materials and Methods

Participants

Fourteen volunteers (seven female), aged 22–34 years, were monetarily compensated for their participation in the study. All had normal or corrected-to-normal vision and were naive to the purpose of the experiments. The experiments were approved by the local ethics committee in accordance with Declaration of Helsinki.

Experimental setup

Stimuli were presented binocular on a computer monitor (resolution, 1280 × 800; frame rate, 59 Hz) driven by an ATI Mobility Radeon HD 3400 graphics card. Experiments were conducted in complete darkness except for the illumination by the monitor. The real-time virtual reality (VR) was created using Vizard 3.0 (Worldviz) and depicted an artificial stone desert consisting of a textured ground plane, 200 scattered stones, and a textured sky (Fig. 1*a*). The orientation of the ground plane texture, the position of the stones, and the starting position of the participant within the VR were randomized in each trial to prevent participants from using any of these as potential cues. The sky was simulated as a 3D dome centered on the participant's current position so that the distance to the horizon was kept constant. The eye height in the VR was adjusted individually to the true eye height of each participant (Daum and Hecht, 2009). Participants used a multidirectional movable joystick (SPEEDLINK) to navigate.

Experimental procedure

The estimation of traveled distances and the estimation of turning angles were tested separately under three different conditions in a production–reproduction task (Fig. 1*a*).

Distance estimation experiment. Each trial started with an instruction for participants to move forward along a linear path while keeping track of their self-displacement. Direction of movement during production was indicated by a visual cue at the horizon. When participants reached the sample distance d_p, movement was automatically stopped and disabled for a few seconds. Subsequently, participants were instructed to reproduce the perceived distance and indicate their final position via button press. In all trials, velocity was kept constant during movement, but changed randomly up to ±60% (scaling factor drawn from a normal distribution) between production and reproduction phases to exclude time estimation strategies to solve the task. To test the effect of prior experience only, the settings for the three conditions were the same except that the sample distances and respective turning angles were drawn from three different underlying uniform sample distributions, specified as small displacements (d_p = [1, 2, 3, 4, 5, 6, 7, 8, 9, 10] m), intermediate displacements (d_p = [5, 6, 7, 8, 9, 10, 11, 12, 13, 14] m) and large displacements range (d_p = [10, 11, 12, 13, 14, 15, 16, 17, 18, 19] m). The sample distributions of the three conditions were chosen to be partially overlapping to test whether displacement estimation behavior differed significantly for the same sample stimulus depending on the previously experienced displacements (Fig. 1*b*). Participants had no knowledge about the amount of displacement they had to reach during the production phase and were naive to the condition in which they were tested.

Turning angle estimation experiment. The stimulus and settings in the angle estimation (AE) experiment were identical to the distance estimation (DE) experiment, with the following exception: participants turned on the spot to a previously indicated direction. Turning direction was kept constant between production and reproduction to preclude the use of external cues to solve the task. The sample turning angles, α_p, for the three prior experience conditions were in analogy drawn from three different sample distributions specified as small displacements (α_p = [10, 20, 30, 40, 50, 60, 70, 80, 90, 100]°), intermediate displacements (α_p = [50, 60, 70, 80, 90, 100, 110, 120, 130, 140]°) and large displacements range (α_p = [90, 100, 110, 120, 130, 140, 150, 160, 170, 180]°).

All participants performed both types of experiment. The three conditions for DE and AE were tested in separate sessions, resulting in six test sessions per participant. Each session lasted between 45 and 60 min and was composed of 200 trials. The first 20 training trials per experimental condition served to familiarize participants with the VR. Feedback on the performance was given after the reproduction by displaying an object in the VR at the correct distance or turning angle and asking subjects to navigate toward this location. In the following 180 test trials, no feedback was given. Only test trials were used for data analysis. Two sessions of the same experiment type, AE or DE, were separated by at least 1 h and up to a few days. Within sessions, participants had a short break of 100 s after 100 and 150 trials. Each sample displacement was repeated 20 times per condition in randomized order. The same trial order within one condition was maintained for all participants. The order in which the three conditions for DE and AE were tested was randomized for each participant.

Analysis of behavioral data

Position and orientation of participants within the VR were sampled at 20 Hz. The displacement between the end of the production phase and time of the button press was calculated as the reproduced displacement d_r and α_r. The estimation error was defined as the difference between the reproduced and produced displacement. Data analysis was conducted in MATLAB R2010b (MathWorks). Statistical differences were assessed using repeated-measures ANOVA (rm-ANOVA). One rm-ANOVA with

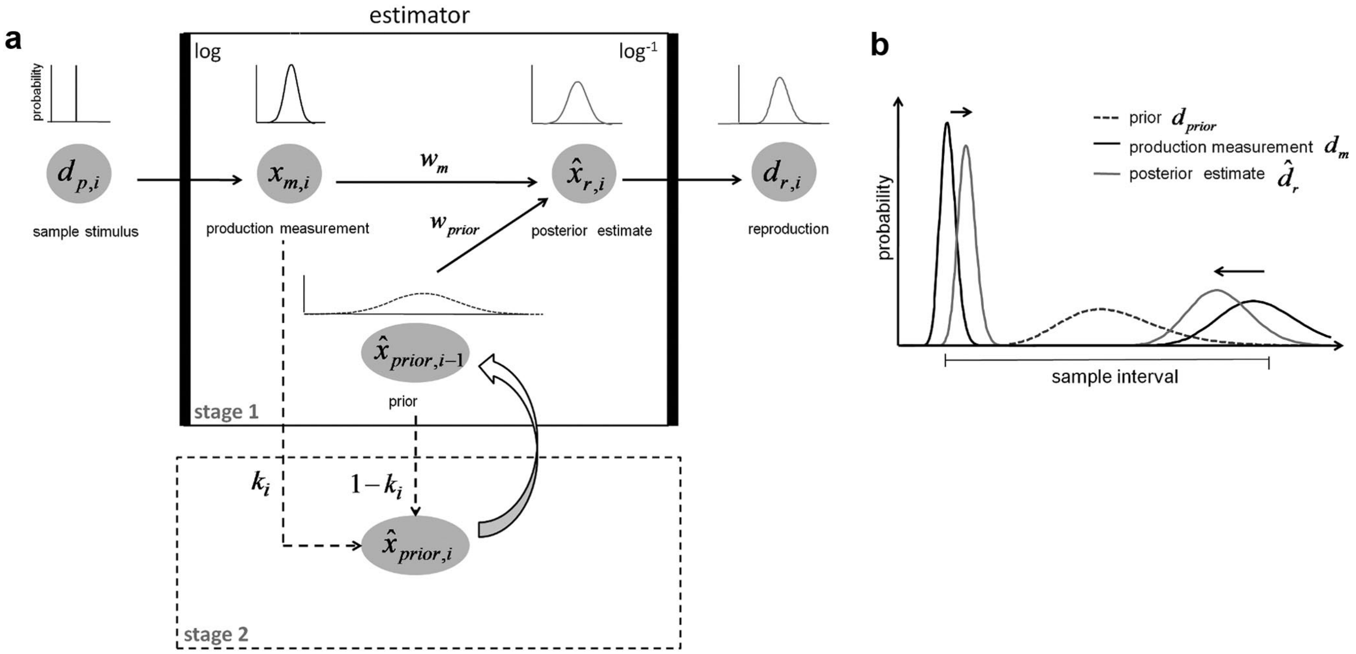

Figure 2. Two-stage Bayesian estimator model on logarithmic scales. ***a***, Schematic estimation process. Stage 1, The produced displacement $d_{p,i}$ in trial i is represented by the Bayesian estimator as a measurement likelihood on logarithmic scales $p(x_{m,i})$. The posterior estimate of displacement is determined by the weighted average of the measurement $x_{m,i}$ and an a priori estimate of displacement $\hat{x}_{prior,i-1}$, with weights of measurement w_m and prior w_{prior}, resulting after backtransform in a reproduced displacement on linear scales $d_{r,i}$. Stage 2, The posterior estimate of the prior in trial i, $\hat{x}_{prior,i}$, is estimated before the next trial according to the weighted average of the a priori estimate $\hat{x}_{prior,i-1}$ and the measurement $x_{m,i}$, with weights k_i and $1 - k_i$, modeled by the discrete formulation of the Kalman filter, where k_i refers to the Kalman gain. The posterior is updated over time to build the prior estimate $\hat{x}_{prior,i-1}$ in the subsequent trial. ***b***, Example for the effect of the Bayesian estimator model on linear scales for a particular sample interval. If the prior d_{prior} is close to the mean of the sample interval, Bayesian fusion with the measurement d_m leads to a posterior estimate d_r that overestimates small displacements and underestimates large displacements. The effect is stronger for larger displacements (black arrows), due to the calculation of the weighted average on logarithmic scales and thus an increased standard deviation with increasing mean on linear scales.

main factors for the overall condition (small displacements, intermediate displacements, large displacements) and displacement ([1:10], [5:14], [10:19] for DE session and [10:100], [50:140], [90:180] for AE sessions) was performed on the signed estimation error to reveal differences in the error magnitude and shape of the curve.

To assess range effects, we also looked for differences in reproduced displacements of samples that were presented in more than one condition (i.e., overlapping samples) using a second rm-ANOVA referred to as duplicated samples comparison with main factor condition (Higher Range vs Lower Range). Thereby, overlapping samples were compared in a single rm-ANOVA. For the factor Higher Range, we used [5:9] m from the intermediate displacements condition and [10:14] m from the large displacements condition, compared with the factor Lower Range including [5:9] m from the smaller displacements condition and [10:14] from the intermediate displacements condition. Note that each measurement was only used once, either for the factor Higher Range or Lower Range. For angular displacements, the factor Higher Range included [50:90]° from the intermediate displacements condition and [100:140]° from the large displacements condition; the Lower Range factor included [50:90]° from the small displacements condition and [100:140]° from the intermediate displacements condition. Linear regression analyses were performed to quantify the relationship between mean and standard deviation. A probability level of $p < 0.05$ was considered significant for all statistical analysis.

Bayesian estimator model

The stimulus displacements for the production phase of the three conditions were entered into a Bayesian estimator model in the same order as in the experiment (Fig. 2). The model assumes Bayesian fusion of measurement and prior experience on logarithmic scales to achieve a final displacement estimate. The single computational steps are as follows.

Logarithmic internal representation of displacement. Weber–Fechner's law proposes the representation of a stimulus size on a logarithmic scale (Fechner, 1860). Several recent psychophysical studies support the notion that human behavior approximately follows this law for numerical quantities (Dehaene et al., 2008), visual motion perception (Zanker, 1995; Jürgens and Becker, 2006; Stocker and Simoncelli, 2006), and locomotor path integration (Durgin et al., 2009). Accordingly, we introduced a modified logarithmic representation of perceived linear or angular displacements, similar to the one previously proposed for motion perception (Stocker and Simoncelli, 2006):

$$x_m = \ln\left(1 + \frac{d_m}{d_0}\right) + n_m, \tag{1}$$

where d_m is the measured displacement on linear scales and x_m is the internal noisy logarithmic representation of the measured displacement. The random variable n_m represents the normally distributed measurement noise $p(n_m) \approx N(0, \sigma_m^2)$. The input stimuli are expressed in virtual meters or degrees. $d_0 \ll 1$ is a small normalization constant, which leads to a unitless internal representation of displacement. For the simulations, we chose an arbitrary fixed value of $d_0 = 0.01$ m for distance estimation and $d_0 = 0.01°$ for angle estimation. The addition of 1 allows for representing a null displacement and may account for the deviation of the Weber–Fechner law at small magnitudes.

Since all represented displacements (d_m/d_0) in our experiment are large compared with 1, we reduced the general description of the transformation in Equation 1 to the simpler form:

$$x_m = \ln\left(\frac{d_m}{d_0}\right) + n_m. \tag{2}$$

Note that d always indicates displacements on linear scale, whereas x refers to the mean of the internal distributions (Fig. 2*a*). Since the distribution of the measurement noise is known, the measured displacement can internally be represented by the likelihood distribution, a Gaussian distribution with $p(x_m) \approx N(x_m, \sigma_m^2)$.

Bayesian fusion of measurement and prior. The probability of having experienced a certain displacement is given by the posterior probability distribution, which depends on the likelihood of measurement or evidence and the prior probability. Assuming that the likelihood functions of prior and measurement are approximately Gaussian, the mean of the posterior distribution on logarithmic scales $\hat{x}_r$ is, according to Bayes' rule, given by a weighted sum of the mean of the prior distribution x_{prior} and the measurement likelihood x_m.

$$\hat{x}_r = w_{prior} \cdot x_{prior} + w_m \cdot x_m \tag{3}$$

with variance

$$\hat{\sigma}_r^2 = \frac{\sigma_m^2 \cdot \sigma_{prior}^2}{\sigma_m^2 + \sigma_{prior}^2}. \tag{4}$$

The weights w_{prior} and w_m add up to unity and depend on the uncertainty of the measurement and prior, measured by the inverse variance of prior and measurement distributions:

$$w_m = 1 - w_{prior} = \frac{1/\sigma_m^2}{1/\sigma_m^2 + 1/\sigma_{prior}^2}. \tag{5}$$

In the proposed model, Bayesian fusion takes place on logarithmic scales (Eqs. 1, 2), thus the reproduced distance on linear scales is determined from the back-transformation of the Gaussian distribution $p(x_r) = N(\hat{x}_r, \sigma_r^2)$, resulting in a lognormal distribution on linear scales.

Up to this point, the model specifies a posterior probability distribution of distances rather than the particular distance that should be reproduced. To determine the distance to be reproduced and thus to execute a specific action the peak, the mean or any specific value of the posterior distribution could be selected, depending on the cost associated with making different types of errors (Doya et al., 2007). Commonly proposed symmetric cost functions (Körding and Wolpert, 2004b) lead to reproduction of one of the location parameters mean, mode, or median of the estimated posterior distribution.

However, in contrast to a normal distribution for the resulting lognormal distribution, these location parameters are no longer equal. The median $\tilde{d}_r$, mean $\bar{d}_r$, and mode $d_{r\,max}$ of the distribution are given by

$$\begin{aligned} \tilde{d}_r &= e^{\hat{x}_r} \cdot d_0 \\ \bar{d}_r &= e^{\hat{x}_r + \hat{\sigma}_r^2/2} \cdot d_0 \\ d_{r\,max} &= e^{\hat{x}_r - \hat{\sigma}_r^2} \cdot d_0, \end{aligned} \tag{6}$$

and thus differ by a shift, which depends on the stimulus distance on linear scales. The variance $\sigma_{d_r}^2$ is given by

$$\sigma_{d_r}^2 = \bar{d}_r^2 \cdot (e^{\hat{\sigma}_r^2} - 1). \tag{7}$$

To account for these differences in the reproduction estimate depending on the cost function, we introduced a shift term, Δx, as additional parameter in the model[a] (see Model fit), so that the reproduced displacement is given as

$$d_r = e^{\hat{x}_r + \Delta x} \cdot d_0. \tag{8}$$

Note that with Equations 3 and 8, the reproduced displacement becomes

$$d_r = e^{\Delta x} \cdot \hat{d}_{prior}^{\,w_{prior}} \cdot \hat{d}_m^{\,w_m}, \tag{9}$$

and thus follows Stevens' power law (Stevens, 1961).

Finally, to account for signal-independent variability of the reproduced displacement caused, for example, by reaction times in handling the response device, the random variable n_c representing normally distributed constant noise $p(n_c) \approx N(0, \sigma_c^2)$ was added to the reproduced displacement on linear scales d_r.

Prior update. In the current study, the Bayesian estimator was tested with two methods to implement an experience-dependent distance prior, referred to as one-stage and two-stage model. In the one-stage model, the prior was implemented as a distribution with a fixed mean centered at the mean of each underlying sample distribution and thereby represented the global statistics of the input stimuli. In the two-stage model, the displacement prior was updated iteratively in an additional computation step dependent on the posterior estimate of the prior in the previous trial and the current measurement of displacement (Fig. 2*a*). The update in each measurement step is modeled by the discrete formulation of the Kalman filter for a 1D first-order system. The state to be estimated and the current measurement at update step *i*, corresponding to trial *i*, are modeled by

$$\begin{aligned} x_{prior,\,i} &= x_{prior,\,i-1} + n_q \\ x_{m,\,i} &= x_{prior,\,i} + n_r. \end{aligned} \tag{10}$$

The random variables n_q and n_r represent the process and measurement noise, respectively. They are assumed to be independent with approximately normal probability distributions $p(n_q) \propto N(0, q)$ and $p(n_r) \propto N(0, r)$. The system defined by Equation 10 thus states that (1) the prior has no intrinsic dynamics and is varying only due to random changes modeled by n_q, and (2) the current measurement is an instantiation of the current prior perturbed by the measurement noise n_r.

For this simple system, the difference equation system of the Kalman filter reduces to

$$\begin{aligned} k_i &= \frac{p_{i-1} + q}{p_{i-1} + q + r} \\ p_i &= k_i \cdot r \\ \hat{x}_{prior,\,i} &= (1 - k_i) \cdot \hat{x}_{prior,\,i-1} + k_i \cdot x_{m,\,i}, \end{aligned} \tag{11}$$

with k_i being the Kalman gain, $\hat{x}_{prior,i-1}$ and $\hat{x}_{prior,i}$ being the a priori and a posteriori estimate of the distance prior at update step *i*, and p_i and p_{i-1} the corresponding variance. Note that it is evident from this equation that the Kalman gain k_i can be interpreted as weight of the measurement depending on measurement noise and the assumed random change of the distance prior. The new estimate of the distance prior is thus a weighted sum of the previous estimate and the current measurement.

In the context of the Bayesian estimator model, we refer to p_i as the estimated variance of the distance prior σ_{prior}^2 and *r* as the measurement variance σ_m^2. The prior for the two-stage model was initialized by the first measurement and reset at the beginning of each new session to account for the lack of prior knowledge of the underlying distribution except for the training trials. Note that after a measurement has been taken, the entire model is deterministic and does not involve any random elements to determine the distance to be reproduced. A preliminary version of the model has been published in abstract form (Glasauer et al., 2009a).

Model fit. The displacements used in the experiment were used in the same order as input d_m for both the one-stage and the two-stage models. The shift term Δx was implemented in both models using Equation 8. In the one-stage model, the single estimate of the prior x_{prior} was modeled as the log-transformed mean value of each underlying sample distribution on linear scales. The weighting of prior w_{prior} and Δx were determined by minimizing the sum of the squares of the residuals of the one-stage estimator model and the individual participants' mean responses for all three conditions simultaneously using the Matlab procedure *lsqnonlin*.

In the two-stage model, the prior x_{prior} was modeled as a continuously varying value determined by the Kalman filter (Eq. 11), which was reset at the beginning of each condition. To quantify the time course of the Kalman gain k_i, which approaches a steady-state value, its time constant τ expressed in trials was determined by fitting an exponential function to k_i. The ratio between measurement and process noise in the Kalman filter *r*/*q* and the shift term Δx were determined by minimizing the sum of the squares of the residuals of the two-stage estimator model and the individual participants' mean responses for all three conditions simultaneously using the Matlab procedure *lsqnonlin*. To be comparable to the one-stage model, the steady-state weighting of the prior w_{prior}, which is

[a] If the displacement to be reproduced was estimated already on logarithmic scales, then the mode, mean, and median of the posterior distribution would be equal and, for commonly used symmetric cost functions, the statistically optimal estimate would be the median of the log-normal distribution. We tested for this possibility and found that the model accounts well for the behavior of subjects when the shift parameter was not significantly different from zero (see Results), but generated worse fits for the remaining participants.

proportional to the ratio r/q, was determined from the result of the fitting procedure and is reported in the Results, below.

Thus, both the one-stage and the two-stage models are each fully determined by two free parameters. However, the one-stage model requires additional input about the prior in each condition, whereas the two-stage model does not. To assess the precision of the fitted parameters, we estimated 95% confidence intervals ($CI_{95\%}$) of all parameters, which were determined from the Jacobian of the parameter surface at the minimum using the Matlab procedure *nlparci*. The coefficient of determination R^2 was estimated to assess the proportion of variability in the mean data that is accounted for by the respective model. To test for a significant difference in the R^2 for individual participants between the two models, we used the Matlab procedure *signtest*.

For both models, the variance of the reproduced displacement $\hat{\sigma}_r^2$ was determined separately from the slope of the linear regression between standard deviation and mean of the reproduced displacements, using Equation 7. The *y*-intercept of the regression was interpreted as being due to constant noise $p(n_c) \approx N(0, \sigma_c)$ (see Bayesian fusion of measurement and prior, above).

Predictions for Bayesian estimation. The proposed Bayesian framework makes specific predictions on the behavior of an optimal estimator that can be tested experimentally. First, assuming independent noise sources for prior experience and measurement, the estimate on logarithmic scales is determined by the weighted average of prior and measurement dependent on the respective reliability. This leads to a power law dependence between input stimulus and reproduced displacement (Eq. 9) in linear space, as proposed by Stevens (1957), where the power function is determined by the individual weighting of the subjects. Second, according to this relationship, the difference between prior and measurement and therefore the effect of the Bayesian fusion becomes more pronounced for larger displacements, meaning that the overestimation and underestimation increase for increasing displacements. This results in a behavior known in the psychophysics literature as range effect (Teghtsoonian and Teghtsoonian, 1978). Third, assuming constant Gaussian noise on logarithmic scales leads to a linear dependence of the mean reproduced displacement and its corresponding standard deviation on linear scales (Eq. 7).

Results

Experience-dependent behavior

Participants' responses show three major characteristics that can be attributed to an estimation process that incorporates knowledge about the underlying sample distribution. These characteristics were tested at the group and single-subject levels.

First, reproduced distances and turning angles exhibited a clear tendency toward the mean of the underlying sample distribution for each of the three sample distributions tested. In each condition, small distances and angles were overestimated and large distances and angles were underestimated (Fig. 3). This can be seen in the overlapping distances ([5:14] m) and angles ([50: 140]°) that were tested in more than one condition for which the duplicated samples comparison reveals a significant difference between conditions (main effect: Higher Range, Lower Range; DE: $F_{(1,13)} = 20.4$, $p < 0.001$; AE: $F_{(1,13)} = 66.6$, $p < 0.001$). Furthermore, the duplicated samples comparison on the single-subject level reveals that 11 of 14 participants in DE sessions and all participants in AE sessions showed a significant dependence of estimation magnitude on the underlying sample distribution (main effect: Higher Range, Lower Range; DE: $F_{(1,13)} = 9.8$–202.0, $p < 0.01$; AE: $F_{(1,13)} = 12.4$–217.0, $p < 0.01$).

Second, the bias toward the center for each sample distribution increased with increasing displacement range. The overestimation and underestimation errors were more pronounced for the conditions with larger displacements (interaction: condition × displacement; DE: $F_{(18,234)} = 4.4$, $p < 0.001$; AE: $F_{(18,234)} = 3.2$, $p < 0.001$). This causes a decrease in the slope between produced and reproduced displacement for increasing sample range. The significant change in the bias, measured by the change in estimation error over the conditions, was found for 13 of 14 participants in DE and AE sessions (interaction: condition × displacement; DE: $F_{(18,234)} = 4.7$–38.3, $p < 0.05$; AE: $F_{(18,234)} = 10.2$–120.1, $p < 0.001$).

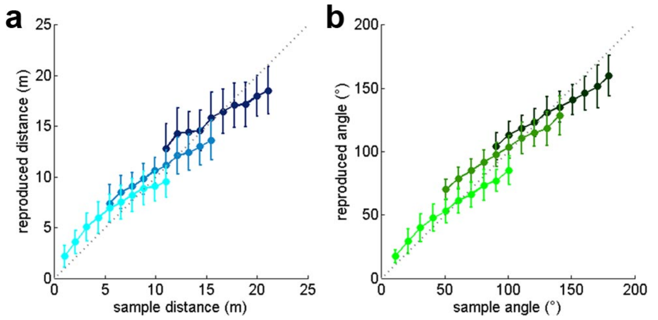

Figure 3. Mean displacement estimation behavior over all participants for the three prior experience conditions (small displacements, intermediate displacements, and large displacements range). ***a***, Mean participants' response in DE sessions in virtual meters (blues). ***b***, Mean participants' response in AE sessions in degrees (greens). Error bars depict the standard deviation between participants' responses. The dotted line indicates were the response and sample stimulus would be equal. In both experiments, the behavior deviated significantly from the line of equality depending on the prior experience condition. Small displacements were underestimated and large displacements were overestimated in all conditions. The bias increased for increasing sample range, being strongest for the large displacements range (dark colors). Small displacements, intermediate displacements, and large displacements range, for DE: turquoise, light blue, and dark blue, respectively; for AE: bright green, olive, dark green, respectively.

Third, the standard deviation of the reproduced displacements was dependent on the sample distribution. The duplicated samples comparison revealed that standard deviations of overlapping samples differed significantly depending on the underlying sample distribution (main effect: Higher Range, Lower Range; DE: $F_{(1,13)} = 11.1$, $p = 0.005$; AE: $F_{(1,13)} = 7.8$ $p = 0.01$). Additionally, we observed a strong correlation between the mean reproduced displacement and the corresponding mean standard deviation for both DE and AE sessions. (DE: linear regression: $r = 0.95$, $p < 0.001$; AE: $r = 0.97$, $p < 0.001$; Fig. 4*a*,*c*). On the single-subject level, the linear regression between standard deviation and mean of reproduced distances yielded a highly significant correlation coefficient *r* for all participants (DE: $p < 0.001$ for 11 of 14 participants, $p < 0.01$ for the remaining three participants, AE: $p < 0.001$ for 11 of 14, $p < 0.01$ for the remaining three participants).

Test of the Bayesian estimator model

The experimental findings support the notion that humans incorporate knowledge about the stimulus properties applied in the current condition into their measurement of displacement and that this behavior is qualitatively in agreement with a Bayesian estimation process.

To evaluate this finding in a quantitative manner, we fit two variants of the Bayesian estimation model to the mean response over all participants and to the individual participants' mean responses using a least-squares fitting method (Fig. 2*a*). The first variant, referred to as one-stage model, tests a fixed prior (for a similar study, see Jazayeri and Shadlen, 2010), that is determined by the mean of each sample distribution and therefore represents prior knowledge that captures the overall statistics of the experiment (model fit group: $w_{\text{prior,DE}} = 0.40$, $CI_{95\%} = [0.33, 0.48]$; $\Delta x_{\text{DE}} = 0.03$, $CI_{95\%} = [0.01, 0.04]$; $w_{\text{prior,AE}} = 0.40$, $CI_{95\%} = [0.36, 0.43]$; $\Delta x_{\text{AE}} = 0.01$, $CI_{95\%} = [0, 0.01]$; individual participants: $\bar{w}_{\text{prior,DE}} = 0.41 \pm 0.14$, range $= [0.20 - 0.61]$; $\Delta\bar{x}_{\text{DE}} =$

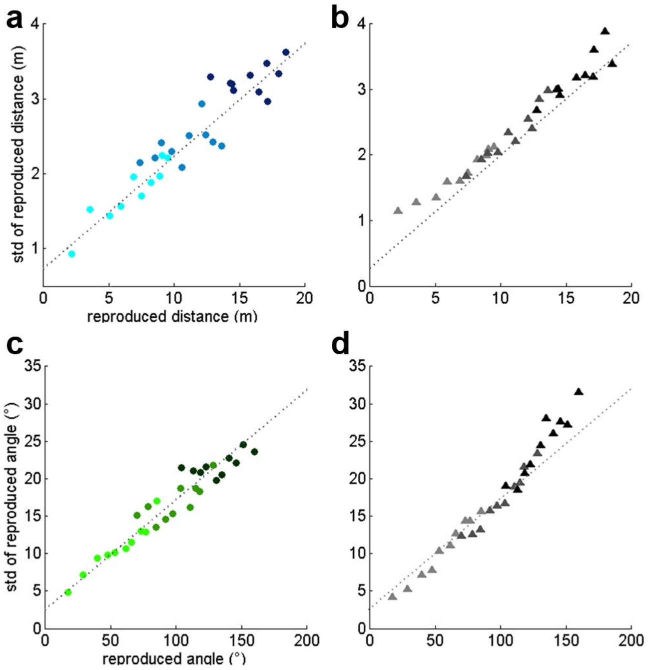

Figure 4. Mean predicted and actual reproduced displacement estimation and corresponding mean standard deviation (std). ***a***, Experimental DE data of mean participants' responses for all three conditions (blue dots) and linear regression (dotted line). ***b***, Two-stage model prediction for DE data (triangles) and same regression (dotted line) as in ***a*** for comparison with the experimental data. ***c***, Experimental AE data of mean participants' responses for all conditions (green dots) and linear regression (dotted line). ***d***, Two-stage model predictions for AE sessions and same regression line as in ***c*** for comparison with the experimental data. Small displacements, intermediate displacements, and large displacements range, for DE: turquoise, light blue, and dark blue, respectively; for AE: bright green, olive, dark green, respectively.

0.02 ± 0.13; $\bar{w}_{\text{prior,AE}} = 0.39 \pm 0.12$, range $= [0.14 - 0.62]$; $\Delta\bar{x}_{\text{AE}} = 0.01 \pm 0.01$). The shift parameter Δx was not significantly different from zero for four of 14 participants in DE and three of 14 participants in AE sessions (remaining participants: DE: five participants, <0; five participants, >0; AE: five participants, <0; six participants, >0).

The second variant, or two-stage model, tests an iteratively updated version of the prior that additionally accounts for variations during the time course of the experiment. This model has, like the one-stage model, two free parameters (model fit group: $w_{\text{prior,DE}} = 0.34$, $\text{CI}_{95\%} = [0.30, 0.58]$; $\Delta x_{\text{DE}} = 0.03$, $\text{CI}_{95\%} = [0.01, 0.05]$; $w_{\text{prior,AE}} = 0.33$, $\text{CI}_{95\%} = [0.28, 0.40]$; $\Delta x_{\text{AE}} = 0.02$, $\text{CI}_{95\%} = [0, 0.04]$; individual participants: $\bar{w}_{\text{prior,DE}} = 0.36 \pm 0.15$, range $= [0.20 - 0.61]$; $\Delta\bar{x}_{\text{DE}} = 0.01 \pm 0.11$; $\bar{w}_{\text{prior,AE}} = 0.32 \pm 0.09$, range $= [0.14 - 0.62]$; $\Delta\bar{x}_{\text{AE}} = 0.02 \pm 0.07$). The shift parameter Δx was not significantly different from zero for six of 14 participants in DE and five of 14 participants in AE sessions (remaining participants: DE: four participants, <0; four participants, >0; AE: three participants, <0; six participants, >0).

The linear relationship between standard deviation and mean of the experimental data was deployed to derive an estimate of the noise sources to simulate the predicted mean reproduction noise of the model. The results for the two-stage model compared with the behavioral data are depicted in Figure 4, *b* and *d*, for the mean of all participants.

Figure 5 compares the experimental data to the mean displacement estimate by the two variants of the fitted Bayesian estimator model. Both variants agree well with the experimental data (coefficient-of-determination, one-stage model, model fit group: $R^2_{\text{AE}} = 0.98$, $R^2_{\text{DE}} = 0.97$; individual participants: $R^2_{\text{AE}} = 0.83 - 0.99$, $R^2_{\text{DE}} = 0.84 - 0.98$; two-stage-model, model fit group: $R^2_{\text{AE}} = 0.98$, $R^2_{\text{DE}} = 0.97$; individual participants: $R^2_{\text{AE}} = 0.80 - 0.99$, $R^2_{\text{DE}} = 0.88 - 0.98$). A nonparametric comparison indicated no significant difference between the R^2 values of individual participants for the one- and two-stage models ($p > 0.1$). However, the prior in the two-stage model arises due to the online estimation of the Kalman filter without any knowledge of the underlying sample distribution, whereas the current estimate of the prior in the one-stage model was set to be the mean of the respective underlying sample distribution. Therefore, the one-stage model requires the incorporation of additional knowledge compared with the two-stage case. Furthermore, the two-stage model with iterative update of the prior accounted for small variations in the data that were captured by the variations of the prior (Fig. 5, insets). Consequently, we considered the two-stage model to be superior to the one-stage model and used it for further analysis.

Figure 6*c* shows an example for a typical time course of the variable prior and measurement in one session. The range of displacements predicted by the prior estimates is smaller than that of the sample stimulus. This leads to a predicted measurement that covers a smaller range of displacements than the input stimuli. The

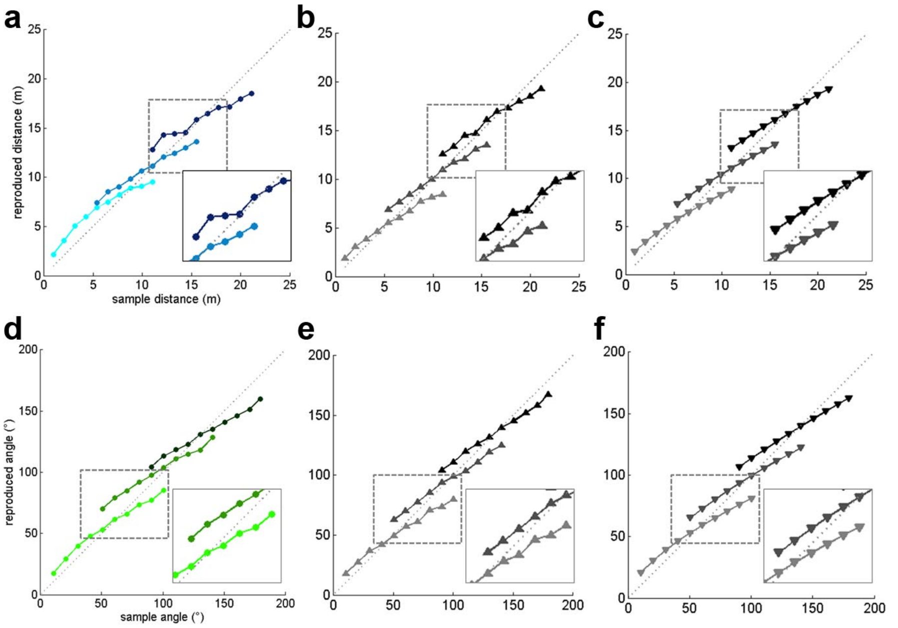

Figure 5. Summary of predicted and actual displacement estimation behavior. ***a***, ***d***, Mean reproduced distances in DE (***a***, blue dots) and AE (***d***, green dots) sessions for all conditions (small displacements, intermediate displacements, and large displacements range, for DE: turquoise, light blue, and dark blue, respectively; for AE: bright green, olive, dark green, respectively). ***b***, ***e***, Corresponding prediction of the fitted two-stage model (upward gray triangles). ***c***, ***f***, Corresponding prediction of the fitted one-stage model (downward gray triangles). The dotted lines in all plots indicate were reproduced estimate and sample displacement would be equal. Insets show small variations in the participants' responses that are captured by the two-stage model with varying prior but not by the one-stage model with fixed prior.

time constant of the evolution of Kalman gain varied between subjects (DE: $\tau_{DE} = [0.2 - 1.6]$ trials, AE: $\tau_{AE} = [0.3 - 2.1]$ trials). These values are similar to the time constants reported for the learning of pointing movements (van Beers, 2009) and shown for learning the mean of a prior distribution in a virtual coin-catching task (Berniker et al., 2010, their Fig. 6).

Predictions on single-subject behavior

Figure 6 compares the model predictions of mean and standard deviation to the individual participant's responses. The model captures individual differences between participants mainly by variation in the weighting of prior and measurement that in turn determine the slope of the predicted response curve. A strong weighting of the prior results in a more pronounced overestimation and underestimation, while a strong weighting of the measurement results in a predicted response that is very similar to the input stimuli (Fig. 6*a*,*b*, line of equality). Thus, the weighting reflects a scale invariant measure of the overall behavioral tendency of subjects in one experiment.

Within participants, behavior was compared by weighting the prior w_{prior} between the DE and AE experiments for each individual participant. We found a significant correlation between the weight in DE compared with AE sessions (linear regression: $r = 0.76, p = 0.001$). Figure 7 shows that the weighting of the prior within participants and between DE and AE conditions was more similar than between participants. In particular, the mean ratio of the AE versus DE weights was approximately equal ($w_{prior,AE}/w_{prior,DE} = 0.90$).

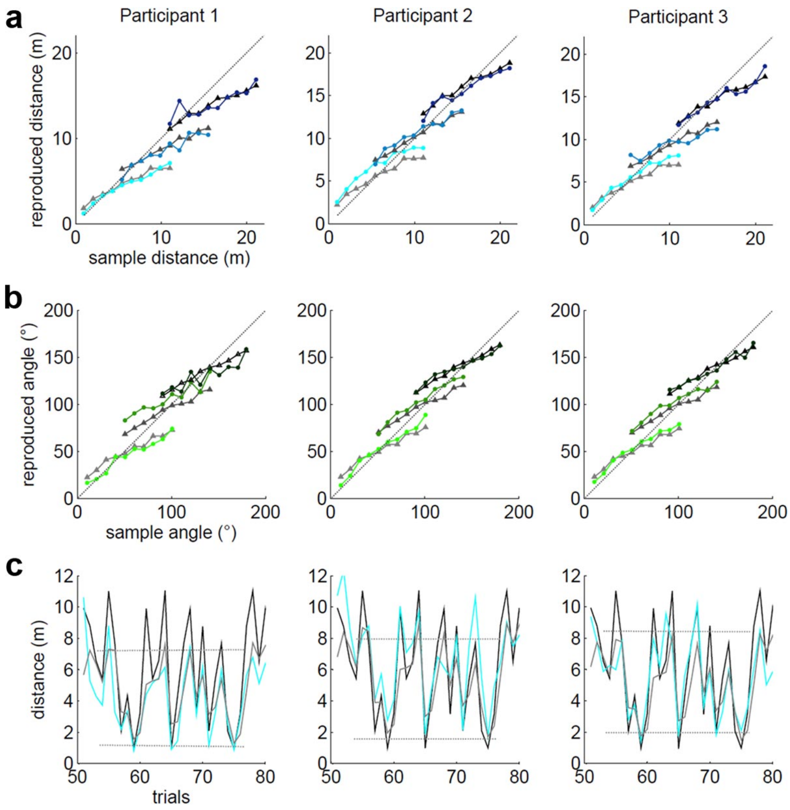

Figure 6. Comparison of individual participants' responses with predicted behavior of the two-stage model. ***a***, Mean responses of three selected participants in DE sessions and model predictions (gray triangles, $w_{prior,Participant1} = 0.39$, $w_{prior,Participant2} = 0.39$, $w_{prior,Participant3} = 0.40$). ***b***, Mean responses of the same participants in AE sessions and model predictions (gray triangles, $w_{prior,Participant1} = 0.15$, $w_{prior,Participant2} = 0.36$, $w_{prior,Participant3} = 0.38$). ***c***, Example of a typical time course for the same participants of sample displacements (black line), reproduced displacements (light blue line), and predicted displacements (gray line) within 30 trials of one DE session (small displacements range, trials: 50–80). Dotted lines indicate the range of displacements covered by the prediction of the model. Small displacements, intermediate displacements, and large displacements range, for DE: turquoise, light blue, and dark blue, respectively; for AE: bright green, olive, dark green, respectively.

Discussion

Human linear and angular displacement estimation is influenced by prior experience. We found that (1) reproduced distances and turning angles were biased toward the center of the underlying sample distribution, (2) the amount of bias increased with increasing sample range, and (3) the standard deviation for all conditions was linearly dependent on the mean reproduced displacement. These three characteristics are well captured by a model of an iterative Bayesian estimator that combines an experience-dependent a priori expectation with the actual noisy measurement to achieve an optimal estimate of displacement. We propose that our results are not limited to displacement estimation, but potentially hold for magnitude reproduction in general.

Behavioral findings in the context of the literature

Previous work on human linear and angular displacement perception found similar results to ours with a tendency to overshoot and undershoot certain displacements (Loomis et al., 1993; Ivanenko et al., 1997; Seemungal et al., 2007; Bergmann et al., 2011). An indication for the influence of prior experience can be found in work that shows that distance estimation and error magnitude vary considerably as a function of changes in the environmental experience (Ziemer et al., 2009) or stimulus range (Teghtsoonian and Teghtsoonian, 1978; Klatzky et al., 1990; Schwartz, 1999). Yet the estimation of distances and turning angles was mostly tested in different studies, because a direct comparison of the two different measures for single participants is difficult. In the present work, however, the model provides a chance to compare the two magnitudes in terms of individual weighting of prior and measurement, which is invariant to the measure of the magnitude. We found that, overall, individual participants seem to weight the prior for distances and turning angles similarly, whereas the differences between participants' weighting were higher. One possible reason for this is that there is a common processing mechanism for magnitudes in general, including the estimation of turning angles and distances as proposed by Walsh (2003). Another possible explanation is that the reliability of the input was very similar because both measures were based on optic flow in the same virtual environment (Frenz and Lappe, 2005; Mossio et al., 2008). However, the degree of reliance on prior information across tasks may also be a general trait that varies among individual subjects.

Experience-dependent Bayesian inference leads to a regression toward the mean

The regression effect, first referred to as the central tendency of judgment (Hollingworth, 1910), in psychophysical magnitude estimation is the tendency to correctly estimate magnitudes close to the center of the stimulus range and misestimate marginal

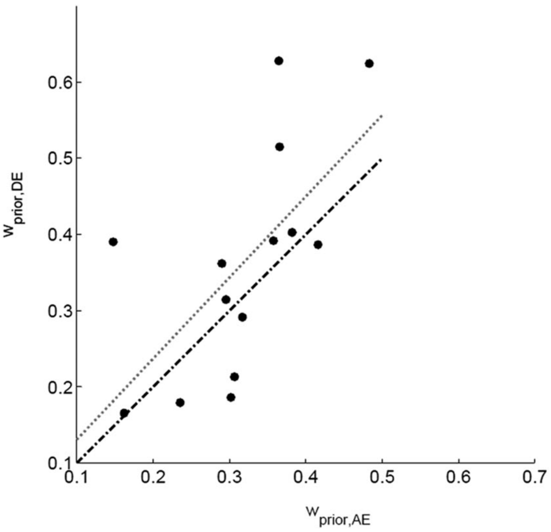

Figure 7. Comparison of the weighting of the prior within participants. Each dot represents the weight w_{prior} in AE and DE sessions for one participant. The gray dotted line indicates the linear regression between the two weights and the black dotted line is the line of equality.

ones: values presented at the lower end of the range are overestimated while those at the upper end show underestimation (Stevens and Greenbaum, 1966; Teghtsoonian and Teghtsoonian, 1978). Stevens (1971) attempted to explain this behavior as the tendency of the observer "to shorten the range of whichever variable he controls." A potential explanation for this tendency is Bayesian fusion of measurement and a priori expectation (Laming, 1999), as shown for displacement estimation in the present work. By multiplying the prior and the likelihood distributions, which correspond to a weighted average of prior and measurement in the Gaussian case, the estimate exhibits a shift from the measurement toward the a priori expectation. As shown in Figure 6*c*, an experience-dependent posterior estimate of randomly presented stimuli covers a smaller range, with displacements close to the center being more likely to occur; this consequently results in a regression toward the center of the sample range.

Dynamic prior knowledge adapts to the range of stimuli presented

Several studies have convincingly demonstrated that humans can use near-optimal strategies to combine stimulus uncertainty and prior information (Mamassian and Landy, 1998; Körding and Wolpert, 2004a; Tassinari et al., 2006). The a priori expectation in Bayesian models is often viewed as a fixed internal tendency that is due to general features in the world, e.g., that slow velocities are more likely to occur than fast ones (Weiss et al., 2002; Stocker and Simoncelli, 2006). Several studies, however, have shown that the a priori estimate can be modulated by short-term experience (Adams et al., 2004; Körding et al., 2004; Miyazaki et al., 2005) and its mean and variance could be learned during the experiment (Guo et al., 2004; Körding and Wolpert, 2004a; Berniker et al., 2010). Jazayeri and Shadlen (2010), for instance, assumed an experience-dependent prior expectation that was modeled as a continuous and fixed distribution, centered around the mean of the sample distribution. Indeed, it makes more sense for such a representation to arise over time. In the present work, we tested the fixed prior against a variable version that continuously updates its expectation with the previous measurement. The assumption behind the proposed updating procedure is that the mean of the stimulus distribution changes slowly over time, but in a way unknown to the system. We show that such an iterative updated prior, modeled by a Kalman filter, accounts for small variations in the data that are most likely due to the order of stimuli presented and cannot be explained by a version with fixed prior. Furthermore, this model provides an explanation for the origin and development of such a prior over time. In particular, the adaption to the underlying sample range for randomly presented stimuli (Teghtsoonian and Teghtsoonian, 1978; Kowal, 1993; Cheng et al., 2010) and also, potentially, the hysteresis effect, which refers to a dependence of the behavior on the order of stimuli, for experiments with nonrandom order (Eisler and Ottander, 1963; Hock et al., 2005) can result from the continuous update of a prior according to the experienced displacements.

Logarithmic Bayesian fusion leads to a direct link between Weber–Fechner and Stevens' power law

In the present work, we suggest that the most parsimonious explanation for the behavior is that displacement is coded internally on a logarithmic scale, as first proposed by Fechner (1860) based on Weber's law, which has shown to hold for human locomotor distance reproduction (Durgin et al., 2009). Similar results could, in principle, also be achieved on linear scales, assuming scalar variability, that is, a rise in the standard deviation with increasing mean (Rakitin et al., 1998; Cantlon et al., 2009). Recent work, however, supports the idea that numerical quantities (Dehaene, 2003; Nieder and Merten, 2007) and visual motion perception (Jürgens and Becker, 2006; Stocker and Simoncelli, 2006) are coded logarithmically in the brain. Note that latter authors assumed that Bayesian integration still takes place in linear space. However, as we have shown here, the power law, or Stevens' law, is a direct consequence of Bayesian integration on a logarithmic scale (Eqs. 2, 3, 9). Thus, as MacKay (1963) has shown before, the Weber–Fechner law and Stevens' law are indeed compatible. The proposed Bayesian fusion also assumes that the variance of a measured magnitude is independent of the magnitude on the logarithmic scale. On linear scales, this leads to a constant increase in standard deviation with increasing mean, as observed in the experimental data and corresponding to scale invariance found in both Weber–Fechner's and Stevens' laws (Chater and Brown, 1999).

Bayesian estimation of displacement, velocity, and time

Bayesian models succeeded in describing a variety of psychophysical data in related domains. The experimental design in the present study was very similar to recent work on interval timing (Jazayeri and Shadlen, 2010), allowing for a direct comparison of the behavioral findings. In particular, the same characteristic features, such as a tendency to the mean of the sample interval and an increase in bias with increasing sample range for estimation of traveled distances and turning angles, were previously reported for interval timing. Jazayeri and Shadlen (2010) tested different probabilistic approaches to combine the two sources of information given by sensory input and prior experience and concluded that a Bayesian observer model is statistically superior to maximum likelihood estimation (Ernst and Banks, 2002) or maximum a posteriori estimation in describing the main features of the behavioral data.

A large range of phenomena in motion perception, such as misestimation of speed and direction, was also successfully described by a Bayesian estimation process based on a prior that favors low speeds (Weiss et al., 2002; Stocker and Simoncelli, 2006). In line with this work, the variability in angular displacement perception has been proposed to be a result of Bayesian fusion of sensory inputs (Butler et al., 2010) and cognitive movement velocity (Jürgens and Becker, 2006). In the present work, however, movement velocity was varied randomly and independently of the experience condition. Thus, we show that the observed effects between conditions de facto depend on the experienced distances or turning angles. Yet the measurement of these displacements in the virtual world is still determined by an integration of optic flow, raising the question of whether the Bayesian estimation process is based on an estimate of displacement or takes place for time and velocity separately and is fused on a higher cognitive level to represent an estimate of displacement. The computational costs of the latter case would be higher for updating more than one magnitude and update of displacements alone in the present work provides a parsimonious explanation for a large number of findings.

Conclusion

From the realization of the iterative Bayesian estimator model, we infer that the systematic errors seen in human path integration behavior are the result of a performance-optimizing estimation process that exploits knowledge about previous behavior and the uncertainty of measurements. The model provides a direct link between Weber–Fechner and Stevens' power law. Consequently, we propose that our results are not limited to displacement estimation, but can potentially provide a unified explanation for commonly seen effects in psychophysical magnitude estimation studies, such as the range, regression effect, and hysteresis effect.

References

Adams WJ, Graf EW, Ernst MO (2004) Experience can change the 'light-from-above' prior. Nat Neurosci 7:1057–1058.

Bergmann J, Krauss E, Münch A, Jungmann R, Oberfeld D, Hecht H (2011) Locomotor and verbal distance judgments in action and vista space. Exp Brain Res 210:13–23.

Berniker M, Voss M, Körding K (2010) Learning priors for Bayesian computations in the nervous system. PLoS One 5:e12686.

Burge J, Ernst MO, Banks MS (2008) The statistical determinants of adaptation rate in human reaching. J Vis 8:20.1–20.19.

Butler JS, Smith ST, Campos JL, Bülthoff HH (2010) Bayesian integration of visual and vestibular signals for heading. J Vis 10:23.

Cantlon JF, Cordes S, Libertus ME, Brannon EM (2009) Comment on "Log or linear? Distinct intuitions of the number scale in Western and Amazonian indigene cultures". Science 323:38; author reply 38.

Chater N, Brown GD (1999) Scale-invariance as a unifying psychological principle. Cognition 69:B17–B24.

Cheng K, Spetch ML, Hoan A (2010) Categories and range effects in human spatial memory. Front Psychol 1:231.

Daum SO, Hecht H (2009) Distance estimation in vista space. Atten Percept Psychophys 71:1127–1137.

Dehaene S (2003) The neural basis of the Weber–Fechner law: a logarithmic mental number line. Trends Cogn Sci 7:145–147.

Dehaene S, Izard V, Spelke E, Pica P (2008) Log or linear? Distinct intuitions of the number scale in Western and Amazonian indigene cultures. Science 320:1217–1220.

Doya K, Ishi S, Pouget A, Rao RPN (2007) Bayesian brain-probabilistic approaches to neural coding. Cambridge, MA: MIT.

Durgin FH, Akagi M, Gallistel CR, Haiken W (2009) The precision of locomotor odometry in humans. Exp Brain Res 193:429–436.

Eisler H, Ottander C (1963) On the problem of hysteresis in psychophysics. J Exp Psychol 65:530–536.

Ernst MO, Banks MS (2002) Humans integrate visual and haptic information in a statistically optimal fashion. Nature 415:429–433.

Etienne AS, Jeffery KJ (2004) Path integration in mammals. Hippocampus 14:180–192.

Fechner GT (1860) Elemente der Psychophysik. Leipzig: Breitkopf and Härtel.

Fetsch CR, Turner AH, DeAngelis GC, Angelaki DE (2009) Dynamic reweighting of visual and vestibular cues during self-motion perception. J Neurosci 29:15601–15612.

Frenz H, Lappe M (2005) Absolute travel distance from optic flow. Vision Res 45:1679–1692.

Fujita N, Klatzky RL, Loomis JM, Golledge RG (1993) The encoding-error model of pathway completion without vision. Geograph Anal 25:295–314.

Glasauer S, Schneider E, Grasso R, Ivanenko YP (2007) Space-time relativity in self-motion reproduction. J Neurophysiol 97:451–461.

Glasauer S, Jahn K, Stein A, Brandt T (2009a) On the origin of systematic errors in a simple navigation task. BMC Neurosci 10[Suppl 1]:P13.

Glasauer S, Stein A, Günther AL, Flanagin VL, Jahn K, Brandt T (2009b) The effect of dual tasks in locomotor path integration. Ann N Y Acad Sci 1164:201–205.

Guo K, Nevado A, Robertson RG, Pulgarin M, Thiele A, Young MP (2004) Effects on orientation perception of manipulating the spatio-temporal prior probability of stimuli. Vision Res 44:2349–2358.

Hock HS, Bukowski L, Nichols DF, Huisman A, Rivera M (2005) Dynamical vs. judgmental comparison: hysteresis effects in motion perception. Spat Vis 18:317–335.

Hollingworth HL (1910) The central tendency of judgment. J Philos Psych Sci Meth 7:461–469.

Ivanenko Y, Grasso R, Israël I, Berthoz A (1997) Spatial orientation in humans: perception of angular whole-body displacements in two-dimensional trajectories. Exp Brain Res 117:419–427.

Jazayeri M, Shadlen MN (2010) Temporal context calibrates interval timing. Nat Neurosci 13:1020–1026.

Jürgens R, Becker W (2006) Perception of angular displacement without landmarks: evidence for Bayesian fusion of vestibular, optokinetic, podokinesthetic, and cognitive information. Exp Brain Res 174:528–543.

Jürgens R, Boss T, Becker W (1999) Estimation of self-turning in the dark: comparison between active and passive rotation. Exp Brain Res 128: 491–504.

Klatzky RL, Loomis JM, Golledge RG, Cicinelli JG, Doherty S, Pellegrino JW (1990) Acquisition of route and survey knowledge in the absence of vision. J Mot Behav 22:19–43.

Knill DC, Pouget A (2004) The Bayesian brain: the role of uncertainty in neural coding and computation. Trends Neurosci 27:712–719.

Körding KP, Wolpert DM (2004a) Bayesian integration in sensorimotor learning. Nature 427:244–247.

Körding KP, Wolpert DM (2004b) The loss function of sensorimotor learning. Proc Natl Acad Sci U S A 101:9839–9842.

Körding KP, Ku SP, Wolpert DM (2004) Bayesian integration in force estimation. J Neurophysiol 92:3161–3165.

Kowal KH (1993) The range effect as a function of stimulus set, presence of a standard, and modulus. Percept Psychophys 54:555–561.

Laming D (1999) Prior expectations in cross-modality matching. Math Social Sciences 38:343–359.

Lappe M, Jenkin M, Harris LR (2007) Travel distance estimation from visual motion by leaky path integration. Exp Brain Res 180:35–48.

Loomis JM, Klatzky RL, Golledge RG, Cicinelli JG, Pellegrino JW, Fry PA (1993) Nonvisual navigation by blind and sighted: assessment of path integration ability. J Exp Psychol Gen 122:73–91.

MacKay DM (1963) Psychophysics of perceived intensity: a theoretical basis for Fechner's and Stevens' laws. Science 139:1213–1216.

Mamassian P, Landy MS (1998) Observer biases in the 3D interpretation of line drawings. Vision Res 38:2817–2832.

Mittelstaedt ML, Glasauer S (1991) Idiothetic navigation in gerbils and humans. Zool Jahrb Abt Physiol 427–435.

Mittelstaedt ML, Mittelstaedt H (1980) Homing by path integration in a mammal. Naturwissenschaften 67:566–567.

Miyazaki M, Nozaki D, Nakajima Y (2005) Testing Bayesian models of human coincidence timing. J Neurophysiol 94:395–399.

Mossio M, Vidal M, Berthoz A (2008) Traveled distances: new insights into the role of optic flow. Vision Res 48:289–303.

Nieder A, Merten K (2007) A labeled-line code for small and large numerosities in the monkey prefrontal cortex. J Neurosci 27:5986–5993.

Rakitin BC, Gibbon J, Penney TB, Malapani C, Hinton SC, Meck WH (1998) Scalar expectancy theory and peak-interval timing in humans. J Exp Psychol Anim Behav Process 24:15–33.

Riecke BE, Veen HAHCv, Bülthoff HH (2002) Visual homing is possible without landmarks: a path integration study in virtual reality. Presence 11:443–473.

Schwartz M (1999) Haptic perception of the distance walked when blindfolded. J Exp Psychol Hum Percept Perform 25:852–865.
Seemungal BM, Glasauer S, Gresty MA, Bronstein AM (2007) Vestibular perception and navigation in the congenitally blind. J Neurophysiol 97:4341–4356.
Stevens SS (1957) On the psychophysical law. Psychol Rev 64:153–181.
Stevens SS (1961) To honor Fechner and repeal his law: a power function, not a log function, describes the operating characteristic of a sensory system. Science 133:80–86.
Stevens SS (1971) Issues in psychophysical measurement. Psychol Rev 78:426–450.
Stevens SS, Greenbaum H (1966) Regression effect in psychophysical judgment. Atten Percept Psychophys 1:439–446.
Stocker AA, Simoncelli EP (2006) Noise characteristics and prior expectations in human visual speed perception. Nat Neurosci 9:578–585.
Tassinari H, Hudson TE, Landy MS (2006) Combining priors and noisy visual cues in a rapid pointing task. J Neurosci 26:10154–10163.
Teghtsoonian R, Teghtsoonian M (1978) Range and regression effects in magnitude scaling. Percept Psychophys 24:305–314.
van Beers RJ (2009) Motor learning is optimally tuned to the properties of motor noise. Neuron 63:406–417.
Walsh V (2003) A theory of magnitude: common cortical metrics of time, space and quantity. Trends Cogn Sci 7:483–488.
Weiss Y, Simoncelli EP, Adelson EH (2002) Motion illusions as optimal percepts. Nat Neurosci 5:598–604.
Zanker JM (1995) Does motion perception follow Weber's law? Perception 24:363–372.
Ziemer CJ, Plumert JM, Cremer JF, Kearney JK (2009) Estimating distance in real and virtual environments: does order make a difference? Atten Percept Psychophys 71:1095–1106.

3 INCORPORATION OF ABSTRACT A-PRIORI ASSUMPTIONS

> Priority is a function of context.
>
> Stephen R. Covey

3.1 SUMMARY

Context matters. The second investigation aims to clarify (1) whether humans are also capable of incorporating other more abstract types of a-priori knowledge, apart from prior experience, and (2) if this behavior can also be explained by a Bayesian framework.
We therefore adjusted the production-reproduction task for linear displacements to test three different experimental conditions. In the first condition test displacements were quasi-randomly drawn from a range of short displacements for the first half of the trials and from a range of larger displacements for the second half of the trials. As in the previous study both ranges were overlapping and the influence of immediate prior experience should cause estimates that are shifted towards the center of the respective sample range. In the second condition the exact same displacements were tested but in a fully randomized order, resulting in a single non-uniform sample range. In this case, we expected no separation into two distinct sample ranges, but a shift of estimates towards the center of the full sample distribution. The third condition introduced an abstract symbolic cue. The exact same order of displacement was used as in the second condition, however, this time each production-reproduction task was preceded by a verbal cue that indicated that the next displacement will either belong to the short or to the long range of displacements.
We assumed three potential scenarios representing how subjects might deal with this additional abstract knowledge. Either they ignore the symbolic cue and rely solely on their short-term prior experience. Then the observed behavior resemble that in condition two. Or subjects make full use the cue to separate the

tested displacements into two distinct ranges. Then the behavior should resemble that of condition one. Or, finally, subjects may make use of the symbolic cue and at the same time also rely on their short-term prior experience which should result in a behavioral performance that resembles a mixture of condition one and two.
We find that behavioral performance in condition three was significantly different from that in condition two, however, no significant difference was found compared to condition one. This suggests that subjects are capable of exploiting the information provided by the abstract verbal cue for their estimate of displacement.
Two alternations of the basic iterative model in chapter 2 are proposed that test two different generative versions of how the additional a-priori knowledge, the sensory input and prior experience are incorporated in the Bayesian estimation process to yield a combined estimate of displacement (cue-combination and categorization). Both versions account equally well for the observed behavior.

3.2 REFERENCE

This work was carried out under the supervision of Stefan Glasauer;
F.H.P. and S.G. designed research; F.H.P., P.M. and S.G. performed research; F.H.P. conducted the experiments; F.H.P. and P.M. analyzed data; F.H.P., P.M. and S.G. wrote the paper.

The paper was published in *Frontiers in Integrative Neuroscience* under the following reference:

FH Petzschner, P Maier and S Glasauer: Combining symbolic cues with sensory input and prior experience in an iterative Bayesian framework *Frontiers in Integrative Neuroscience*; 6:58, 1-18, 2012.

frontiers in
INTEGRATIVE NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 13 August 2012
doi: 10.3389/fnint.2012.00058

Combining symbolic cues with sensory input and prior experience in an iterative Bayesian framework

Frederike H. Petzschner [1,2,3,4] ****, Paul Maier*** [1,2] ***and Stefan Glasauer*** [1,2,3,4]

[1] *Institute for Clinical Neurosciences, Ludwig-Maximilians-University Munich, Munich, Germany*
[2] *Bernstein Center for Computational Neuroscience Munich, Munich, Germany*
[3] *Graduate School of Systemic Neurosciences, Ludwig-Maximilians-University-Munich, Munich, Germany*
[4] *Integrated Center for Research and Treatment of Vertigo, Ludwig-Maximilians-University Munich, Munich, Germany*

Edited by:
Zhuanghua Shi, Ludwig-Maximilians-Universität München, Germany

Reviewed by:
Toemme Noesselt, Otto-von-Guericke-University, Germany
David R. Wozny, Carnegie Mellon University, USA

****Correspondence:***
Frederike H. Petzschner, Institute for Clinical Neurosciences, Ludwig-Maximilians-Universität, Marchioninistrasse 23, 81377 München, Germany.
e-mail: fpetzschner@lrz.uni-muenchen.dex

Perception and action are the result of an integration of various sources of information, such as current sensory input, prior experience, or the context in which a stimulus occurs. Often, the interpretation is not trivial hence needs to be learned from the co-occurrence of stimuli. Yet, how do we combine such diverse information to guide our action? Here we use a distance production-reproduction task to investigate the influence of auxiliary, symbolic cues, sensory input, and prior experience on human performance under three different conditions that vary in the information provided. Our results indicate that subjects can (1) learn the mapping of a verbal, symbolic cue onto the stimulus dimension and (2) integrate symbolic information and prior experience into their estimate of displacements. The behavioral results are explained by to two distinct generative models that represent different structural approaches of how a Bayesian observer would combine prior experience, sensory input, and symbolic cue information into a single estimate of displacement. The first model interprets the symbolic cue in the context of categorization, assuming that it reflects information about a distinct underlying stimulus range (categorical model). The second model applies a multi-modal integration approach and treats the symbolic cue as additional sensory input to the system, which is combined with the current sensory measurement and the subjects' prior experience (cue-combination model). Notably, both models account equally well for the observed behavior despite their different structural assumptions. The present work thus provides evidence that humans can interpret abstract symbolic information and combine it with other types of information such as sensory input and prior experience. The similar explanatory power of the two models further suggest that issues such as categorization and cue-combination could be explained by alternative probabilistic approaches.

Keywords: pre-cueing, path integration, cue-combination, multi-modal, categorization, experience-dependent prior, magnitude reproduction, iterative Bayes

INTRODUCTION

Because the demands in natural tasks are highly complex but sensory information is corrupted by noise, humans are versed in exploiting contextual information. To improve efficiency, reduce the amount of computational costs, and allow fast adaption to the outside world, we infer existing dependencies and combine relevant information to guide our perception and action. The sources of information can vary from the simultaneous input coming from different senses (Ernst and Bülthoff, 2004; Angelaki et al., 2009) or distinct input from one sensory modality (Jacobs, 1999; Stone et al., 2009), over short and long-term experience (Adams et al., 2004; Stocker and Simoncelli, 2006; Verstynen and Sabes, 2011), to abstract expectations and contextual cues in the environment (Langer and Bülthoff, 2001).

A possible framework for combining these diverse sources of uncertain information is offered by Bayesian probability theory, which has proven applicable to several of the mentioned issues. It provides a normative, mathematical description of how various sources of information can be merged to obtain a statistically optimal estimate of their cause in the presence of uncertainty. One of the most common applications of the Bayesian approach is multimodal cue integration, where the provided information about a stimulus results from different sensory modalities, such as vision, audition, or proprioception (Ernst and Banks, 2002; Battaglia et al., 2003; Körding et al., 2007).

Senses, however, are not the only source of information that determines our perception. Contextual and symbolic cues can also contribute as a new source of information. In visual search paradigms, contextual cues are known to influence reaction times (e.g., Müller et al., 2003; Vincent, 2011). The context can also lead to an internal organization of stimuli into distinct categories that influence perception by leading to an increased ability to discriminate between categories at the expense of discriminability within categories. Examples for category effects range from the perception of speech sounds (Liberman et al., 1957) or colors (Davidoff et al., 1999) to facial expressions (Etcoff and Magee, 1992). A

Bayesian explanation for category effects in speech perception was offered by Feldman et al. (2009). However, their solution only treats implicit predefined categories, not auxiliary contextual cues providing information about these categories, e.g., pre-cueing.

Another type of contextual information comes from the preceding occurrence of a stimulus in the form of prior experience. Bayesian probability theory has been successfully applied to a broad spectrum of studies exploring the effect of short or long-term experience on our current percept (Adams et al., 2004; Stocker and Simoncelli, 2006; Verstynen and Sabes, 2011). For human estimation of distances and turning angles in a production-reproduction task, we have recently shown that the effect of prior experience results in a varying bias depending on the underlying sample range (Petzschner and Glasauer, 2011). The participants' behavior was best explained by an iterative Bayesian estimate derived from the current noisy measurement merged with information from short-term prior experience, which is updated on a trial by trial basis.

Sensory input is often embedded not just in the temporal context of prior experience, but occurs together with other indirect cues that provide a contextual environment helping to interpret the sensory input. These indirect or symbolic cues join together with sensory input and experience to yield a uniform percept. While there is a considerable body of research on multi-modal sensory fusion, the mechanisms of integration of symbolic cues into sensory perception are less well understood.

The present work aims to clarify the role of auxiliary contextual cues on behavior that is known to be influenced by prior experience. We extended our distance production-reproduction task (Petzschner and Glasauer, 2011) to include a symbolic cue that supplied additional, but initially uncertain information about the stimulus value. The symbolic cue values were provided as a written instruction prior to each trial and indicated whether the distance to be reproduced would be "short" or "long." The cue values corresponded to two ranges of distances. We investigated whether (1) subjects could use such a symbolic cue that provided reliable but imprecise information about the sample distances and (2) how this abstract information influenced their estimation process. To evaluate the behavioral results in the cue condition we used two control conditions that mimicked the extreme cases of cue usage. In the first control condition, we presented participants with exactly the same distances in the same order, but without the symbolic cue. In the second control condition the "short" and "long" ranges of displacements were presented in a separate order. Thus, if subjects ignored the symbolic cue, we expected that the performance in the cue condition would resemble that of the first control condition. If subjects however separated their estimates based on the symbolic cue, the behavior should be similar to the second control condition.

We then compare the behavioral data to predictions of two distinct Bayesian observer models, the *categorical* and the *cue-combination model*, which are founded on qualitatively different assumptions about the causal relationship between the sensory stimulus and the symbolic cue and consequently, about how the mapping of the symbolic cue to the stimulus dimension is learned during the experiment. Both models are based on our previously published *basic iterative model* (Petzschner and Glasauer, 2011, see **Figure 1A**) and generate a combined estimate of the distance to be reproduced given the observed stimulus, the symbolic cue, and prior experience. In addition, in both models Kalman filters are used to dynamically update the prior experience and to learn the relation between sensory stimulus and symbolic cue.

The two models differ in how the symbolic cue is merged with prior experience and sensory input into a distance estimate. This difference corresponds to different assumptions about the causal outside world structures between the stimulus, the measurement, and the symbolic cue (see **Figure 1**). In the *categorical model*, the idea is that the symbolic cue helps to identify an underlying stimulus category (Feldman et al., 2009). The model is based on the assumption that in the outside world, in each trial one of two categories is chosen, which determines the range of test distances. The test distance, which is drawn randomly from the respective category, leads to a noisy distance measurement. In addition, the symbolic cue signifies the chosen category with a certain reliability (**Figure 1B**). In the *cue-combination model*, it is assumed that the symbolic cue provides additional information similar to a sensory signal from a different modality (e.g., Ernst and Banks, 2002). The cue-combination model has a different view on the outside world. As our previous basic iterative model (Petzschner and Glasauer, 2011), it assumes the test distances are drawn from a single range, instead of distinct categories. The chosen test distance leads to a noisy distance measurement and to a noisy cue signal, which determines the symbolic cue (**Figure 1C**).

MATERIALS AND METHODS

PARTICIPANTS

Twenty volunteers (nine female) aged 20–29, who had all normal or corrected-to-normal vision and were naive to the purpose of the experiments, took part in the study. Participation was monetarily compensated. The experiments were approved by the local ethics committee and conducted in accordance with Declaration of Helsinki.

EXPERIMENTAL SETUP

Stimuli were viewed binocularly on a PnP monitor driven by an NVIDIA GeForce 8800 GTX graphics card at a frame rate of 60 Hz and with a monitor resolution of 1920 × 1200. All experiments were carried out in complete darkness except for the illumination by the monitor. The real-time virtual reality (VR) was created using Vizard 3.0 (Worldviz, http://www.worldviz.com/ and depicted the same artificial stone desert as described in Petzschner and Glasauer (2011), consisting of a textured ground plane, 200 scattered stones that served as pictorial depth cues, and a textured sky (**Figure 2**). The orientation of the ground plane texture, the position of the stones, and the starting position of the participant within the VR were randomized in each trial to prevent participants from using landmark cues to calibrate their estimate of displacement. The sky was simulated as a 3D dome centered on the participant's current position and thus the distance to the horizon was kept constant. In the VR each participant's eye height was adjusted individually to his/her true eye height. A multi-directional movable joystick (SPEEDLINK) was used to change the position with a constant speed.

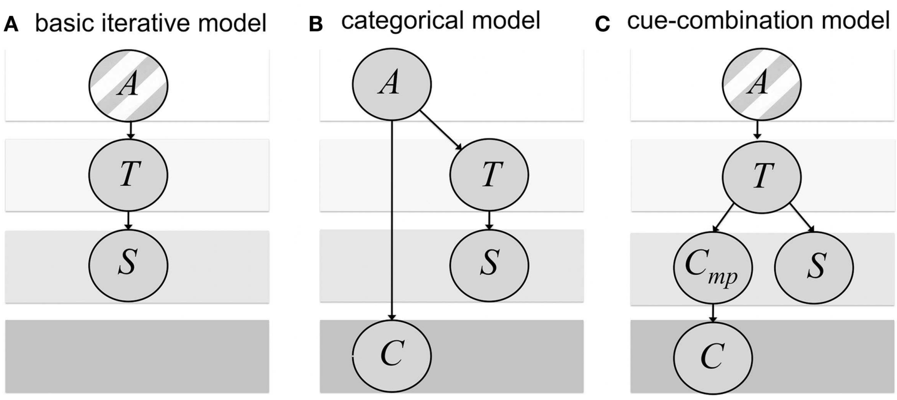

FIGURE 1 | Bayesian networks of the generative probabilistic models corresponding to the estimation part (i.e., dependence on previous trials not shown). The assumed probabilistic dependencies are shown as arrows. **(A)** Basic iterative model as described in Petzschner and Glasauer (2011). The stimulus *S* is a noisy measurement of the target distance *T* that is drawn from a single underlying category *A*. **(B)** Categorical model: the target distance *T* and the discrete symbolic cue *C* depend on the choice of the underlying category *A*. Again, the stimulus *S* is a noisy measurement of the target distance *T*. **(C)** Cue-combination model: The stimulus *S* and cue signal C_{mp} represent both independent noisy measurements of the target distance *T* that is drawn from a single underlying category *A*. The cue signal C_{mp} is mapped to the symbolic cue *C*. The striped background in **(A,C)** indicates that *T* is assumed to be drawn from a single category *A* in contrast to **(B)** were target distance and cue depend on the choice of the underlying category.

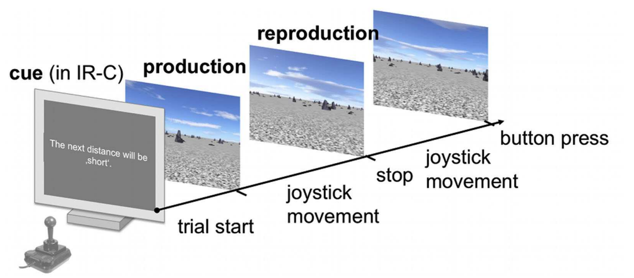

FIGURE 2 | Schematic time course of a single trial. Subjects had to subsequently produce and reproduce a sample distance in a virtual reality using the joystick to change their position with a constant speed. The final position was indicated via button press. In the IR-C condition each production-reproduction block was preceded by a symbolic cue that declared the upcoming sample displacement to be either "short" or "long." No symbolic cue was displayed in the IR-NC and BR-NC conditions.

EXPERIMENTAL PROCEDURE

Subjects had to estimate traveled distances in a production-reproduction task in three different experimental conditions, "blocked-ranges, no cue" (BR-NC), "interleaved-ranges, no cue" (IR-NC), and "interleaved-ranges, cue" (IR-C). The task remained the same for all three conditions.

Task

In each trial subjects were asked to "produce" a certain sample distance, by using a joystick to move forward through the virtual environment on a linear path toward the direction of a visual object at the horizon of the virtual world until they were automatically stopped for 2.25 s. During that time they received an instruction to subsequently "reproduce" the same amount of displacement that they had experienced during the production phase. Throughout the reproduction phase subjects continued moving in the same direction as in the production phase and indicated via button press when they thought they had covered the same distance as in the production phase. In the condition with cues the symbolic cue was presented before the production phase. **Figure 2** displays a schematic overview of the time course of events in a single trial. In all trials velocity was kept constant during one movement, but changed randomly up to ±60% (scaling factors between joystick output and constant VR velocity were drawn from a normal distribution) between production and reproduction phase to exclude time estimation strategies to solve the task.

Experimental conditions

Each experimental condition consisted of 110 trials. The first 10 trials per condition were training trials and served to familiarize participants with the task and VR. During these 10 trials, feedback on the performance was given after the reproduction phase by asking subjects to navigate toward an object that was displayed at the correct distance in the VR. The following 100 trials were test trials without any feedback. Only test trials were used for data analysis. After 50 trials subjects had a short break of 100 s to relax their hands. During that time the subjects did not leave their position and the room remained dark. Different experimental conditions were separated by a break for no less than 15 min outside the room of the experiment. In all three conditions the overall number of repetitions for each sample distance remained the same, thus the overall distribution of samples was the same for all three conditions. The same trial order within one condition as well as the same order of cues in the cued condition was maintained for all participants. The three experimental conditions were performed in a randomized order.

"Blocked-ranges, no cue" condition. In the BR-NC condition the 100 test distances were drawn in two blocks from two different underlying uniform sample distributions referred to as "short" range ([5, 7, 9, 11, 13] m) and "long" range ([11, 13, 15, 17, 19] m).

In the first block of 50 trials the sample distances were randomly drawn from the "short" range distribution; sample distances for the second block of 50 trials were randomly drawn from the "long" range distribution. The two blocks were separated by a short break of 100 s. Within each range each sample distance was repeated 10 times in a randomized order. Note that the 11 and 13-m distances appeared in both the "short" and "long" range distribution, and were thus repeated 20 times in the overall condition. Thus, we refer to these displacements as overlapping samples. Subjects received no additional information about the underlying sample distribution (**Figure 3A**).

"Interleaved-ranges, no cue" condition. In the IR-NC condition the same sample distances of the two distributions were tested as in the BR-NC condition, however in a interleaved order resulting in one randomized, non-uniform sample distribution [5, 7, 9, 11, 13, 15, 17, 19] m. All samples were repeated 10 times during the overall condition, except the 11 and 13-m distance, which were again repeated 20 times. As above subjects received no additional information about the underlying sample distribution (**Figure 3B**).

"Interleaved-ranges, cue" condition. In the IR-C condition sample distances were tested in the exact same order as in the

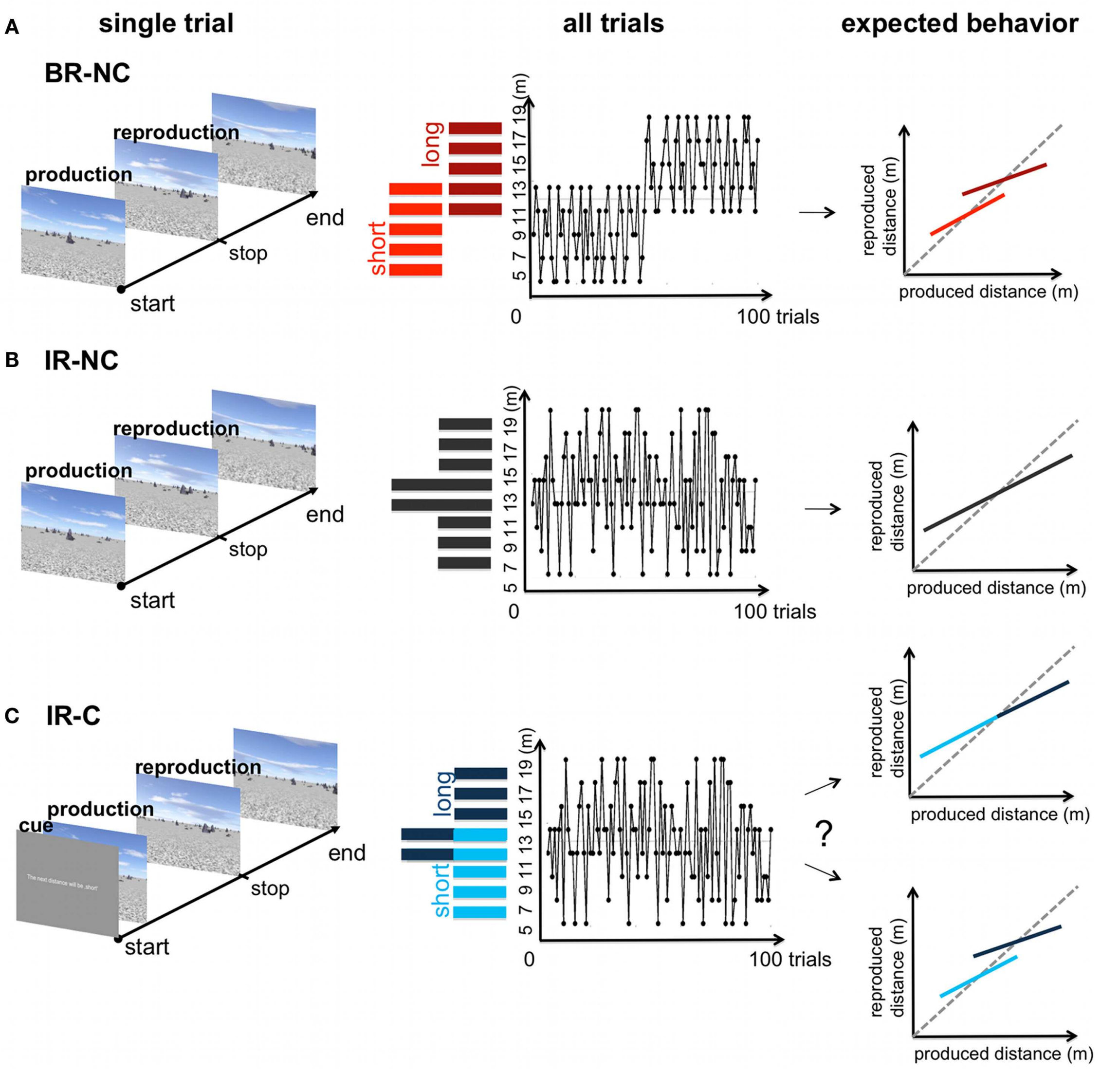

FIGURE 3 | Overview of the three experimental conditions. Left: time course of one trial in the distance production-reproduction task. Middle: distribution and trial sequence for the blocked and interleaved-ranges. Right: Potential behavioral response. **(A)** BR-NC condition: the two sample ranges were tested in a blocked order. In the first half of the trials a range of "short" displacements was tested, in the second half of the condition a range of "long" distances was tested. Both ranges were overlapping for two distances (11 and 13 m) **(B)** IR-NC condition: The same displacements as in **(A)** where tested in the production-reproduction task, but in a interleaved order resulting in one non-uniform range of randomized sample displacements. **(C)** IR-C condition: displacements were tested in the exact same order as in **(B)**, but each trial started with a symbolic cue that indicated either a "short" or "long" displacement. No further information was provided. Depending on the influence of the symbolic cue the resulting behavior could range between the extreme cases mimicked in **(A,B)**.

IR-NC condition based on one non-uniform sample distribution (**Figure 3C**). However this time subjects were told that there are two different types of samples referred to as "short" and "long" distances and that, in order to improve their performance, they would receive a written, symbolic cue that indicated which type the upcoming distance would belong to. No further information on the meaning of "short" and "long" was provided. At the beginning of each trial the sample distance was assigned on the screen to belong to one of the two types ("The next test distance will be short" or "The next test distance will be long"). All distances ranging from 5 to 9 m and one half of the 11 and 13-m distance samples were announced as being "short," all distances ranging from 15 to 19 m and the other half of the 11 and 13-m distances were announced as being "long." Thus the symbolic cue was always valid, except for distances 11 and 13 m, where the same distance could either be referred to as "short" or "long." Consequently, the separation provided by the symbolic cue is comparable to the two temporally separate ranges in the BR-NC condition.

DATA ANALYSIS

Participants' position and orientation within the VR were sampled at 20 Hz. The reproduced displacement was calculated as the difference between the position at the time of the button press and the produced displacement.

To test for differences in the behavior that are due to the use of the underlying sample range or the written symbolic cue, trials in all three conditions were split into two groups, the ranges "short" and "long." For the BR-NC condition, where the two distributions were tested consecutively, this was achieved by splitting the trials into two halves ("short": trials 1–50; "long": trials 51–100). In both the IR-NC and IR-C condition trials were split according to the symbolic cue ("short" and "long") given in the IR-C condition. Note that we also split the IR-NC condition in order to provide a direct comparison of the same trials with and without symbolic cue.

Differences in the behavioral data for the two ranges can be easily examined by comparing across those displacements that were tested in both ranges (11 and 13 m). Thus we refer to the comparison of 11 and 13 m between the "short" and "long" range as "overlapping samples comparison."

Data analysis was conducted in MATLAB R2010b (MathWorks). Statistical differences were assessed using repeated-measures analysis of variance (rm-ANOVA). A probability level of $p < 0.05$ was considered significant for all statistical analysis. To assess differences between conditions and ranges we used rm-ANOVA for the "overlapping samples comparison" with the within-subjects factors *condition* (BR-NC, IR-NC, IR-C), *range* ("short" vs. "long") and *distance* (two distances, 11 and 13 m). Since the use of the symbolic cue should have an effect not just on the "overlapping samples," but also on the whole set of presented distances, we tested the difference between conditions by a second rm-ANOVA for the mean reproduction error with the within-subject factors *condition* (BR-NC, IR-NC, IR-C) and *distance* (10 distances, see "Blocked-Ranges, No Cue" Condition).

MODELING

In our previous study we proposed a model of iterative Bayesian estimation that explained subjects performance in a distance production-reproduction task by the incorporation of prior experience into the estimation process (Petzschner and Glasauer, 2011). This basic iterative model is applied to explain the data for the two conditions without symbolic cue (BR-NC and IR-NC) in the present work (**Figure 1A**). For the symbolic cue condition (IR-C) the model must be extended to incorporate information that is not only driven by prior experience but the symbolic cue itself. Important for such an extension is the interpretation of the symbolic cue. Neither the symbolic cue itself nor the experimental instruction specified (1) the value or range of values in the stimulus dimension it corresponds to, and (2) the proportion of trials in which the symbolic cue is actually valid.

As mentioned in the Introduction, we propose two qualitatively different ideas how the symbolic cue could be interpreted, how the mapping of the symbolic cue to the stimulus dimension is learned, and how it is finally integrated into the estimation process. The first interpretation, referred to as categorical model, assumes that the symbolic cue C is an indicator for a category A that determines the distribution from which the target distance T, that is the distance to be reproduced, is being drawn (**Figure 1B**). This interpretation corresponds largely to the categorical model proposed by Feldman et al. (2009), except that in their model there is no symbolic cue provided to the observer. The second interpretation, referred to as cue-combination model, assumes that the target T is drawn from one single distribution and the symbolic cue C provides additional evidence about T just like a sensory cue from another modality (**Figure 1C**). Thus, this second interpretation leads to a multimodal fusion model in which one sensory input S, the stimulus measurement, is continuous and the other sensory input C, the symbolic cue, is discrete.

In the following, the two models are described in detail. Each model has three free parameters, which are explained in the respective section. We first describe the estimation part that fuses sensory measurement, symbolic cue, and prior experience. We then separately describe the update part that implements a discrete Kalman Filter as iterative Bayesian algorithm to update cue-related priors (categorical model) or calibrate likelihoods (cue-combination model).

The estimation part of the two models is also illustrated in **Figure 4** by displaying how the prior information, the symbolic cue, and the sensory likelihood function are transformed into a posterior distribution, which determines the reproduced distance.

We use a mathematical notation where we refer to random variables with upper case letters (e.g., A,S,T,C), to values for discrete variables such as cue and category with indexed lower case letters (e.g., c_i), and to values for continuous variables such as the sensory input with lower case letters (e.g., s). Furthermore, we abbreviate notations such as $P(T,A = a_i \mid S,C)$ to $P(T,a_i \mid S,C)$.

Categorical model

The categorical model follows Feldman et al. (2009) for the definition of the distributions. We assume that the target distance T is

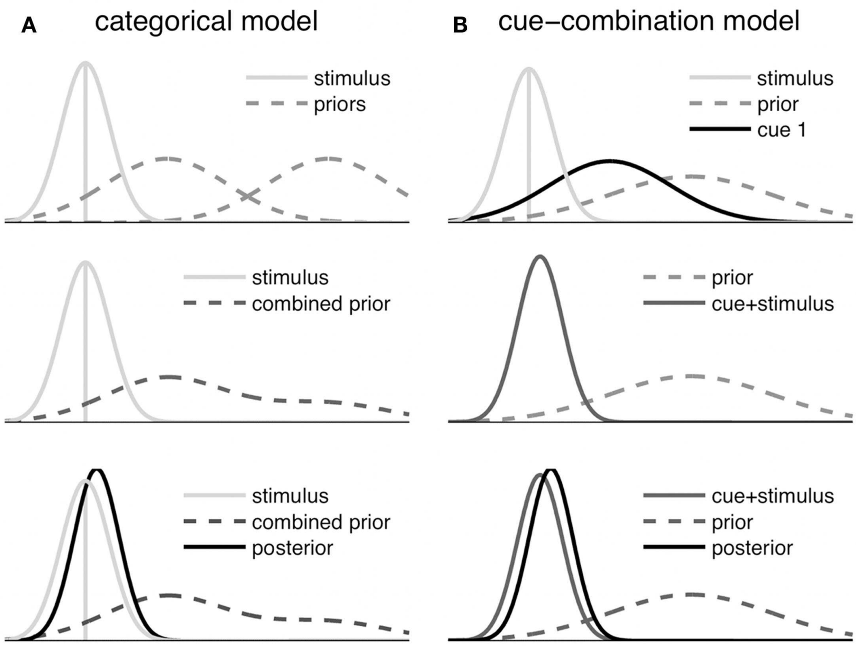

FIGURE 4 | Schematic illustration of the Bayesian fusion in the symbolic cue models. Only the estimation step is shown, which does not include updating based on information from previous trials. **(A)** Categorical model: the category priors are merged after weighting each prior with the conditional probability of the respective category given the sensory input and the symbolic cue. Then the resulting Gaussian mixture distribution (combined prior) is fused with the stimulus measurement to derive the posterior. **(B)** Cue-combination model: first the stimulus likelihood and the likelihood corresponding to the current symbolic cue signal are fused. Then this fused signal (cue + stimulus) is combined with the prior, yielding the posterior.

drawn from a normally distributed category

$$T|A \sim N\left(\mu_A, \sigma_A^2\right) \tag{1}$$

and that categories $A = a_i$ have individual means μ_{a_i}, but share the same variance σ_A^2. Our generative model assumes that categories a_i themselves are drawn uniformly from one of n possible categories ($n = 2$ in the present experiment, see **Figure 4** top left). Due to measurement noise, T cannot be sensed directly, but only the noisy measurement S with the conditional Gaussian distribution

$$S|T \sim N\left(T, \sigma_S^2\right). \tag{2}$$

In addition to the direct stimulus measurement S, participants are presented with a symbolic cue value c_j, which provides information about the underlying category. Nevertheless there is some uncertainty associated with the symbolic cue. Accordingly the cue reliability, that is, the probability of the correct symbolic cue value being presented, given a certain category a_j, is specified as $p_C = P(c_j \mid a_j)$ and assumed to be constant over trials. Accordingly the probability of being presented with a wrong symbolic cue out of $n-1$ remaining cues, is

$$P\left(c_j|a_i\right) = \frac{1 - p_C}{n - 1}. \tag{3}$$

To reproduce the target distance T, we are interested in the posterior distribution $P(T \mid S,C)$. To infer this posterior distribution, we first calculate the probability $P(T,A \mid S,C)$, which can be derived by applying Bayes' law to the complete joint distribution $P(T,A,S,C)$, and then marginalize over the category A:

$$P\left(T|S, C\right) = \sum_i^n P\left(T, a_i|S, C\right). \tag{4}$$

We show in the Appendix that, with the conditional dependency assumptions for this model (see **Figure 1B**), we can rewrite the posterior as

$$P\left(T|S, C\right) = \sum_i^n P\left(T|S, a_i\right) \cdot P\left(a_i|S, C\right). \tag{5}$$

The category-dependent posteriors $P(T \mid S,a_i)$, which now are independent of the symbolic cue C, are weighted by the posterior probabilities $P(a_i \mid S,C)$ of the categories given stimulus S and symbolic cue C.

To infer the target distance we compute the mean of the posterior $P(T \mid S,C)$. Analogous to the equation above, the mean of the posterior can be computed as weighted sum of conditional expectations of the category-dependent posteriors, where the weights are again the posteriors of the categories.

$$E\left[T|s, c_j\right] = \sum_i^n P\left(a_i|s, c_j\right) E\left[T|s, a_i\right]. \tag{6}$$

We show in the Appendix that this can be reformulated as

$$E\left[T|s,c_j\right] = w_\mathrm{m}s + (1-w_\mathrm{m})\sum_i^n P\left(a_i|s,c_j\right)\cdot\mu_{a_i}. \quad (7)$$

That is, a weighted sum of the category means μ_{a_i} forms the mean of a Gaussian mixture distribution (see **Figure 4A** middle), and this mean is summed with measurement s weighted by w_m. The measurement weight w_m is determined by the measurement and category variances:

$$w_\mathrm{m} = \frac{\sigma_A^2}{\sigma_A^2+\sigma_S^2} \qquad 1-w_\mathrm{m} = \frac{\sigma_S^2}{\sigma_A^2+\sigma_S^2}. \quad (8)$$

Thus, w_m is solely determined by the ratio σ_A^2/σ_S^2, which is one of the free parameters of the model. In the Appendix we show that the posteriors of the categories can be rewritten to

$$P\left(a_i|s,c_j\right) = P\left(c_j|a_i\right)\cdot\alpha_{i,j}\left(s\right) \quad (9)$$

and thus depend on cue reliability and a measurement-dependent factor $\alpha_{i,j}(s)$:

$$\alpha_{i,j}\left(s\right) = \frac{P\left(s|a_i\right)}{p_\mathrm{C}P\left(s|a_j\right)+\frac{1-p_\mathrm{C}}{n-1}\sum_{k\neq j}^n P\left(s|a_k\right)}. \quad (10)$$

Here we exploit the specific form of the cue reliability and assume the categories to be uniformly distributed. The marginalization over T results in a normal distribution $P(S\,|\,A)$ with

$$S|A \sim N\left(\mu_A,\sigma_S^2+\sigma_A^2\right). \quad (11)$$

Applying the assumption for the cue reliability to the posterior expectation, we finally have

$$E\left[T|s,c_j\right] = w_\mathrm{m}s + (1-w_\mathrm{m}) \times\left(p_\mathrm{C}\cdot\alpha_{j,j}\left(s\right)\cdot\mu_{a_j} + \frac{1-p_\mathrm{C}}{n-1}\sum_{i\neq j}^n \alpha_{i,j}\left(s\right)\cdot\mu_{a_i}\right). \quad (12)$$

The term within the large brackets is composed of the mean of the correct category weighted by the cue reliability and the weighted sum of all other category means.

The effect of this weighting is to select or suppress the correct category, depending on the cue reliability parameter p_C. The latter would correspond to a deliberately misleading symbolic cue. Furthermore, the influence of the symbolic cue is balanced by the probability of the measurement depending on the category, which appears in $\alpha_{i,j}(s)$.

In Feldman et al. (2009), the symbolic cue indicating the category is not provided, which corresponds to an uninformative symbolic cue. We can reflect this in our model by setting $P(c_j\,|\,a_i) = 1/n$ for any *i,j*. We show in the Appendix that this indeed removes the dependency of the category posterior on the symbolic cue, yielding

$$E\left[T|s\right] = w_\mathrm{m}s + (1-w_\mathrm{m})\sum_i^n P\left(a_i|s\right)\cdot\mu_{a_i}. \quad (13)$$

This corresponds to Eqs 10 and 11 in Feldman et al. (2009) for equal category variance.

The posterior of T is a Gaussian mixture distribution, whose mean is not necessarily equal to its mode. However, the Gaussian measurement likelihood typically dominates the posterior, because its variance is small compared to the combined variance of the prior distributions corresponding to the categories. This yields a near Gaussian posterior as illustrated in **Figure 4**.

Cue-combination model

Instead of assuming that the symbolic cue signifies a category of sensory stimuli, it can also be conceived as providing additional information about the location of the stimulus in the sensory dimension. Under this assumption, the target distance T is drawn from a single distribution

$$T \sim N\left(\mu_T,\sigma_T^2\right) \quad (14)$$

with the stimulus S being a noisy reading of T

$$S|T \sim N\left(T,\sigma_S^2\right). \quad (15)$$

The intuition behind the cue-combination model is that the same mechanism of multi-modal sensory fusion (e.g., Ernst and Banks, 2002), which the brain might use to combine different sensory modalities, is used to merge sensory and symbolic information. From an observer point of view, this requires an inference mechanism that maps the symbolic cue C to a continuous cue signal C_mp. We call this signal the mapped cue. This signal is then merged with the sensory signal S and prior T in the usual Bayesian fashion. From a generative point of view, this inference inverts the causal relationships assumed for the outside world (see **Figure 1C**). In particular, C_mp is discretized by a step function to yield C. Our update mechanism, described further below, learns to map each cue value c_i to a cue signal value c_mp that falls into the corresponding range. This corresponds to learning the thresholds of the step function. This mapping is deterministic, thus the cue signal becomes a known quantity, similar to actual observations. We can therefore derive the estimation step using C_mp only, leaving out C.

The cue signal C_mp has a likelihood function that corresponds to the average location and dispersion associated with the symbolic cue (see **Figure 4B**)

$$C_\mathrm{mp}|T \sim N\left(\mu_\mathrm{C}\left(T\right),\sigma_\mathrm{C}^2\right). \quad (16)$$

Note that C_mp depends on T in a more complex way than S, reflected by the non-linear mapping $\mu_\mathrm{C}(T)$ We treat the cue signal C_mp the same way as the observation S. The mapping of the symbolic cue to the cue signal depends on the value of C and is updated iteratively. This updating can be understood as learning or calibration of the symbolic cue values (see Iterative update).

The optimal estimate of the target distance T is provided by a sensory fusion of the stimulus, the cue signal, and the prior

$$P\left(T|S, C_{\mathrm{mp}}\right) \propto P\left(S|T\right) \cdot P\left(C_{\mathrm{mp}}|T\right) \cdot P\left(T\right). \quad (17)$$

With w_{m} as weight for the measurement s and w_{fu} as weight for the fused signal composed of mapped cue c_{mp} and measurement s, the mean for the posterior is computed as follows:

$$E\left[T|s, c_{\mathrm{mp}}\right] = \left(1 - w_{\mathrm{fu}}\right) \cdot \mu_T + w_{\mathrm{fu}} \cdot \left(\left(1 - w_{\mathrm{m}}\right) \cdot c_{\mathrm{mp}} + w_{\mathrm{m}} \cdot s\right). \quad (18)$$

The weights w_{fu} and w_{m} result from the variances of target, stimulus, and symbolic cue:

$$w_{\mathrm{fu}} = \frac{\sigma_T^2}{\sigma_T^2 + \sigma_{CS}^2} \qquad w_{\mathrm{m}} = \frac{\sigma_C^2}{\sigma_S^2 + \sigma_C^2}. \quad (19)$$

The combined variance σ_{CS}^2 of symbolic cue and stimulus is

$$\sigma_{CS}^2 = \frac{\sigma_C^2 \sigma_S^2}{\sigma_C^2 + \sigma_S^2}. \quad (20)$$

Note that in the indicies we wrote C instead of C_{mp} for brevity. A more detailed derivation of the expectation of the posterior is provided in the Appendix. In short, since prior, combined likelihood, and their product are Gaussians, the mean of the posterior is given by a weighted sum of prior mean and the weighted sum of mapped cue and measurement (see **Figure 4** right).

Iterative update

Prior experience as well as cue mapping are not available at the start of the experiment but need to be acquired and updated over the course of the trials. Such updating on a trial by trial basis can be achieved by a discrete Kalman filter updating internal states at each time step. In our case, the states correspond to the means of the two categories in case of the categorical model, to the means of the two symbolic cue likelihoods for the cue-combination model, and to the distance prior in case of the previously published basic iterative model (see Petzschner and Glasauer, 2011).

In both models, the symbolic cue is used to decide which category mean will be updated or which symbolic cue likelihood will be learned. The updating of the category means is an extension of our *basic iterative model* from one single category to multiple categories (see also Feldman et al., 2009). The iterative updating of the mean of the symbolic cue likelihood can be interpreted as learning the non-linear mapping of the symbolic cue to the stimulus dimension or as calibration of the symbolic cue in terms of a distance.

For Gaussian noise and linear dynamics, the Kalman filter yields an estimate of the current state. The current state is estimated based on the current observation and the estimate of the state at the previous time step, taking into account a deterministic temporal evolution of the state. The state x to be updated and the current measurement y at trial i are described by the system equations

$$\begin{aligned} x_i &= x_{i-1} + n_q \\ y_i &= x_i + n_r. \end{aligned} \quad (21)$$

The random variables n_q and n_r represent the process and measurement noise, which are assumed to be independent with Gaussian probability distributions $P(n_q) \approx N(0, q)$ and $P(n_r) \approx N(0, r)$. The temporal evolution of the state x defined by these equations can be seen as a random walk governed by the process noise. The measurement y is a noisy version of x.

For such a simple system, it can be shown that the difference equation system of the Kalman filter reduces to

$$\begin{aligned} k_i &= \frac{p_{i-1} + q}{p_{i-1} + q + r} \\ p_i &= k_i \cdot r \\ \hat{x}_i &= (1 - k_i) \cdot \hat{x}_{i-1} + k_i \cdot y_i \end{aligned} \quad (22)$$

with k_i being the Kalman gain, $\hat{x}_{i-1}$ and $\hat{x}_i$ being the *a priori* and *a posteriori* estimate of the state (e.g., a category mean) at trial i, and p_{i-1} the corresponding variance of that quantity. Note that it is evident from this equation that the Kalman gain k_i can be interpreted as weight of the measurement depending on measurement noise and the assumed random change of the estimated quantity, such as a category mean. The new estimate is thus a weighted sum of the previous estimate and the current measurement.

The update for the categorical model employs a Kalman filter for each category mean to be estimated, yielding equations indexed by j:

$$\mu_{a_j,i} = \left(1 - k_i^j\right) \cdot \mu_{a_j,i-1} + k_i^j \cdot s_i. \quad (23)$$

For two categories we consequently have two Kalman filters, one for each category mean. The variances σ_A^2 and σ_S^2 correspond to quantities p_i and r, respectively. Note that the ratio of the two variances only depends on the ratio q/r, which is one of the free model parameters.

The cue-combination model uses three Kalman filters to calibrate the two symbolic cue likelihoods and to update the prior for the target distance T using the same general form of update equations as described above.

$$c^j_{\mathrm{mp},i} = \left(1 - k_i^j\right) \cdot c^j_{\mathrm{mp},i-1} + k_i^j \cdot s_i \quad (24)$$

$$\mu_{T,i} = \left(1 - k_i^T\right) \cdot \mu_{T,i-1} + k_i^T \cdot s_i. \quad (25)$$

The calibration of the symbolic cue likelihoods yields the mapped cues used in the estimation.

Logarithmic stimulus representation

There is some indication that magnitudes are internally represented in the brain on a log-scale (Fechner, 1860; Dehaene, 2003; Jürgens and Becker, 2006; Stocker and Simoncelli, 2006; Durgin et al., 2009). In Petzschner and Glasauer (2011) we showed that defining a Bayes-optimal observer on log-scales leads to an elegant combination of Steven's power law with the Weber–Fechner law (Fechner, 1860; Stevens, 1961). The estimates in our models in the present work are again computed based on simplified logarithmic representations of the presented stimuli. In conjunction with that stands an additional parameter that can represent different optimal decision strategies in subjects. We shortly recap the idea here and refer to Petzschner and Glasauer (2011) for a detailed treatment. The logarithmic representation is given as

$$s = ln\left(\frac{d_{\mathrm{m}}}{d_0}\right) + n_{\mathrm{m}}. \tag{26}$$

The internal representation of the measurement s is computed as the natural logarithm of the measurement on linear scales, d_{m}. In the present work, d_{m} is given in virtual meters. To achieve a unit-less representation, d_{m} is normalized with the small constant $d_0 \ll 1$. The random variable n_{m} represents the normally distributed measurement noise $P(n_m) \approx N(0, \sigma_S^2)$.

The estimate x_{est}, corresponding to $E[T \mid s, c_j]$ for the categorical model and $E[T \mid s, c_{\mathrm{mp}}]$ for the cue-combination model, is a log-scale value. It is transformed back to a linear scale with

$$d_r = e^{x_{\mathrm{est}} + \Delta x} \cdot d_0. \tag{27}$$

The result is the linear scale reproduction d_r in virtual meters. We assume here that, apart from this transformation and possibly additional noise, the reproduction in subjects corresponds to the estimate.

The value Δx accounts for different decision strategies of the subjects. A decision strategy collapses the posterior distribution into a single value, the estimate, which is optimal in the sense that it minimizes the expected loss due to the deviation from the real value (the real distance in our case). Typical decision strategies use the mean, median, or mode of a distribution as optimal (loss-minimal) estimate, which correspond to three typical loss functions (Körding and Wolpert, 2004). While these values are equal for normal distributions, they are different in our case, since the normal distribution transfers into a log-normal distribution after back-transformation. For the log-normal distribution mean, median, and mode differ by a linear shift of x_{est}. Therefore, by introducing an additional parameter Δx in our models, we account for different types of loss functions. We call this parameter the shift term.

Model fit

To analyze how well our models explain the experimental results, we fitted their free parameters such that the difference between model output and subject responses was minimized.

The free parameters in the categorical model are the cue reliability p_{C}, the ratio σ_A^2/σ_S^2 of the noise in the target distances and the measurement noise, and the shift term Δx reflecting the loss function of the Bayesian estimator. The ratio σ_A^2/σ_S^2 determines the weight of the measurement w_{m} relative to the category priors. This weighting schema reflects how subjects may put more weight on whichever quantity has less variance.

The free parameters of the cue-combination model are the shift term Δx and two ratios. The first is the ratio of target distance noise to the combined noise in measurement and continuous cue signal, σ_T^2/σ_S^2. The second is the ratio of the noise in the cue signal to the measurement noise, σ_C^2/σ_S^2. Analogous to the categorical model, these ratios determine the relative weights w_{fu} and w_{m}, respectively. The first is the weight of the combined measurement and cue signal relative to the prior, the second the measurement weight relative to the cue signal.

The basic iterative model has two free parameters, the shift term Δx and the ratio of target distance noise to measurement noise, σ_T^2/σ_S^2. This ratio determines the measurement weight w_{m} of this model.

For the IR-C condition we fitted the category and cue-combination models to the responses of each single subject. That is, for each subject two sets of parameters were generated, corresponding to the two models. For the other two conditions, our models reduce to the iterative Bayesian estimation model (Petzschner and Glasauer, 2011), which we fitted in these cases. All models were fitted by minimizing the squared differences of model output and subject response in each trial using the Matlab function *lsqnonlin*.

The correct order of sample displacements over all trials in one condition was used as input to the models. Kalman filters in the models were initialized with the first observation, that is the first produced distance of the subject in the given condition.

To assess the precision of the fitted parameters, we estimated 95% confidence intervals of all parameters that were determined from the Jacobian of the parameter surface at the minimum using the Matlab function *nlparci*.

Model comparison

We compared the models' goodness of fit by comparing their coefficients of determination R^2. The coefficient of determination assesses the proportion of variability in the mean data that is accounted for by the respective model. To test for a significant difference in the R^2 of the two model fits across subjects we used the non-parametric Wilcoxon signed rank test (Matlab procedure *signrank*).

RESULTS

BEHAVIORAL DATA

In order to test the effect of an additional symbolic cue on the estimation of distances we used three experimental conditions. One condition tested the cue influence directly (IR-C condition) while the other two served as reference conditions for the extreme cases of the cue effect, i.e., ignoring the cue (IR-NC) or using the symbolic cue as perfectly reliable indicator for the stimulus range (BR-NC). The average results of all three conditions are presented in **Figure 5** (left side).

Differences between conditions can be assessed by comparing the estimation of overlapping samples, that is, displacements that were assigned to the "short" as well as to the "long" distribution. However, assigning distances to a short or long range should not only affect the overlapping distances, but the estimation and consequently the reproduction errors for *all* distances presented. Condition-dependent differences in distance reproduction should occur either due to the influence of short-term prior experience or induced by the symbolic cue.

Comparison of distance errors

The comparison of the distance reproduction error shows a main effect of *distance* [$F(9,38) = 136.2$, $p < 0.0001$] together with a highly significant interaction of *condition* and *distance* [$F(18,342) = 3.45$, $p < 0.0001$]. This interaction is due to a clear separation of error patterns between conditions, which can be seen in **Figure 6** where the differences between errors in the interleaved condition (IR-NC) to the other two conditions are shown. Note that in both conditions where the ranges were separated either temporally (BR-NC) or by the symbolic cue (IR-C), the errors in the low range correspond on average to overshoots, while the errors in the high range correspond to undershoots with respect to those in the interleaved condition without cue (IR-NC). This correspondence of error patterns also confirms that the symbolic cue causes changes in distance estimation analogous to those found during temporal dissociation of the two ranges. However the effect in the IR-C condition is not as strong as in the BR-NC condition. Separate *post hoc* rm-ANOVAS with only two conditions shows that for IR-C versus BR-NC this interaction vanishes [$F(9,171) = 1.82$, $p = 0.068$ n.s.], while it remains highly significant for IR-C and IR-NC [$F(9,171) = 3.36$, $p = 0.0008$]. Thus, while in IR-C and IR-NC all distance stimuli were the same in magnitude and order, the reproduced distances are clearly different, which shows that the symbolic cue was used by the subjects in a way very similar to exploiting the temporal separation of the two ranges in the BR-NC condition.

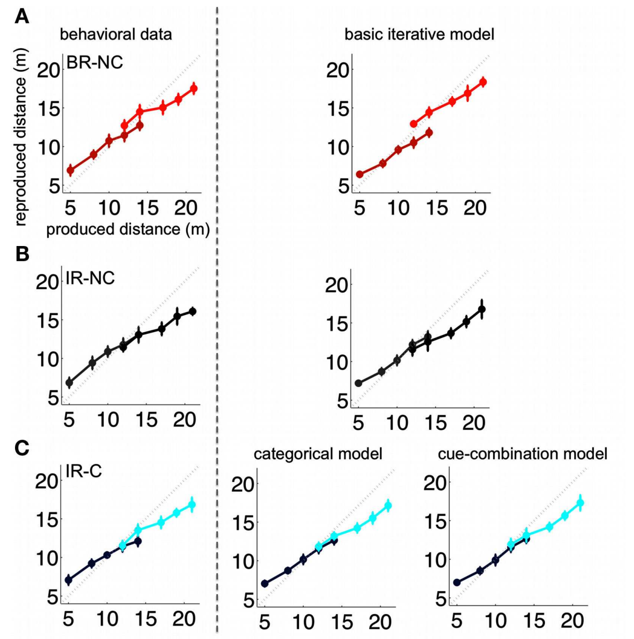

FIGURE 5 | Group mean of all subjects (left) and respective model predictions (right). The group mean corresponds to the mean taken over all subjects for the whole trial sequence. Models were accordingly fitted to the resulting "mean trial sequence." The rows **(A–C)** show the results for the three conditions BR-NC, IR-NC, and IR-C respectively. The IR-NC and BR-NC were fitted with the basic iterative model introduced in Petzschner and Glasauer (2011). Predictions for the cue condition IR-C were generated with the categorical as well as cue-combination model. Error bars depict the standard deviation of the reproduced distances across trials.

Overlapping samples

The results for the whole range of distances are also supported by the overlapping samples comparison, which reveals a significant interaction of *condition* × *range* (short/long) for all experimental conditions [$F(2,38) = 11.9$, $p = 0.0001$]. This implies a significant difference in the estimation of the two overlapping distances depending on the experimental condition (see also, **Figure 5**).

Separate ANOVAS with only two conditions revealed the individual relationships between the conditions. Differences in subjects' behavior based solely on temporal order were determined based on the comparison of the IR-NC and BR-NC conditions (for a detailed description, see Materials and Methods).

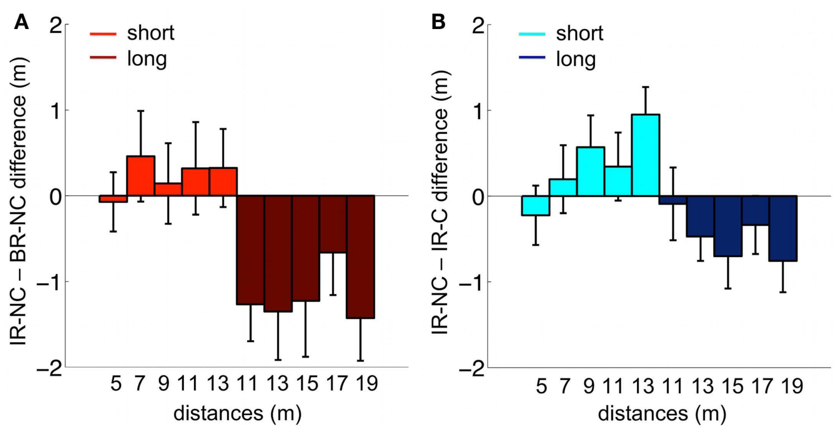

FIGURE 6 | Mean behavioral differences between conditions. (A) Difference between the IR-NC and the BR-NC conditions for the mean reproduced distances. **(B)** Difference between the IR-NC and the IR-C conditions for mean reproduced distances. The only difference between the two conditions was the symbolic cue. Colors code the "short" and "long" range of displacements respectively.

In analogy with previous results, we find a significant interaction of *condition* × *range* in the overlapping samples comparison [$F(1,19) = 26.5$, $p < 0.001$], which confirms that temporal order affects distance reproduction. By testing for the interaction between the IR-C and IR-NC condition, we assessed exclusively cue-based differences in subjects' behavior. Again we find a significant *condition* × *range* interaction for the overlapping samples comparison [$F(1,19) = 8.8$, $p < 0.01$]. To compare performance when the sample ranges were either separated by time or symbolic cue, we performed an rm-ANOVA for the IR-C and BR-NC condition. In this case we find no significant difference between conditions in the overlapping samples comparison [interaction: *condition* × *range*; $F(1,19) = 3.6$, $p > 0.05$ n.s]. Thus, as found above, the symbolic cue leads to a behavior that resembles the performance exhibited for presenting the stimuli in ranges separated by time as in the BR-NC condition.

The *post hoc* analysis of the individual conditions supports the results of the condition comparison. The rm-ANOVA of the overlapping samples comparison reveals a significant difference for the estimation of the overlapping samples comparison in the BR-NC condition [main effect: *range* ("short" vs. "long") $F(1,19) = 25.7, p < 0.001$] but no significant difference of the overlapping samples comparison in the IR-NC condition where no separation between the ranges was provided [main effect: *range* ("short" vs. "long") $F(1,19) = 1.3$, $p > 0.05$]. Finally, the symbolic cue in the IR-C condition caused a significant difference in behavior based on the assigned range [overlapping samples comparison: main effect: *range* ("short" vs. "long"); $F(1,19) = 9.3$, $p < 0.01$].

MODELING

Our results show that the symbolic cue significantly affects the reproduction of the stimuli in a way that is more similar to the behavior in the BR-NC condition than to the one in the IR-C condition. This raises the question how the knowledge about the symbolic cue is incorporated into the estimation process. We compare our two models by fitting them to the responses of each single subject and also to the mean responses over all subjects computed for the overall time course of trials, which we refer to as "group mean." **Figure 5** depicts this group mean and the group mean fits of our models.

Categorical model for condition IR-C

The categorical model assumes that the target distances presented in each trial stem from one of two categories, and that the symbolic cue informs about the given category in that trial. The three free parameters of this model, the cue reliability, the measurement weight, and the shift term, were estimated by a least squared fit (group mean fit, $R^2 = 0.92$: $p_C = 0.74$, $CI_{95\%} = [0.70\ 0.78]$; $w_m = 0.33$; $\Delta x = -0.04$, $CI_{95\%} = [-0.05\ -0.03]$; individual participants fit: $\overline{p_C} = 0.76 \pm 0.13$, $range = [0.57\ 1.00]$; $\overline{w_m} = 0.32 \pm 0.11$, $range = [0.06\ 0.48]$; $\overline{\Delta x} = -0.06 \pm 0.19$, $range = [-0.67\ 0.26]$). The shift terms were not normally distributed over all subjects (Lillifors test, $p = 0.02$). Yet they show a unimodal distribution with a peak close to the shift corresponding to choosing the median of the posterior distribution as an estimate.

Cue-combination model for condition IR-C

In contrast to the categorical model, the cue-combination model assumes that target distances are drawn from one underlying distribution and treats the symbolic cue as a second sensory input to the system. Its three free parameters are the measurement weight, the fusion weight, and the shift term. Analogous to the categorical model they were fit using a least squares method (group mean fit, $R^2 = 0.91$: $w_m = 0.38$; $w_{fu} = 0.55$; $\Delta x = -0.05$, $CI_{95\%} = [-0.07\ -0.03]$; individual participants fit: $\overline{w_m} = 0.39 \pm 0.10$, $range = [0.18\ 0.50]$; $\overline{w_{fu}} = 0.54 \pm 0.16$, $range = [0.25\ 0.81]$; $\Delta x = -0.07 \pm 0.19$, $range = [-0.67\ 0.25]$). As in the case of the categorical model, shift terms fitted for the cue-combination model were not normally distributed over all subjects (Lillifors test, $p = 0.03$), yet showed a unimodal distribution with a peak near the shift corresponding to the median.

Basic iterative model for conditions IR-NC and BR-NC

If the symbolic cue is abandoned, the two new models reduce to the basic iterative model. For comparison, we fitted this model on the two non-cue conditions IR-NC and BR-NC. The model has two free parameters, which have been fitted for each of these two conditions individually (IR-NC group mean fit: $w_m = 0.33$; $\Delta x = -0.05$, $CI_{95\%} = [-0.07\ -0.03]$; IR-NC individual participants fit: $\overline{w_m} = 0.33 \pm 0.13$, $range = [0.03\ 0.48]$; $\overline{\Delta x} = -0.07 \pm 0.22$, $range = [-0.67\ 0.37]$; BR-NC group mean fit: $w_m = 0.34$; $\Delta x = -0.04$, $CI_{95\%} = [-0.06\ -0.02]$; BR-NC individual participants fit: $\overline{w_m} = 0.33 \pm 0.09$, $range = [0.14\ 0.49]$; $\overline{\Delta x} = -0.04 \pm 0.12$, $range = [-0.29\ 0.26]$).

Model comparison

To compare the categorical and cue-combination model, we computed R^2 values for individual participant fits (see **Figure 7**) in the IR-C condition (categorical model: $\overline{R^2} = 0.54 \pm 0.15$, $range = [0.31\ 0.88]$; cue-combination model: $\overline{R^2} = 0.54 \pm 0.15$, $range = [0.24\ 0.88]$) as well as for the group mean fits (categorical model: $R^2 = 0.92$; cue-combination model: $R^2 = 0.91$). In the other two conditions without a symbolic cue, our existing Bayesian estimator model shows similar goodness of the individual participant fits (IR-NC: $\overline{R^2} = 0.45 \pm 0.18$, $range = [0.05\ 0.72]$; BR-NC: $\overline{R^2} = 0.51 \pm 0.27$, $range = [-0.40\ 0.85]$). And as in the IR-C condition the group mean fit turns out to be better (IR-NC: $R^2 = 0.87$; BR-NC: $R^2 = 0.88$) than the individual estimates.

In comparing the goodness of fit of the categorical and the cue-combination model (non-parametric Wilcoxon signed rank test), no significant difference between the two models could be found ($p > 0.45$). We also tested whether the small differences of the subject-by-subject R^2 values that can be seen in **Figure 7** are related to the subjects' response biases and variances. However, we could not find any significant correlations (Spearman ranks test, $p > 0.13$).

DISCUSSION

The context in which a stimulus occurs can contain additional relevant information about the stimulus itself. It is thus advantageous to combine all types of available information, in order to use the composite as an estimate of the stimulus. Here we demonstrate that this fusion of information takes place in distance estimation

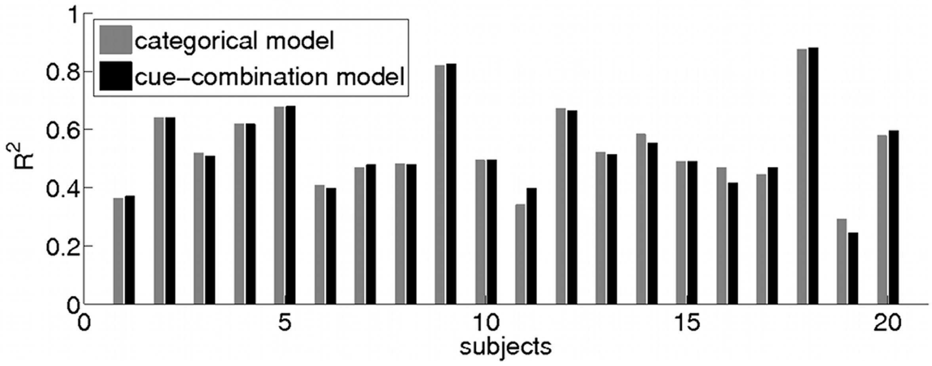

FIGURE 7 | Model Comparison. Bar plot of individual R^2 values of the model fit for the categorical (gray) and cue-combination model (black) to each subjects' behavior (1–20) in the IR-C condition. A comparison of the goodness of fit of the categorical and the cue-combination model revealed no significant difference between the two models.

by path integration, where subjects incorporated prior experience and abstract information provided by a symbolic cue into their current estimate of displacement. We proposed two generative Bayesian models that describe this fusion of information based on two distinct assumptions – categorization and cue-combination.

CUE-BASED RANGE AND REGRESSION EFFECTS

The influence of the symbolic cue on distance estimation behavior was assessed by comparing the cue condition (IR-C) to two reference conditions. Both mimicked the two possible extreme cases of cue usage. The no cue condition BR-NC tested two overlapping ranges of stimuli that were blocked in time, in order to change the respective prior experience of subjects and mimicked the case in which the pre-cueing by the words "short" or "long" would lead to a full separation of stimuli into two groups of events or categories. The IR-NC condition combined these two ranges to a single distribution of distances. The order and magnitude of stimuli was exactly the same as in the IR-C condition, thus replicating the cue condition for the case where the symbolic cue would be fully ignored.

In all three experimental conditions we observed a tendency to bias toward certain displacements, also referred to as regression effect (Hollingworth, 1910). In the no-cue conditions BR-NC and IR-NC the bias depended on the respective underlying sample distribution and could be explained by incorporation of short-term prior experience into the current estimate of displacements, as shown in our previous study (Petzschner and Glasauer, 2011). The behavior in the cue condition IR-C did not resemble that of the IR-NC condition although the order and size of sample displacements was the same. It was rather reflecting the behavior observed for two distinct sample ranges in the BR-NC condition, even though the effect was smaller.

Thus, the bias in the cue condition cannot be explained exclusively by the use of prior experience. This led to the question of how the additional symbolic cue information is processed. One possible explanation comes from the studies on categorization effects (Huttenlocher et al., 1991; Cheng et al., 2010). If there is uncertainty in the stimulus metric, then information about stimulus categories can be incorporated into the estimation process (Huttenlocher et al., 1991; Feldman et al., 2009). In our case, the symbolic cue could cause a sorting of stimuli into categories such that the expectation about the upcoming stimulus varies depending on whether subjects assume the stimulus to be drawn from the "short" or "long" category. We elaborated on this idea in the categorical model.

Another possible explanation, which we pursued in our cue-combination model, comes from a different field of research – multi-modal sensory cue-combination (Ernst and Banks, 2002; Ernst and Bülthoff, 2004). Similar to our findings, von Hopffgarten and Bremmer (2011) showed in a recent study on self-motion reproduction that subjects are capable of learning an abstract relationship between a novel cue and the stimulus and exploit that information to improve their performance. In their study, the frequency of a simultaneous auditory signal indicated movement speed and was used by the subjects to improve self-motion reproduction. Their study provides evidence that subjects learned the initially unknown frequency-velocity mapping provided by the auditory cue, comparable to the mapping of the symbolic cue to distance in our present experiment. Von Hopffgarten and Bremmer argued that the observed behavior could be interpreted by "sensory combination" (Ernst and Bülthoff, 2004), where the auditory input served as an additional, non-redundant cue.

CATEGORICAL MODEL

The categorical model is based on the assumption that the stimulus comes from one of two distinct, but perhaps overlapping, categories of stimuli, each represented by its own probability distribution (Feldman et al., 2009). Accordingly the symbolic cue provides information about the respective category. The order of events in this generative model is as follows (**Figure 1B**): (1) the category is chosen, (2) the information about the category is provided as symbolic cue, and (3) the stimulus is drawn from the distribution corresponding to the category. Note that the symbolic cue does not necessarily provide reliable information about the category. Hence, the prediction of the symbolic cue for a respective category is not always correct. The model represents this uncertainty with a trial-independent probability that we refer to as cue reliability.

Since the categories are unknown, they have to be learned from the symbolic cue values ("short" and "long" in the present experiment) and the stimulus presentation. Note that the semantic interpretation of the cue values is not sufficient to determine the categories, since the cue values do not specify the ranges; they only denote an order within the presented stimuli, i.e., that a "short" distance probably is shorter than a "long" one. Learning is achieved by iterative Bayesian estimation analogous to Petzschner and Glasauer (2011). Our categorical model is thus an extension of the model of Feldman et al. (2009) to explain the so-called perceptual magnet effect in speech perception. In contrast to their model, where no pre-cueing was done and the categories were assumed to be fixed, our model provides the symbolic cue values as additional uncertain information about the category and allows learning of the category means during the course of the experiment. The variance of the prior distributions could also be learned during the experiment (Berniker et al., 2010). However, in the present study we assume that it is, apart from an initialization phase, constant throughout the experiment. For other categorization tasks, such as understanding of speech, it has been proposed that the learning of weighting of acoustic cues for categorization might take place during development (Toscano and McMurray, 2010).

The combination of categorical information with the measured stimulus value was also proposed in a model by Huttenlocher et al. (1991) for estimating spatial location. In their model categorical information is used in two distinct ways. First the remembered stimulus measurement is weighted with categorical prototype information and second the resulting estimates are constrained to fall within the category boundaries. In our model estimates are not artificially restricted to certain boundaries, even though the weighting with the learned mean of the respective category will bias them toward this mean. Hence, our estimation process explains the tendency to bias toward the category means, which is reported in a variety of psychophysical studies. This *central tendency bias*, *schema*, or *range effect*, causes a tendency of estimates to be biased toward the category they where assigned to (Hollingworth, 1910; Johnson and Vickers, 1987; Cheng et al., 2010).

The category model can be extended to an arbitrary number of categories. However, introducing new categories or new cue values during the experiment would not only require learning of that category, but also re-computing of the relative weights of the other categories. In other words, a new category or new cue value should directly affect the other categories.

In the present work the number of categories is predefined and given by the number of cue values, but under many other circumstances this is not the case. Recent work (e.g., Lucas and Griffiths, 2010) addresses the question of how we determine the number of categories in the context of learning of causal structures. While this is not required in the present study, our cue-combination model, which is independent of the number of cues, may well be capable of dynamically adapting to new cue values added during the course of the experiment. This could be considered as a weaker form of structural learning.

CUE-COMBINATION MODEL

In contrast to the categorical model, the cue-combination model assumes that the stimulus comes from one continuous range of stimuli and the pre-cueing provides additional evidence about where in this range the current stimulus can be found. This idea is similar to common models in sensory cue-combination, where the sensory inputs from a common source are fused in order to build a unified percept of its origin (Ernst and Banks, 2002; Körding et al., 2007). In terms of a generative model, the order of events in the cue-combination model is as follows (**Figure 1C**): (1) the stimulus is drawn from the underlying distribution, and (2) the symbolic cue is determined from this stimulus by some mapping. In our current implementation, this mapping is assumed to be probabilistic. Therefore, a large stimulus is assumed to cause the respective symbolic cue value in most of the cases, but at some occasions it can also lead to the other cue value. Since the mapping between stimulus and cue value is not prespecified, it has to be learned over the course of the experiment. This is achieved by iteratively adapting the mean of the likelihood function associated to each symbolic cue value. In addition to the unknown mapping, the underlying stimulus distribution is learned during the experiment (as in Petzschner and Glasauer, 2011).

A more intuitive explanation of the cue-combination model is provided from the observer point of view. Given the stimulus and an additional corresponding cue one aims to combine these two sources of information in an optimal manner. This would require that the cue can be related to a certain displacement value. This can be achieved by learning the relation between the current stimulus distance and the respective cue on a trial by trial basis. We refer to this process as mapping in the present model.

The mapping of the symbolic cue values to the stimulus dimension does not require knowledge about the possible number of cue values. Rather, the adaptation is similar to cue calibration, e.g., learning the transformation between one stimulus dimension and another (Burge et al., 2010; Zaidel et al., 2011). Thus, in contrast to the category model, adding another symbolic cue value during the experiment would not require a change in the mapping of the previously presented cues. This makes the model more flexible to changes than the categorical model.

MODEL COMPARISON

Interestingly, the results of the categorical and cue-combination model are very similar, although the underlying assumptions are substantially different. The categorical model is based on an intuitive assumption about how the stimuli presented to the subjects are generated: it assumes that there are two distinct categories, from which the stimuli are drawn. This corresponds, for example, to the categories in speech production, where a certain syllable is produced or understood based on a distinct category. The cue-combination model does not assume such an underlying structure, but rather treats the symbolic cue as additional modality. Consequently, the cue-combination model is more flexible to changes in cueing while, at least for our experiment, being equally powerful in explaining the data compared to the categorical model. The main reason for the similar performance of both models is, apart from the experimental setting, the iterative updating of the "meaning" associated with the symbolic cue, which leads to very similar sources of information regarding the range of stimuli denoted by the cues. This information is, in both models, weighted by reliability either in form of a variance associated with the symbolic cue or a probability that the symbolic cue is accurate. Thus, both models can fairly well describe the behavior observed in our experiments: our participants used the symbolic cue, they were able to associate them with the stimulus magnitude, but they did not completely trust them, as evidenced by the difference between the IR-C and BR-NC conditions.

Similarly, both models would also have performed equally well in predicting the two outcomes of cue usage mimicked in the IR-NC and BR-NC condition (**Figure 3**). If the information provided by the cue would not be incorporated into the estimate of the displacements this would have resulted in a cue weighting close to zero reflected by either a cue reliability that is close to 0.5 in the categorical model or a very high cue variability in the cue-combination model. An extreme cue usage, as mimicked by the BR-NC condition, would have an opposite effect on the respective parameters.

This raises the question under which circumstances the two models would make different predictions. One major difference between the two estimation processes lies in the different means of incorporating prior knowledge. Consider **Figures 4A,B**. While the categorical model uses a combined prior that is driven by the occurrence of all respective cues, the cue-combination incorporates a global prior that only depends on short-term prior experience of the stimuli independently of the corresponding cues. We used the parameters derived from the fit of the experimental data in this paper to test how these differences could lead to differing predictions of the cue-combination model and categorical model under specific circumstances.

Imagine the case where the two ranges are clearly separated. Due to the influence of the experience driven prior the cue-combination model would be biased by the full range of all displacements causing a global underestimation in the high range and an overestimation of the short range of stimuli. In contrary, the combined prior in the categorical model would show two discrete peaks at the center of the respective categories and thus lead to an estimate that is, for both ranges, centered closer to the single category means. However, this strong bias in the cue-combination model would only become evident if we assume a constant variance of the prior. If the variance of the prior is also updated on a trial by trial basis (Berniker et al., 2010; Verstynen and Sabes, 2011), both models would again become similar.

Yet another case in which both models differ should become obvious when omitting the cue in some catch trials. The cue-combination model would then reduce to the basic model and rely on the global unimodal prior, thus resulting in a global tendency to the mean (Petzschner and Glasauer, 2011). In contrast, the categorical model works with two prior distributions even when the cue is missing. In that case, our categorical model reduces to the category model of Feldman et al. (2009) and would exhibit the perceptual magnet effect, which biases the reproduction toward the category means.

Another difference should be observable in cases where the presentation of the cue and stimulus is not fully randomized. Consider a case where the "long" cue is repeatedly presented in a block with a long displacement. The cue-combination model would show a quick adaption of the global prior to these long displacements, which would result in reproduction values biased toward the long displacements. The categorical model would predict a much weaker adaption to the block as it still incorporates all potential cues, the long as well as the short ones. In that respect the categorical model seems to have a longer memory and less flexibility for fast changes.

Finally, both models become mathematically equivalent for a specific parameter combination. This is the case if the variance of the global prior in the cue-combination model becomes large enough and the cue reliability in the categorical model is set to unity. That is, for the categorical model we have to set $p_C = 1$ in Eqs 10 and 12. For the cue-combination model, we set $\sigma_T^2 = \infty$ in Eq. 20 so that $w_{fu} = 1$. Then the conditional expectations for both models (Eqs 12 and 19) become equivalent.

ITERATIVE LEARNING AND CALIBRATION

The no cue conditions demonstrated that subjects incorporated knowledge about the stimulus history into their current estimate of displacement. We model this iterative learning of prior knowledge by a discrete Kalman filter. In our previous work we showed that this online update of prior experience explains small variation in the data that a fixed prior could not account for (Petzschner and Glasauer, 2011). That humans are indeed capable of learning not only the mean but also the variance of an experience driven prior distribution was also recently shown (Berniker et al., 2010; Verstynen and Sabes, 2011).

The significant influence of the symbolic cue on the behavioral performance in the cue condition further shows that most subjects also included this information into their estimate of displacement. As mentioned above, the semantic interpretation of the cue values was not sufficient to allow such a fusion of cue values and sensory stimulus. Thus, subjects had to learn how to associate both. The cue-combination model interprets this learning as a mapping of the cue values onto the stimulus dimension. That an abstract, even arbitrary, mapping of different types of information can be acquired during the course of an experiment was also shown by Ernst (2007). In his study subjects were trained with stimuli that usually are unrelated in the world, such as the luminance of an object and its stiffness, but which in the experiment had a fixed mapping. He showed that subjects learned to integrate the two formerly unrelated signals, similar to the mapping in our models. Calibration is, however, not only necessary between unrelated stimulus dimensions, but also between those which are normally related, such as visual and vestibular signals indicating self-motion. A recent study could show that such a calibration is independent of the reliability of the cue (Zaidel et al., 2011), which corresponds to the learning or calibration implemented in our models.

CONCLUSION

Natural human action and perception profits from the incorporation of contextual information. We show that in addition to the previously found influence of prior experience, humans are also capable of using non-metric information, in the form of a symbolic cue, for their estimate of displacement, even if the mapping of the symbolic cue onto the stimulus dimension has to be acquired during the experiment. Two substantially different models of how this information enters the estimation process led to equally good fits to the experimental data. This result sheds new light on the modeling of behavioral problems such as categorization, cue-combination, and trial-to-trial dependencies.

ACKNOWLEDGMENTS

This research was supported by the BMBF (grants IFB 01EO0901 and BCCN 01GQ0440). We thank Virginia Flanagin and Paul MacNeilage for valuable comments on a previous version of the manuscript.

REFERENCES

Adams, W. J., Graf, E. W., and Ernst, M. O. (2004). Experience can change the 'light-from-above' prior. *Nat. Neurosci.* 7, 1057–1058.

Angelaki, D. E., Gu, Y., and DeAngelis, G. C. (2009). Multisensory integration: psychophysics, neurophysiology, and computation. *Curr. Opin. Neurobiol.* 19, 452–458.

Battaglia, P. W., Jacobs, R. A., and Aslin, R. N. (2003). Bayesian integration of visual and auditory signals for spatial localization. *J. Opt. Soc. Am. A. Opt. Image Sci. Vis.* 20, 1391–1397.

Berniker, M., Voss, M., and Körding, K. (2010). Learning priors of Bayesian computations in the nervous system. *PLoS ONE* 5, e12686. doi:10.1371/journal.pone.0012686

Burge, J., Girshick, A. R., and Banks, M. S. (2010). Visual-haptic adaptation is determined by relative reliability. *J. Neurosci.* 30, 7714–7721.

Cheng, K., Spetch, M. L., and Hoan, A. (2010). Categories and range effects in human spatial memory. *Front. Psychol.* 1:231. doi:10.3389/fpsyg.2010.00231

Davidoff, J., Davies, I., and Roberson, D. (1999). Colour categories in a stone-age tribe. *Nature* 398, 203–204.

Dehaene, S. (2003). The neural basis of the Weber-Fechner law: a logarithmic mental number line. *Trends Cogn Sci. (Regul. Ed.)* 7, 145–147.

Durgin, F. H., Akagi, M., Gallistel, C. R., and Haiken, W. (2009). The precision of locomotor odometry in humans. *Exp. Brain Res.* 193, 429–436.

Ernst, M. O. (2007). Learning to integrate arbitrary signals from vision and touch. *J. Vis.* 7, 1–14.

Ernst, M. O., and Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* 415, 429–433.

Ernst, M. O., and Bülthoff, H. H. (2004). Merging the senses into a robust percept. *Trends Cogn. Sci. (Regul. Ed.)* 8, 162–169.

Etcoff, N. L., and Magee, J. J. (1992). Categorical perception of facial expressions. *Cognition* 44, 227–240.

Fechner, G.T. (1860). *Elemente der Psychophysik*. Leipzig: Breitkopf and Härtel.

Feldman, N. H., Griffiths, T. L., and Morgan, J. L. (2009). The influence of categories on perception: explaining the perceptual magnet effect as optimal statistical inference. *Psychol. Rev.* 116, 752–782.

Hollingworth, H. L. (1910). The central tendency of judgment. *J. Philos. Psychol. Sci. Methods* 7, 461–469.

Huttenlocher, J., Hedges, L. V., and Duncan, S. (1991). Categories and particulars: prototype effects in estimating spatial location. *Psychol. Rev.* 98, 352–376.

Jacobs, R. A. (1999). Optimal integration of texture and motion cues to depth. *Vision Res.* 39, 3621–3629.

Johnson, J., and Vickers, Z. (1987). Avoiding the centering bias or range effect when determining an optimum level of sweetness in lemonade. *J. Sens. Stud.* 2, 283–292.

Jürgens, R., and Becker, W. (2006). Perception of angular displacement without landmarks: evidence for Bayesian fusion of vestibular, optokinetic, podokinesthetic, and cognitive information. *Exp. Brain Res.* 174, 528–543.

Körding, K. P., Beierholm, U., Ma, W. J., Quartz, S., Tenenbaum, J. B., and Shams, L. (2007). Causal inference in multisensory perception. *PLoS ONE* 2, e943. doi:10.1371/journal.pone.0000943

Körding, K. P., and Wolpert, D. M. (2004). The loss function of sensorimotor learning. *Proc. Natl. Acad. Sci. U.S.A.* 101, 9839–9842.

Langer, M. S., and Bülthoff, H. H. (2001). A prior for global convexity in local shape-from-shading. *Perception* 30, 403–410.

Liberman, A. M., Harris, K. S., Hoffman, H. S., and Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *J. Exp. Psychol.* 54, 358–368.

Lucas, C. G., and Griffiths, T. L. (2010). Learning the form of causal relationships using hierarchical Bayesian models. *Cogn. Sci.* 34, 113–147.

Müller, H. J., Reimann, B., and Krummenacher, J. (2003). Visual search for singleton feature targets across dimensions: stimulus- and expectancy-driven effects in dimensional weighting. *J. Exp. Psychol. Hum. Percept. Perform.* 29, 1021–1035.

Petzschner, F. H., and Glasauer, S. (2011). Iterative Bayesian estimation as an explanation for range and regression effects: a study on human path integration. *J. Neurosci.* 31, 17220–17229.

Stevens, S. S. (1961). To honor Fechner and repeal his law: a power function, not a log function, describes the operating characteristic of a sensory system. *Science* 133, 80–86.

Stocker, A. A., and Simoncelli, E. P. (2006). Noise characteristics and prior expectations in human visual speed perception. *Nat. Neurosci.* 9, 578–585.

Stone, J. V., Kerrigan, I. S., and Porrill, J. (2009). Where is the light? Bayesian perceptual priors for lighting direction. *Proc. Biol. Sci.* 276, 1797–1804.

Toscano, J. C., and McMurray, B. (2010). Cue integration with categories: weighting acoustic cues in speech using unsupervised learning and distributional statistics. *Cogn. Sci.* 34, 434–464.

Verstynen, T., and Sabes, P. N. (2011). How each movement changes the next: an experimental and theoretical study of fast adaptive priors in reaching. *J. Neurosci.* 31, 10050–10059.

Vincent, B. (2011). Covert visual search: prior beliefs are optimally combined with sensory evidence. *J. Vis.* 11, 25.

von Hopffgarten, A., and Bremmer, F. (2011). Self-motion reproduction can be affected by associated auditory cues. *Seeing Perceiving* 24, 203–222.

Zaidel, A., Turner, A. H., and Angelaki, D. E. (2011). Multisensory calibration is independent of cue reliability. *J. Neurosci.* 31, 13949–13962.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 11 May 2012; accepted: 24 July 2012; published online: 13 August 2012.

Citation: Petzschner FH, Maier P and Glasauer S (2012) Combining symbolic cues with sensory input and prior experience in an iterative Bayesian framework. Front. Integr. Neurosci. ***6****:58. doi: 10.3389/fnint.2012.00058*

APPENDIX

CATEGORICAL MODEL

Our categorical model shall infer the target distance T from given measurement s and cue c_j. Here we derive the posterior distribution over T from known distributions, along with its conditional expectation. The posterior is given as

$$P(T|S,C) = \sum_i^n P(T, a_i|S,C). \tag{A1}$$

We have to marginalize over the categories because they are unknown. The key idea is now to express the posterior as a weighted sum over distributions of which we can easily compute the expectation.

We first factorize the posterior within the sum to obtain a distribution of which we can easily compute its expectation. As it turns out, it is the posterior of T given S and the category.

$$P(T, a_i|S,C) = P(T|a_i, S, C) \cdot P(a_i|S,C). \tag{A2}$$

Due to model assumptions concerning the factorization, the posterior of T does not depend on C once the category A is given. The full joint distribution for our model, according to our assumptions (**Figure 1B**), factorizes as follows:

$$P(T, A, S, C) = P(S|T)\,P(T|A)\,P(C|A)\,P(A). \tag{A3}$$

We use this factorization of the full joint marginalize out T.

$$P(A,S,C) = \int P(C|A)\,P(S|t)\,P(t|A)\,P(A)\,dt = P(C|A)\,P(A)\int P(S,t|A)\,dt = P(C|A)\,P(A)\,P(S|A). \tag{A4}$$

Then, by applying Bayes' theorem we, see

$$P(T|a_i,S,C) = \frac{P(T,a_i,S,C)}{P(a_i,S,C)} = \frac{P(S|T)\,P(C|a_i)\,P(T|a_i)\,P(a_i)}{P(C|a_i)\,P(a_i)\,P(S|a_i)} = \frac{P(S|T)\,P(T|a_i)\,P(a_i)}{P(S|a_i)\,P(a_i)} = \frac{P(S|T,a_i)\,P(T,a_i)}{P(S,a_i)} = P(T|S,a_i). \tag{A5}$$

We, see that all factors depending on C cancel each other out. The posterior thus does not depend on C, given A. Following the definition of the expectation, we now have

$$E\left[T|s,c_j\right] = \int t \sum_i^n P\left(t, a_i|s,c_j\right)dt = \int t \sum_i^n P(t|s,a_i)\cdot P\left(a_i|s,c_j\right)dt \tag{A6}$$

with s and c_j being the known distance measurement and the known cue. We, see that the posterior of the category $P(a_i \mid s,c_j)$ does not depend on t, which means we can do the following reordering. We pull t into the sum, exchange sum and integral and pull out the posterior of the category:

$$E\left[T|s,c_j\right] = \sum_i^n P\left(a_i|s,c_j\right)\int t\,P(t|s,a_i)\,dt. \tag{A7}$$

The integral expresses the expectation of the category-dependent posterior of T. Thus we have

$$E\left[T|s,c_j\right] = \sum_i^n P\left(a_i|s,c_j\right)E\left[T|s,a_i\right]. \tag{A8}$$

We will now express the posterior of the category, the weights in the above sum, through known distributions.

$$P(A|S,C) = \frac{P(A,S,C)}{P(S,C)} = \frac{P(C|A)\,P(A)\,P(S|A)}{P(S,C)} = \frac{P(C|A)\,P(A)\,P(S|A)}{\sum_k P(C|a_k)\,P(S|a_k)\,P(a_k)}. \tag{A9}$$

We reused the factorization of $P(A,S,C)$ derived above. In the denominator we use it to marginalize out A. All distributions appearing in the result are known from our model assumptions, except $P(S \mid A)$. It results from integrating over t:

$$P(S|A) = \int P(S, t|A)\, dt = \int P(S|t)\, P(t|A)\, dt. \tag{A10}$$

Since both $P(S \mid t)$ and $P(t \mid A)$ are normally distributed, integrating over t yield the following distribution:

$$S|A \sim N\left(\mu_A, \sigma_S^2 + \sigma_A^2\right). \tag{A11}$$

Remember that we assume equal variances for all categories.

A special case of this model is if no cue is present. This corresponds to a case where all cues appear with equal probability independently of the given category, $P(c_j \mid a_i) = 1/n$, and thus tell us nothing about the category. This leads to the posterior of A becoming independent of C:

$$P(A|S, C) = \frac{1/nP(A)\, P(S|A)}{\sum_k 1/nP(S|a_k)\, P(a_k)} = \frac{P(A)\, P(S|A)}{P(S)} = P(A|S). \tag{A12}$$

With the posterior of the category now being independent of C, the expectation for T given measurement and cue in Eq. 33 above reduces to Eq. 29 in Feldman et al. (2009). The Feldman model is thus a special case of our model.

It remains to compute the expectations of the category-dependent posteriors, $E[T \mid s,a_i]$. These posteriors result from standard Bayesian fusion of the likelihood for S and the prior $P(T \mid A)$, as we have shown above:

$$P(T|S, A) = \frac{P(S|T)\, P(T|A)}{P(S|A)} = \frac{P(S|T, A)\, P(T|A)}{P(S|A)}. \tag{A13}$$

Since both the likelihood $P(S \mid T)$ and the prior $P(T \mid A)$ are normally distributed, the posterior is normally distributed with mean and variance

$$\mu_i = \frac{s\sigma_A^2 + \mu_{a_i}\sigma_S^2}{\sigma_S^2 + \sigma_A^2} \qquad \sigma_i^2 = \frac{\sigma_S^2\sigma_A^2}{\sigma_S^2 + \sigma_A^2}. \tag{A14}$$

That allows us to write the expectation as

$$\begin{aligned} E\left[T|s, c_j\right] &= \sum_i^n P\left(a_i|s, c_j\right) \frac{s\sigma_A^2 + \mu_{a_i}\sigma_S^2}{\sigma_S^2 + \sigma_A^2} = \sum_i^n P\left(a_i|s, c_j\right) \left(\frac{\sigma_A^2}{\sigma_S^2 + \sigma_A^2} s + \frac{\sigma_S^2}{\sigma_S^2 + \sigma_A^2} \mu_{a_i}\right) \\ &= \sum_i^n P\left(a_i|s, c_j\right) w_{\mathrm{m}} s + \sum_i^n P\left(a_i|s, c_j\right)(1 - w_{\mathrm{m}})\mu_{a_i} = w_{\mathrm{m}} s + (1 - w_{\mathrm{m}}) \sum_i^n P\left(a_i|s, c_j\right) \mu_{a_i}. \end{aligned} \tag{A15}$$

This can be further simplified using another two of our model assumptions. First, we assume that, *a priori*, categories are uniformly distributed, that is $P(a_i) = 1/n$. Second, we assume that the correct cue appears with some probability $P(c_j \mid a_i) = p_C$, $j = i$, while the remaining wrong cues appear with equal probabilities $P(c_j \mid a_i) = (1 - p_C)/(n-1)$. First we rewrite the posterior of the category:

$$\begin{aligned} P\left(a_i|s, c_j\right) &= \frac{P\left(c_j|a_i\right) \cdot P(a_i)\, P(s|a_i)}{\sum_k P\left(c_j|a_k\right) P(s|a_k)\, P(a_k)} \\ &= P\left(c_j|a_i\right) \cdot \frac{1/nP(s|a_i)}{p_C P\left(s|a_j\right) 1/n + \frac{1-p_C}{n-1} \sum_{k \neq j} P(s|a_k)\, 1/n} \\ &= P\left(c_j|a_i\right) \cdot \frac{P(s|a_i)}{p_C P\left(s|a_j\right) + \frac{1-p_C}{n-1} \sum_{k \neq j} P(s|a_k)} = P\left(c_j|a_i\right) \cdot \alpha_{i,j}(s). \end{aligned} \tag{A16}$$

This results in

$$E\left[T|s, c_j\right] = w_{\mathrm{m}} s + (1 - w_{\mathrm{m}}) \sum_i^n P\left(c_j|a_i\right) \alpha_{i,j}(s) \mu_{A_i}. \tag{A17}$$

Then we can further rewrite by again replacing $P(c_j \mid a_i)$ to get

$$E\left[T|s,c_j\right] = w_m s + (1-w_m)\left(p_C \alpha_{j,j}(s)\mu_{a_j} + \sum_{i\neq j}^{n} \frac{1-p_C}{n-1}\alpha_{i,j}(s)\mu_{a_i}\right)$$
$$= w_m s + (1-w_m)\left(p_C \alpha_{j,j}(s)\mu_{a_j} + \frac{1-p_C}{n-1}\sum_{i\neq j}^{n} \alpha_{i,j}(s)\mu_{a_i}\right). \quad \text{(A18)}$$

CUE-COMBINATION MODEL

Our assumptions about the (conditional) distributions of target distance T, stimulus S and mapped cue C_{mp} (the cue signal) lead to the following factorization of the model's full joint probability:

$$P\left(T,S,C_{mp}\right) = P(S|T)\,P\left(C_{mp}|T\right)P(T). \quad \text{(A19)}$$

From this, the posterior follows immediately:

$$P\left(T|S,C_{mp}\right) = \frac{P\left(T,S,C_{mp}\right)}{P\left(S,C_{mp}\right)} = \alpha P(S|T)\,P\left(C_{mp}|T\right)P(T) \propto P(S|T)\,P\left(C_{mp}|T\right)P(T). \quad \text{(A20)}$$

We can, see that the posterior density function is, apart from the proportionality factor $\alpha = 1/P(S,C_{mp})$, a product of Gaussians. First, we combine the two likelihood density functions. Following the product rule for Gaussians, this product yields a Gaussian with the following parameters:

$$\mu_{CS} = \frac{s\sigma_C^2 + c\sigma_S^2}{\sigma_S^2+\sigma_C^2} \qquad \sigma_{CS}^2 = \frac{\sigma_S^2\sigma_C^2}{\sigma_S^2+\sigma_C^2}. \quad \text{(A21)}$$

In the indices we write C instead of C_{mp} for brevity and better readability, e.g., μ_{CS} instead of $\mu_{C_{mp}S}$. A problem is that the (unknown) mean of the symbolic cue likelihood, $\mu_C(T)$, depends non-linearly on T. However, we may assume that this dependence is approximately linear. Remember that our model receives as input a discrete cue, which steers a calibration process (implemented by a Kalman filter), whose output in each trial is then interpreted as additional measurement c_{mp} of the mapped cue. This output of the calibration process closely follows the stimuli from either the long or the short range, depending on the discrete cue. The ranges themselves do not change and therefore a normal distribution with fixed $\mu_C(T)$ (after a short calibration period) can approximate the dispersion of the c_{mp} values.

The product of the combined likelihood density function with the density of the prior $P(T)$ is again a product of Gaussians resulting in a Gaussian. The density of the posterior is thus Gaussian with parameters

$$\mu_{posterior} = \frac{\mu_T\sigma_{CS}^2 + \mu_{CS}\sigma_T^2}{\sigma_{CS}^2+\sigma_T^2} \qquad \sigma_{posterior}^2 = \frac{\sigma_{CS}^2\sigma_T^2}{\sigma_{CS}^2+\sigma_T^2}. \quad \text{(A22)}$$

According to standard probability theory, the expectation of the posterior is then given as

$$E\left[T|s,c_{mp}\right] = \mu_{posterior} = \frac{\sigma_{CS}^2}{\sigma_{CS}^2+\sigma_T^2}\mu_T + \frac{\sigma_T^2}{\sigma_{CS}^2+\sigma_T^2}\mu_{CS} = (1-w_{fu})\,\mu_T + w_{fu}\mu_{CS}$$
$$= (1-w_{fu})\,\mu_T + w_{fu}\frac{s\sigma_C^2 + c_{mp}\sigma_S^2}{\sigma_S^2+\sigma_C^2} = (1-w_{fu})\,\mu_T + w_{fu}\left(\frac{\sigma_C^2}{\sigma_S^2+\sigma_C^2}s + \frac{\sigma_S^2}{\sigma_S^2+\sigma_C^2}c_{mp}\right) \quad \text{(A23)}$$
$$= (1-w_{fu})\,\mu_T + w_{fu}\left(w_m s + (1-w_m)\,c_{mp}\right) = (1-w_{fu})\,\mu_T + w_{fu}\left((1-w_m)\,c_{mp} + w_m s\right)$$

with the weights

$$w_{fu} = \frac{\sigma_T^2}{\sigma_{CS}^2+\sigma_T^2} \qquad 1-w_{fu} = 1-\frac{\sigma_T^2}{\sigma_{CS}^2+\sigma_T^2} = \frac{\sigma_{CS}^2+\sigma_T^2-\sigma_T^2}{\sigma_{CS}^2+\sigma_T^2} = \frac{\sigma_{CS}^2}{\sigma_{CS}^2+\sigma_T^2}$$
$$w_m = \frac{\sigma_C^2}{\sigma_S^2+\sigma_C^2} \qquad 1-w_m = 1-\frac{\sigma_C^2}{\sigma_S^2+\sigma_C^2} = \frac{\sigma_S^2+\sigma_C^2-\sigma_C^2}{\sigma_S^2+\sigma_C^2} = \frac{\sigma_S^2}{\sigma_S^2+\sigma_C^2} \quad \text{(A24)}$$

4 DISCUSSION AND CONCLUSION

> There are things known and there are things unknown, and in between are the doors of perception.
>
> — Aldous Huxley

In this thesis systematic effects in human path integration behavior are studied. We find that human estimation of distances and turning angles exhibits behavioral characteristics that are omnipresent in human magnitude estimation (range effect, regression effect, scalar variability). We provide evidence that these characteristics can be explained by the use of a-priori assumptions for the current estimate of displacement. Finally, we show that this influences performance in way that is predicted by iterative Bayesian estimation models that implement the Weber-Fechner law and are capable of incorporating and updating a-priori knowledge.

4.1 A BAYESIAN APPROACH TO MAGNITUDE ESTIMATION

While this work focussed on human path integration, it is likely that the iterative Bayesian approach presented can provide a comprehensive account for the systematic characteristics in magnitude estimation behavior. Since, the characteristics explained by the model are ubiquitous across magnitude estimation[1]. Three key assumptions are essential to the approach's success in modeling this behavior:

- **Logarithmic Scaling**: Measured magnitudes are internally represented either as normal distributions with constant

1 The parallels in magnitude estimation behavior across different modalities can be nicely visualized in the comparison between figure 2 in Jazayeri and Shadlen [2010], which depicts human performance in time estimation for different test ranges and figure 3 in Petzschner and Glasauer [2011], which depicts nearly indistinguishable behavioral characteristics for human distance and turning angle estimation.

standard deviation on a logarithmic scale or as normal distributions with linear increase of standard deviation on a linear scale.

- **Bayesian Inference**: Estimated magnitudes are the result of Bayesian inference based on the current measurement and a-priori assumptions.
- **Online Update of Prior Knowledge**: The a-priori assumptions are based on past experience and/or additional information about the stimulus and are constantly updated.

In the following three subsections, I will discuss the general validity and implications of each assumption of the Bayesian approach in the context of magnitude estimation.

4.1.1 Magnitude Estimation on Logarithmic Scales

The *logarithmic scaling* assumption is rooted in the early findings of Weber [1850] who discovered that the just-noticeable difference between two stimulus magnitudes increases with the stimulus size. The accompanying linear increase in the standard deviation of estimates with the mean is termed constant coefficient-of-variation or scalar variability, corresponding to the two competing mathematical explanations of the phenomenon: Either magnitudes are, as proposed by Fechner, represented on a logarithmic scale with constant noise leading to an proportional increase in standard deviation with the mean on linear scales; or they are represented on linear scales with a corresponding linear increase in the standard deviation of the internal noise[2] [Cantlon et al., 2009; Feigenson et al., 2004; Gibbon, 1977]. In other words, either we have a tendency to compress the range of magnitudes onto a logarithmic scale such that larger magnitudes are represented with a lower resolution; or our estimates simply become noisier with increasing stimulus magnitude [Cantlon et al., 2009; Dehaene et al., 2008]. It was pointed out that the two theories are hardly discriminable, as the '[...] *linear coding with scalar variability and logarithmic coding with fixed variability lead to the same metric of number similarity, and therefore to the same behavior.*' [Dehaene, 2001]. Some support for the logarithmic coding hypothesis has recently come from animal studies. Nieder and Miller [2003] tested both behavioral and neuronal representations of numerical information in the prefrontal cortex of rhesus monkeys. They found a nonlinear compression of numerical scale, as it would

2 This corresponds to the usual assumption about signal-dependent noise in the motor system.

be postulated by the psychophysical laws, both on a neuronal level as well as on a behavioral level [Dehaene, 2003; Nieder and Miller, 2003; Nover et al., 2005].
For our iterative Bayes model, both coding types would be feasible and should in principle yield similar results. We chose to use the logarithmic transformation of magnitudes to account for the linear increase in standard deviation, because it represents an implementation of the Weber-Fechner law and leads to simplified computations of the Bayesian estimate.

4.1.2 Magnitude Estimation as Bayesian Inference

A core assumption of our approach is that systematic behavior in magnitude estimation can be described by Bayesian inference. This does not, however, imply that magnitude estimation is mechanistically Bayesian estimation. There is currently an ongoing debate among neuroscientists, related to this issue. The controversy is about to what extent perception can be regarded as Bayesian inference[3].
Some proponents of Bayes promote the conception that perception *is* Bayesian inference [Friston and Stephan, 2007; Knill and Pouget, 2004; Knill and Richards, 1996]. Support for this view comes from the remarkable successful of Bayesian estimation at explaining all kinds of perceptual processes including motion perception [Stocker and Simoncelli, 2006], visual illusions [Weiss et al., 2002], sensory-motor learning [Körding et al., 2004], sensory cue-combination [Ernst and Banks, 2002], and specific cases of magnitude estimation [Jazayeri and Shadlen, 2010; Jürgens and Becker, 2006]. These studies motivated conclusions that *'human perception is close to the Bayesian optimal suggesting the Bayesian process may be a fundamental element of sensory processing'* [Körding and Wolpert, 2006] or, analogously, that there are *'myriad ways in which human observers behave as optimal Bayesian observers'* [Knill and Pouget, 2004].
Generalizing statements like these brought opponents to the stage. They question whether the Bayesian approach can effectively explain perceptual processes and have put forward some important criticisms [Bowers and Davis, 2012a; Colombo and Series, 2012; Jones and Love, 2011]:
First, even if humans behave according to Bayesian estimation, it does not mean that this algorithm is implemented in the brain [Bowers and Davis, 2012a]. To prove the implementation of the

3 In fact, there is currently a heated debate on this issue in the *Psychological bulletin* [Bowers and Davis, 2012a,b; Griffiths et al., 2012].

Bayesian framework in the brain, one would need to show that populations of neurons can both represent probability distributions and carry out the necessary Bayesian computations. There have been various theoretical attempts to model Bayesian computations on a neuronal level [Beck et al., 2008; Denève et al., 2007; Ma and Pouget, 2008], but there is still little experimental evidence from neuronal recordings for an actual implementation [Fetsch et al., 2012].

De facto, only a subgroup of Bayes proponents have this strong *theoretical* view[4] [Bowers and Davis, 2012a]. The majority uses the approach as a measure of performance. Their *methodological* view does not imply an actual implementation of Bayesian computations but gives teleological explanations of *why* it may make *sense* to show a specific behavior [Bowers and Davis, 2012a; Griffiths et al., 2012].

The work presented here is of this methodological type. On the basis of behavioral data it cannot be distinguished whether humans carry out Bayesian magnitude estimation[5] or any alternative computation at the algorithmic level. However, the behavioral findings can still provide insights to features of the underlying processes, inasmuch as any algorithmic account of magnitude estimation should factor in the incorporation of a-priori evidence that results in the characteristic behavior.

The second criticism concerns the flexibility of the approach. Bowers and Davis [2012a] argue that *'there are too many arbitrary ways that priors, likelihoods, utility functions, etc., can be altered in a Bayesian theory post hoc'*. The claim can be illustrated by a Gedankenexperiment: Bayes' theorem yields the probability of a hypothesis to be true, given evidence and a-priori assumptions. It is, thus, highly depended on each of these components. Consider, for example, that we want to determine whether it will be raining on a particular day by looking at the sky. Our subjective belief (posterior) would depend on both the likelihood (*'It is cloudy outside.'* or *'It is sunny.'*) and our prior (e.g. the general rain probability in Munich or Seattle). Consequently, a change in either of the components would greatly affect our weather forecast. Similarly, a Bayes model that does not define both the likelihood and the prior based on actual experimental evidence, but uses priors (or likelihoods) which lead to a best fit of the data, cannot provide mechanistic insights. A typical

4 Terms are chosen according to Bowers and Davis [2012a]. For an argument against this devision, see Griffiths et al. [2012].

5 In the sense that the brain represents priors and likelihoods, and their respective reliability in form of probability distributions and multiplies these distributions to come up with an a-posteriori estimate.

example for such an *arbitrary* adjustment of priors is a Bayesian approach to speed perception [Weiss et al., 2002]. In that study, the authors suggest that humans have a prior that represents a preference for low speeds. They chose this prior as it best explained the behavioral data, but admitted that there is *'no direct evidence (either from first principles or from empirical measurements) that this assumption is correct.'* [Weiss et al., 2002].
Yet, the criticism described above is not inherent to Bayes models, but holds for any model that depends on the adjustment of free parameters. Often, there is no justification for the choice of the free parameters. The reason why this has become such an important issue in the case of Bayesian estimation is that these models usually come up with an interpretation of prior, likelihood, or utility function, while many alternative modeling approaches have no need to justify their free parameters [Griffiths et al., 2012].
Fortunately, often the prior and likelihood are constrained by fundamental knowledge [Vilares and Kording, 2011]. In the study in chapter 2, the prior and likelihood are derived from the preceding and the current test stimuli alone [Petzschner and Glasauer, 2011]. Not every teleological explanation provides, however, a correct mechanistic explanation of the behavior. The second study in chapter 3 reveals that, with additional information sources being available, different generative models can account equally well for the same behavior [Petzschner et al., 2012].

4.1.3 Online Update of a-priori Knowledge

What distinguishes the present framework from most other Bayesian approaches to human magnitudes estimation is the continuous update of prior knowledge.
Priors in Bayesian models are often fixed quantities or *natural statistics*, such as the prior that 'light comes from above' (see figure 3b) [Howard et al., 2009; Mamassian et al., 2002]. And even if it is shown that the mean and weighting of the prior can change due to experience [Adams et al., 2004; Tassinari et al., 2006], there are only few accounts of how these priors are learned and adapted [Berniker et al., 2010; Verstynen and Sabes, 2011].
In magnitude estimation, stimuli can be spread over wide ranges and the behavior is highly dependent on these ranges (e.g. range effect). Therefore, it seems necessary to implement an adaptive component, that allows us to change our a-priori assumptions. The two Bayesian approaches to magnitude estimation that are the most closely related to this work avoid an adaptive compo-

nent [Jazayeri and Shadlen, 2010; Jürgens and Becker, 2006]. Instead, they use task-specific default priors to model the behavior. These priors are located somewhere in the middle between the largest and smallest stimulus magnitude of the respective test range. As we have shown, such a default value can account for the range and regression effect, but fails to explain additional variability due to the order in which stimuli are presented (see chapter 1.2, *Sequential Effects*). In addition, they lack an explanation of how this task-specific knowledge is acquired, as subjects cannot have the knowledge of the full stimulus range at hand.
The strength of the framework presented in this thesis is that the online incorporation of prior experience explains a number of systematic behavioral characteristics, without assuming any explicit knowledge of the underlying stimulus range. Current estimates are biased by the preceding trials, which results automatically in a shift of the responses towards the center of the underlying sample distribution. This explains how the regression effect can already be observed in the first few trials [Stevens and Greenbaum, 1966] and in experiments where each subject performs only a single judgement for a specific magnitude [Stevens and Poulton, 1956]. In addition, an online-updated prior provides the basis for sensorimotor learning without feedback.
Notably, a-priori knowledge is not restricted to prior experience. The study in chapter 3 has shown that abstract explicit knowledge (provided by a written, verbal cues) influences the estimation process similar to experience. This suggests that humans can exploit and incorporate all kinds of information sources for perceptual decision making [Michel and Jacobs, 2008; Petzschner et al., 2012].

4.2 TOWARDS A UNIFYING ACCOUNT OF MAGNITUDE ESTIMATION

So far I have focussed on the Bayesian approach as a modeling framework for magnitude estimation behavior. Here, I will discuss the implications of our work for the interpretation of the classical psychophysical laws and its relation to other unifying principles in magnitude estimation.

4.2.1 Implications for the Psychophysical Laws

Bayes links Weber-Fechner and Stevens. The most well-known principles in magnitudes estimation were probably proposed by Fechner and Stevens [Fechner, 1860; Stevens, 1961]. Both of them believed that a simple mathematical relationship exists between the intensity of a stimulus and the subjective sensation that it causes. Furthermore, this relationship can be derived from the stimulus-response function in magnitude estimation. Although their methodological approaches as well as their conclusions relating sensation to physical magnitude were fundamentally different, there have been a number of attempts to unify the two approaches [Ekman, 1964; MacKay, 1963].
The *Bayesian* approach presented in this thesis reveals a new unexpected link between the Weber-Fechner and Stevens' power-law. When we assume that the Weber-Fechner law holds and coding of magnitudes takes place on logarithmic scales [Fechner, 1860], then the proposed Bayesian estimation takes also place on logarithmic scales. Hence, to reproduce a magnitude on linear scales the logarithmic estimate needs to be back-transformed. This transformation yields a power-law dependence between the responded magnitude R and the stimulus intensity I of the form[6]

$$R = c \cdot I_p^{w_p} I^{w_s} \tag{4.1}$$

where w_p and w_s are the respective weighting of prior intensity I_p and the current stimulus intensity I.
This closely resembles Stevens' power-law, where the response R is assumed to be a direct reflection to the subjective sensation S [Stevens, 1961]:

$$S = c \cdot I^n \tag{4.2}$$

A Bayesian estimation on logarithmic scales thus results in a power-law relation on linear scales. The exponent of this *Bayesian* power function is, in contrast to Stevens' empirical exponent n, derived from a concrete theoretical framework.

Log versus power-function. How can we interpret this link between logarithmic and power-law? Let us consider the different cases in which we observe a power or a logarithmic function.
Logarithmic stimulus-response relationships are commonly observed when subjects are asked to discriminate between stimulus magnitudes. Stevens referred to these methods as *indirect*

6 Nomenclature for equation 9 in chapter 2 is adjusted to that of Stevens' power-law in section 1.1.1. The constant c in equation 4.1 and 4.2 reflects different values.

scaling. The power-law, in contrast, is usually observed in *direct scaling* methods, where magnitude estimates are made on a ratio scale [Stevens, 1956]. Examples of the latter case include magnitude judgments based on assigning numbers in proportion to the sensation evoked (*magnitude estimation*[7]), matching the magnitude of a stimulus in one modality to the magnitude of another modality (*cross-modality matching*), or matching stimuli of the same modality (*uni-modal matching*[8]).

Stevens believed that the *direct scaling* methods yield the *stimulus transformation function* relating the stimulus to the subjective sensation, $S = f_1(I)$. Therefore, he concluded that the power function, empirically obtained from direct scaling methods, provides a measure of subjective sensation.

A number of scientists, in particular Shepard [1981], already pointed out that there is a flaw in Stevens' description. The subjective sensation itself is a hidden variable and accordingly cannot be measured directly. What is measured, is the response to the stimulus intensity and this response is linked to the subjective sensation via a second transform, the *response transformation function*, $R = f_2(S)$. Accordingly, the measured stimulus-response relationship, which in Stevens' case corresponds to a power function, is reflected by a third function $R = f_3(I) = f_2[f_1(I)]$. Stevens' conclusion that this power function allows for a direct measure of subjective sensation would therefore imply that f_2 is linear. But as Shepard and others correctly pointed out, this has not been proved to be the case [Attneave, 1962; Gescheider, 1988; MacKay, 1963; Shepard, 1981].

The *Bayesian* approach, in fact, proposes that there are different transformation steps in the process of magnitude estimation, and even more important that these transformation steps can explain the different stimulus-response relationship observed with direct and indirect scaling methods. After a logarithmic *stimulus transformation*, the measured stimulus and a-priori knowledge are combined to yield an *estimate* of the magnitude, still on logarithmic scales. In order to give a response on an absolute scale (e.g. the stimulus scale) a back-transformation (*response transformation function*) of the estimate is necessary that leads to a power-law dependence between stimulus and response (equation 4.1). This last transformation is only required if subjects have to make a response on an absolute scale. If, however, the response requires a mere comparison of two stimuli, then the estimate on logarithmic scales should be sufficient to solve the task.

7 Here, the term corresponds to the method proposed by Stevens.
8 The production-reproduction paradigm used to test path integration throughout this thesis can be considered as a type of unimodal matching.

In other words, a power-function is predicted in all cases where a back-transformation is necessary, that is, when subjects are asked to adjust or reproduce a magnitude on its own physical scale or that of another modality (including the number scale), as in the case of Stevens' *direct scaling* methods. If the response, however, does not require a back-transformation into the stimulus space, e.g. in a discrimination task, where the response is 'longer'-'shorter' or 'bigger'-'smaller', then one would observe a logarithmic stimulus-response relationship as predicted by the Weber-Fechner law and observed with *indirect scaling* methods. Accordingly, the stimulus-response function observed depends on the transformations required to solve the task at hand. The view further implies that the response in any magnitude estimation task - direct or indirect - will always reflect an estimate of magnitude and not subjective sensation.

A re-interpretation of Stevens' power-law exponents. As outlined briefly above, the exponents of the power function in the Bayes approach are derived from a concrete mathematical framework. Therefore, Stevens' fitted power-law exponents can be re-interpreted as the weighting of prior and measurement and consequently as an uncertainty associated with the magnitude measurements (compare equation 4.1 and 4.2).
The first evidence for a link between uncertainty and power-law exponents comes from a number of studies that reported an increase of the regression effect with increasing task difficulty [Kappauf, 1975; Sheldon, 1973; Stevens and Greenbaum, 1966; Thomas, 1949]. The iterative Bayesian estimation model explains the regression towards the center of the underlying sample range by the incorporation of the immediate prior experience. The weighting of prior experience, and consequently the regression, should increase if the sensory information provided is experienced as being less reliable. This the case when the task difficulty increases.
The second evidence comes from the range effect itself. If we assume that stimuli are represented on logarithmic scales, then their respective uncertainty increases with the stimulus magnitude (scalar variability), without increasing the difficulty of the task itself. This leads to a stronger regression towards the center for larger stimulus ranges, known as the range effect [Teghtsoonian, 1973; Teghtsoonian and Teghtsoonian, 1978].

Limits of the Bayesian approach. The Bayesian approach should prove applicable to a wide range of magnitude estimation data, for which Stevens determined a power-law relationship - with

one important exception. The exponent of the Bayes model is determined by a relative weighting of measurement and prior. It can, therefore, not be bigger than one. Nevertheless, Stevens found modalities with exponents that exceed one, e.g. the estimation of pain from electric shock [Stevens, 1960b].
However, this does not necessarily disqualify the Bayesian approach for these modalities, as the fitted exponents might actually not yield an absolute modality-specific value. The most common methods for determining the exponents are magnitude estimation[9] and cross-modal matching. Both require the transformation from one modality to the scale of another modality (in the case of magnitude estimation the number scale). This matching might yield fitted exponents that differ from the modality-specific ones [Attneave, 1962; Ekman, 1964]: If we assume that two given modalities yield a logarithmic dependence such that

$$S_1 = c_1 + n_1 \cdot \ln(I_1) \qquad S_2 = c_2 + n_2 \cdot \ln(I_2) \tag{4.3}$$

and the task is to match one modality with the other (e.g. to match a number to the loudness of a tone), than setting $S_1 = S_2$ yields

$$I_1 = e^{\frac{c_2 - c_1}{n_1}} \cdot I_2^{n_2/n_1} \tag{4.4}$$

Consequently, the exponent obtained from cross-modality matching and magnitude estimation might be a fraction of the modality-specific exponents[10] n_1 and n_2 [Attneave, 1962; Wagenaar, 1975].
Our iterative Bayesian estimation model can be extended in a similar way to cross-modal matching. Here, we would assume that the matching takes place on logarithmic scales. The modality to be matched S_1 is thus adjusted to the estimate[11] of the experienced modality $\hat{S}_2$, such that $\hat{S}_2 = n \cdot S_1$. This would yield

$$I_1 = e^{\frac{c}{n}} \cdot I_2^{w_{s,2}/n} \tag{4.5}$$

where n is a constant, that represents a linear scaling transformation from one modality to the other on logarithmic scales and

9 Here, the term corresponds again to the method proposed by Stevens.

10 It is not necessary to assume that the modalities yield a logarithmic relationship, a similar relation is obtained if we assume that they yield a power-law relation, $S = c \cdot I^n$, then the analog matching yields: $c_1 \cdot I_1^{n_1} = c_2 \cdot I_2^{n_2}$ and consequently $I_1 = \sqrt[n_1]{\frac{c_2}{c_1}} \cdot I_2^{n_2/n_1}$ [Attneave, 1962]. In all equations, c refers to different constant terms.

11 $\hat{S}_2$ reflects already the weighted Bayesian estimate (see chapter 2, equation 3).

$w_{s,2}$ is the weighting of the stimulus intensity I_2.
Accordingly, if the exponent of the matching modality is smaller that one, then the fitted exponent would yield a value that is larger than the modality-specific value and also larger than one. This holds true independent of whether we assume a logarithmic, a power-function or a Bayesian estimation stage for the matching. In other words, it is possible that Stevens' exponent are larger than one, although the modality-specific exponents are smaller than one. In these cases, the Bayesian approach would still be valid.

4.2.2 Magnitude Estimation across Modalities

Subject-specific Variability. Part of the fascination for magnitude estimation is driven by the consistency in the behavior across modalities. The incorporation of prior experience can explain some of this consistency, but not the subject-dependent correlations observed in chapter 2. There, we separately tested the estimation of two distinct magnitudes, turning angles and distances. A comparison of the model predictions revealed that subjects put nearly equal weights on angular and linear displacement measurements and that this correlation in weighting was significant for the whole group of subjects. This finding seems trivial at first glance, because the virtual environment did not change and one would assume that subjects received a similar sensory input in both tasks. In fact, however, a rotational movement yields a completely different pattern of optic flow than linear motion. A recent study found an age-related increase of thresholds for translational motion, but no effect on radial flow processing, suggesting specialized neuronal processing mechanisms for the different types of optic flow [Billino et al., 2008]. Therefore, the correlation in behavior turns out to be rather surprising. If it is not the global input, then where does the correlation come from?
The weighting of prior experience and current measurement is determined by their associated uncertainty. This uncertainty could be a task-specific parameter that is not directly related to the global, but possibly to local optic flow detection. Otherwise, it could also be a subject-dependent parameter reflecting similar processing mechanisms for magnitude estimation across modalities.
Support for the latter hypothesis is provided by studies of magnitude estimation which reported a subject-dependent correla-

tion in the Stevens' power-law exponents for circle size, number, and line length or weights, tastes, and smells [Jones and Marcus, 1961; Robinson, 1976]. An extension to Stevens' power-law was proposed that includes a subject-dependent variable in the exponent to capture this correlation [Jones and Marcus, 1961]. This factor enters the exponent at the same location as the weights in the Bayesian approach. Therefore this parameter could analogously reflect some kind of subject-dependent uncertainty.

A common metric of magnitudes in the brain? The subject-dependent correlations and a variety of other studies linking space, time, and quantities (e.g. [Casini and Macar, 1997; Dehaene et al., 1998; Glasauer et al., 2009]) have provided evidence for a generalized magnitude system across modalities. For instance, Dehaene et al. [2008] concluded that *'mathematical objects may find their ultimate origin in basic intuitions of space, time, and number that have been internalized through millions of years of evolution in a structured environment and that emerge early in ontogeny, independently of education'*. Similar ideas have been proposed by Walsh [2003]. In his *Theory Of Magnitude* (ATOM), he suggests that space, quantity, and time are linked by *'a common metric for action'* and that the *'apparent specializations for time, space, and quantity develop from a single magnitude system operating from birth'*. The area in the brain where this common processing of magnitudes is assumed take place is the parietal cortex [Bueti and Walsh, 2009].
The problem in testing ideas like this is that it often is tricky to compare modalities, like sound, pressure, or time as they are measured in different arbitrary units [Poulton, 1968]. The Bayesian approach provides a unit-less measure of behavior in form of an individual weighting of prior and measurement, thus allowing for a direct comparison between different types of magnitudes. The unit-less measure can be used to reveal similarities in the estimation of magnitudes across modalities, not only on a behavioral level, but also on a neuronal level, making the approach potentially relevant for the analysis of neuroimaging studies.

4.3 CONCLUDING REMARKS

The convergence of experimental findings on human magnitude estimation behavior across modalities and tasks encourages me to believe that a unifying approach exists for human magnitude estimation. The framework presented in this thesis can provide the basis for such an approach. It re-interpretes some of the classical findings in human magnitude estimation. It relates the two seemingly contradicting psychophysical laws (log and power function) and brings together the characteristic effects (regression effect, range effect, within-subject consistencies) in a single mathematical framework. Furthermore, context-dependent variability in behavior becomes a tractable and interpretable problem and the individual weighting can be used as a unit-less measure of subject-specific behavior.

This work also opens many research avenues. So far, the framework has only been applied to human path integration. Although, it seems applicable to a wide range of magnitude estimation data, its performance remains to be tested for other modalities. Furthermore, the conclusion that characteristic magnitude estimation behavior is the result of an incorporation of a-priori knowledge, leads to a number of related questions. Is this incorporation of knowledge task-dependent, i.e. is information only used when an action is required? Also, the current iterative Bayesian model for path integration incorporates that a given state-of-the-world changes on a very slow time scale. The prediction would, therefore, be that some of the characteristic errors in path integration are driven by the fact that most experiments simulate a very rapidly changing environment. What would consequently happen if the experiment tested more natural, slowly changing stimuli? I believe that some of these, and related issues can be assessed solely based on behavioral studies.

An overarching understanding of the processes underlying magnitude estimation will, however, require the contribution of a wide range of technologies, including electrophysiology and neuroimaging. In particular, we currently have very little knowledge on how the mechanistic processes and different transformation stages of magnitude estimation are realized at the neuronal level. An area that, I believe, might play a key role in this process is the parietal cortex, as this brain region is associated with both, the processing of time, magnitude, and number and the coding of prior expectation and evidence.

BIBLIOGRAPHY

Adams W J, Graf E W and Ernst M O. Experience can change the 'light-from-above' prior. Nature Neuroscience, 7(10):1057–1058, 2004.

Aiba T S and Stevens S S. Relation of brightness to duration and luminance under light- and dark-adaptation. Vision Research, 4(7):391–401, 1964.

Attneave F. Perception and related areas. In Koch S, editor, Psychology: A study of a science, pages 619–659. McGraw-Hill, New York, 1962.

Baird J C. Sensation and judgement: complementarity theory of psychophysics. Lawrence Erlbaum Associates, Mahwah, NJ, 1997.

Barlow H B. A method of determining the overall quantum efficiency of visual discriminations. Journal of Physiology, 160:155–168, 1962.

Beck J M, Ma W J, Kiani R, Hanks T, Churchland A K, Roitman J, Shadlen M N, Latham P E and Pouget A. Probabilistic population codes for Bayesian decision making. Neuron, 60(6):1142–1152, 2008.

Berniker M and Körding K. Bayesian approaches to sensory integration for motor control. Wiley Interdisciplinary Reviews: Cognitive Science, 2(4):419–428, 2011.

Berniker M, Voss M and Körding K. Learning priors for Bayesian computations in the nervous system. PLoS ONE, 5(9):e12686, 2010.

Billino J, Bremmer F and Gegenfurtner K R. Differential aging of motion processing mechanisms: evidence against general perceptual decline. Vision Research, 48(10):1254–1261, 2008.

Bowers J S and Davis C J. Bayesian just-so stories in psychology and neuroscience. Psychological Bulletin, 138(3):389–414, 2012a.

Bowers J S and Davis C J. Is that what Bayesians believe? reply to Griffiths, Chater, Norris, and Pouget (2012). Psychological Bulletin, 138(3):423–426, 2012b.

Bueti D and Walsh V. The parietal cortex and the representation of time, space, number and other magnitudes. Philosophical Transactions of the Royal Society: Biological Sciences, 364(1525):1831–1840, 2009.

Cantlon J F, Cordes S, Libertus M E and Brannon E M. Comment on "log or linear? Distinct intuitions of the number scale in Western and Amazonian indigene cultures". Science, 323(5910):38, 2009.

Casini L and Macar F. Effects of attention manipulation on perceived duration and intensity in the visual modality. Memory & Cognition, 25:812–818, 1997.

Colombo M and Series P. Bayes in the brain-on Bayesian modelling in neuroscience. The British Journal for the Philosophy of Science, 63:697–723, 2012.

Cross D V. Sequential dependencies and regression in psychophysical judgments. Perception & Psychophysics, 14(3):547–552, 1973.

Dehaene S. Substracting pigeons: logarithmic or linear? Psychological Science, 12(3):244–246, 2001.

Dehaene S. The neural basis of the Weber-Fechner law: a logarithmic mental number line. Trends in Cognitive Sciences, 7(4):145–147, 2003.

Dehaene S, Dehaene-Lambertz G and Cohen L. Abstract representations of numbers in the animal and human brain. Trends in Neurosciences, 21(8):355–361, 1998.

Dehaene S, Izard V, Spelke E and Pica P. Log or linear? Distinct intuitions of the number scale in Western and Amazonian Indigene cultures. Science, 320(5880):1217–1220, 2008.

Denève S, Duhamel J R and Pouget A. Optimal sensorimotor integration in recurrent cortical networks: a neural implementation of Kalman filters. The Journal of Neuroscience, 27(21):5744–5756, 2007.

Doya K, Ishi S, Pouget A and Rao R P N. Bayesian brain: Probabilistic approaches to neural coding. MIT Press, Massachusetts, 2007.

Ekman G. Is the power law a special case of Fecher's law? Perceptual & Motor Skills, 19:730, 1964.

Ekman G, Hosman B, Lindman R, Ljungberg L and Akesson C A. Inter-individual differences in scaling performance. Perceptual & Motor Skills, 26:815–823, 1968.

Ernst M O and Banks M S. Humans integrate visual and haptic information in a statistically optimal fashion. Nature, 415(6870):429–433, 2002.

Fechner G T. Elemente der Psychophysik. Breitkopf und Härtel, Leipzig, 1860.

Feigenson L, Dehaene S and Spelke E. Core systems of number. Trends in Cognitive Sciences, 8(7):307–314, 2004.

Fetsch C R, Pouget A, DeAngelis G C and Angelaki D E. Neural correlates of reliability-based cue weighting during multisensory integration. Nature Neuroscience, 15(1):146–54, 2012.

Friston K J and Stephan K E. Free-energy and the brain. Synthese, 159(3):417–458, 2007.

Garner W R. An informational analysis of absolute judgments of loudness. Journal of Experimental Psychology, 46(5):373–380, 1953.

Gescheider G a. Psychophysical scaling. Annual Review of Psychology, 39:169–200, 1988.

Gescheider G A and Collins A A. The measurement of loudness in individual children and adults by absolute magnitude estimation and cross-modality matching. Journal of Acoustical Society of America, 85:2012–2020, 1989.

Gibbon J. Scalar expectancy theory and Weber's law in animal timing. Psychological Review, 84(3):279–325, 1977.

Glasauer S, Schneider E, Grasso R and Ivanenko Y P. Space-time relativity in self-motion reproduction. Journal of Neurophysiology, 97:451–461, 2007.

Glasauer S, Stein A, Guenther A L, Flanagin V L, Jahn K and Brandt T. The effect of dual tasks in locomotor path integration. Annals of the New York Academy of Sciences, 1164:201–205, 2009.

Green D M and Swets J A. Signal detection theory and psychophysics. Wiley, New York, 1966.

Grenander U. Stochastic processes and statistical inference. Arkiv for Matematik, 1:195–277, 1950.

Griffiths T L, Chater N, Norris D and Pouget A. How the Bayesians got their beliefs (and what those beliefs actually are): comment on Bowers and Davis (2012). Journal of Experimental Psychology General, 138(3):415–422, 2012.

Hollingworth H L. The central tendency of judgment. The Journal of Philosophy, Psychology and Scientific Methods, 7(17):461 – 469, 1910.

Howard I S, Ingram J N, Körding K P and Wolpert D M. Statistics of natural movements are reflected in motor errors. Journal of Neurophysiology, 102(3):1902–1910, 2009.

Jazayeri M and Shadlen M N. Temporal context calibrates interval timing. Nature Neuroscience, 13(8):1020–1026, 2010.

Jones F N and Marcus M J. The subject effect in judgments of subjective magnitude. Journal of Experimental Psychology, 61:40–44, 1961.

Jones F N and Woskow M. On the relationship between estimates of magnitude of loudness and pitch. American Journal of Psychology, 75:669–671, 1962.

Jones M and Love B C. Bayesian fundamentalism or enlightenment? On the explanatory status and theoretical contributions of Bayesian models of cognition. The Behavioral and Brain Science, 34(4):169–188, 2011.

Jürgens R and Becker W. Perception of angular displacement without landmarks: evidence for Bayesian fusion of vestibular, optokinetic, podokinesthetic, and cognitive information. Experimental Brain Research, 174(3):528–543, 2006.

Kalman R E. A new approach to linear filtering and prediction problems. Transaction of the ASME—Journal of Basic Engineering, 82(Series D):35–45, 1960.

Kappauf W E. Regression effect in judgments to determine equal response contours. Perception & Psychophysics, 17(4):405–410, 1975.

Kersten D and Mamassian P. Ideal observer theory. Encyclopedia of Neuroscience, 5:89–95, 2009.

Kersten D, Mamassian P and Yuille A. Object perception as Bayesian inference. Annual Review of Psychology, 55:271–304, 2004.

King B M. Odor intensity measured by an audio method. Journal of Food Science, 51(5):1340–1344, 1986.

Klatzky R L, Loomis J M, Golledge R G, Cicinelli J G, Doherty S and Pellegrino J W. Aquisition of route and survey knowledge in the absence of vision. Journal of Motor Behavior, 22:19–43, 1990.

Knill D C and Pouget A. The Bayesian brain: the role of uncertainty in neural coding and computation. Trends in Neurosciences, 27(12):712–719, 2004.

Knill D C and Richards W. Perception as Bayesian inference. Cambridge University Press, Cambridge, 1996.

Körding K P, Ku S p and Wolpert D M. Bayesian integration in force estimation. Journal of Neurophysiology, 92(5):3161–3165, 2004.

Körding K P and Wolpert D M. The loss function of sensorimotor learning. Proceedings of the National Academy of Sciences of the United States of America, 101(26):9839–9842, 2004.

Körding K P and Wolpert D M. Bayesian decision theory in sensorimotor control. Trends in Cognitive Sciences, 10(7):319–326, 2006.

Laming D. The measurement of sensation. Oxford University Press, Oxford, 1997.

Laming D. Prior expectations in cross-modality matching. Mathematical Social Sciences, 38(3):343–359, 1999.

Landy M S, Maloney L T, Johnston E B and Young M. Measurement and modeling of depth cue combination: in defense of weak fusion. Vision Research, 35(3):389–412, 1995.

Lappe M, Stiels M, Frenz H and Loomis J M. Keeping track of the distance from home by leaky integration along veering paths. Experimental Brain Research, 212:81–89, 2011.

Ma W J and Pouget A. Linking neurons to behavior inmultisensory perception: A computational review. Brain Research, 1242:4–12, 2008.

MacKay D M. Psychophysics of perceived intensity: A theoretical basis for Fechner's and Stevens' laws. Science, 139(3560):1213–1216, 1963.

Maloney L T. Statistical theory and biological vision. In Heyer D and Mausfeld R, editors, Perception and the physical world, pages 145–189. John Wiley & Sons, Ltd, Chichester, UK, 2002.

Mamassian P, Landy M and Maloney L T. Bayesian modelling of visual perception. In Rao R P N, Olshausen B A and Lewicki M S, editors, Probabilistic models of the brain: Perception and neural function, pages 13–36. MIT Press, Cambridge, 2002.

Masin S C, Zudini V and Antonelli M. Early alternative derivations of Fechner's law. Journal of the History of the Behavioral Sciences, 45(1):56–65, 2009.

Michel M M and Jacobs R A. Learning optimal integration of arbitrary features in a perceptual discrimination task. Journal of Vision, 8(2):1–16, 2008.

Michell J. Psychophysics, intensive magnitudes, and the psychometricians' fallacy. Studies in History and Philosophy of Biological and Biomedical Sciences, 17:414–432, 2006.

Miyazaki M. Testing Bayesian models of human coincidence timing. Journal of Neurophysiology, 94:395–399, 2005.

Neyman J and Pearson E S. On the problem of the most efficient tests of statistical hypotheses. Philosophical Transactions of the Royal Society of London, 231:289–337, 1933.

Nieder A and Miller E K. Coding of cognitive magnitude: compressed scaling of numerical information in the primate prefrontal cortex. Neuron, 37:149–157, 2003.

Nover H, Anderson C H and DeAngelis G C. A logarithmic, scale-invariant representation of speed in macaque middle temporal area accounts for speed discrimination performance. The Journal of Neuroscience, 25(43):10049–10060, 2005.

Peterson W W, Birdsall T G and Fox W C. The theory of signal detectability. Proceedings of the IRE Professional Group on Information Theory, 4:171–212, 1954.

Petzschner F H and Glasauer S. Iterative Bayesian estimation as an explanation for range and regression effects: A study on human path integration. The Journal of Neuroscience, 31(47):17220–17229, 2011.

Petzschner F H, Maier P and Glasauer S. Combining symbolic cues with sensory input and prior experience in an iterative Bayesian framework. Frontiers in Integrative Neuroscience, 6(58):1–18, 2012.

Poggio T, Torre V and Koch C. Computational vision and regularization theory. Nature, 317(6035):314–319, 1985.

Poulton E C. The new psychophysics: Six models for magnitude estimation. Psychological Bulletin, 69(1):1–19, 1968.

Poulton E C. This week's citation classic: The new psychophysics: six models for magnitude estimation. Citation Classics, 20:83, 1981.

Poulton E C. Bias in quantifying judgments. Laurence Erlbaum Associates, Hilldale, 1989.

Rice S O. Mathematical analysis of random noise. Bell Systems Technical Journal, 23:282–332, 1944.

Robinson G H. Biasing power law exponents by magnitude estimation instructions. Perception & Psychophysics, 19(1):80–84, 1976.

Rouder J N and Morey R D. The nature of psychological thresholds. Psychological Review, 116(3):655–60, 2009.

Rule S J. Subject differences in exponents of psychophysical power functions. Perceptual & Motor Skills, 23:1125–1126, 1966.

Schwartz M. Haptic perception of the distance walked when blindfolded. Journal of Experimental Psychology: Human Perceptual Performance, 25:852–865, 1999.

Shannon C E. A mathematical theory of communication. Bell System Technical Journal, 27:379–423, 1948.

Sheldon P E. Equal-onset contours of vibrotactile stimuli. Perception & Psychophysics, 13(3):403–407, 1973.

Shepard R N. On the status of the 'direct' psychophysical measurement. Journal of Mathematical Psychology, 24:21–57, 1981.

Stevens S S. On the theory of scales of measurement. Science, 103:677–680, 1946.

Stevens S S. Mathematics, measurement and psychophysics. In Handbook of experimental psychology, pages 1–49. John Wiley & Sons, Ltd, 1951.

Stevens S S. The direct estimation of sensory magnitudes - loudness. The American Journal of Psychology, 69(1):1–25, 1956.

Stevens S S. On the psychophysical law. Psychological Review, 64(3):153–81, 1957.

Stevens S S. On the new psychophysics. Scandinavian Journal of Psychology, 1(1):27–35, 1960a.

Stevens S S. The psychophysics of sensory function. American Scientist, 48(2):226–253, 1960b.

Stevens S S. To honor Fechner and repeal his law: A power function, not a log function, describes the operating characteristic of a sensory system. Science, 133(3446):80–86, 1961.

Stevens S S. A metric for the social consensus. Science, 151:530–541, 1966a.

Stevens S S. Concerning the measurement of brightness. Journal of the Optical Society of America, 56(8):1135–1136, 1966b.

Stevens S S. Neural events and the psychophysical law. Science, 170:1043 – 50, 1970.

Stevens S S and Greenbaum H B. Regression effect in psychophysical judgment. Perception & Psychophysics, 1:439 – 446, 1966.

Stevens S S and Guirao M. Subjective scaling of length and area and the matching of length to loudness and brightness. Journal of Experimental Psychology, 66:177–186, 1963.

Stevens S S and Guirao M. Scaling of apparent viscosity. Science, 145(3634):1157 – 1158, 1964.

Stevens S S and Harris J R. The scaling of subjective roughness and smoothness. Journal of Experimental Psychology, 64:489–494, 1962.

Stevens S S and Poulton E C. The estimation of loudness by unpracticed observers. Journal of Experimental Psychology, 51:71–78, 1956.

Stocker A A and Simoncelli E P. Noise characteristics and prior expectations in human visual speed perception. Nature Neuroscience, 9(4):578–85, 2006.

Tassinari H, Hudson T E and Landy M S. Combining priors and noisy visual cues in a rapid pointing task. The Journal of Neuroscience, 26(40):10154–10163, 2006.

Teghtsoonian M and Teghtsoonian R. How repeatable are Stevens's power law exponents for individual subjects? Perception & Psychophysics, 10(3):147–149, 1971.

Teghtsoonian R. Range effects in psychophysical scaling and a revision of Stevens' law. The American Journal of Psychology, 86(1):3–27, 1973.

Teghtsoonian R and Teghtsoonian M. Range and regression effects in magnitude scaling. Perception & Psychophysics, 24(4):305–314, 1978.

Thomas G J. Equal-volume judgments of tones. The American Journal of Psychology, 62(2):182–201, 1949.

Verstynen T and Sabes P N. How each movement changes the next: an experimental and theoretical study of fast adaptive priors in reaching. The Journal of Neuroscience, 31(27):10050–10049, 2011.

Vilares I and Kording K. Bayesian models: The structure of the world, uncertainty, behavior, and the brain. Annals of the New York Academy of Sciences, 1224(1):22–39, 2011.

von Helmholtz H. Handbuch der physiologischen Optik. Leopold Voss, Leipzig, 1867.

von Helmholtz H. Die Tathsachen der Wahrnehmung. In Vorträge und Reden, pages 213 – 247. Vieweg, Braunschweig, 1878.

Wagenaar W A. Stevens vs Fechner: A plea for dismissal of the case. Acta Psychologica, 39:225–235, 1975.

Wald A. Contributions to the theory of statistical estimation and testing hypotheses. The Annals of Mathematical Statistics, 10:299 – 326, 1939.

Walsh V. A theory of magnitude: common cortical metrics of time, space and quantity. Trends in Cognitive Sciences, 7(11):483–488, 2003.

Warren R M and Warren R P. Helmholtz on perception: Its physiology and development. John Wiley & Sons, Ltd, New York, 1986.

Weber E H. Der Tastsinn und das Gemeingefühl. In Wagner R, editor, In Handwörterbuch der Physiologie, pages 481–588. Vieweg, Braunschweig, 1850.

Weiss Y, Simoncelli E P and Adelson E H. Motion illusions as optimal percepts. Nature Neuroscience, 5(6):598–604, 2002.

Welch G and Bishop G. An introduction to the Kalman filter. Technical report, University of North Carolina, Chapel Hill, 2002.

Westheimer G. Was Helmholtz a Bayesian? Perception, 37(5):642–650, 2008.

Zwislocki J J. Group and individual relations between sensation. Perception & Psychophysics, 33(5):460–468, 1983.

ACKNOWLEDGEMENTS

Leider lässt sich wahrhafte Dankbarkeit nicht in Worten ausdrücken.

— Johann Wolfgang von Goethe

This dissertation would not have been possible without the guidance and the help of several individuals who in one way or another contributed to this work.

First and foremost, I would like to express my gratitude to my first supervisor Stefan Glasauer, who has taught my science in different ways, evoked my interest for the *'old'* research and who always encouraged me to become an independent thinker.

I am deeply grateful to my two remaining thesis committee members. I want to sincerely thank Virginia Flanagin, for her unfailing help, advice and support and Andreas Herz, who provided me with valuable guidance and unique opportunities.

I want to thank the GSN, the BCCN Munich and the IFB-LMU for providing the permission and funding to undertake this research.

Special thanks go to many friends and colleagues in the Forschungshaus and the GSN for the scientific and non-scientific discussions, advice and continuous support. Especially, I want thank my office mates, Christopher, Nadine and Leni, you guys made work fun, Melanie for a deep friendship starting on our very first day, Paul for the great collaboration. I want to thank all proof-readers for their extremely valuable comments, Alex for bringing this thesis design to my attention and finally Jackson, TBT and She Hulk for a memorable time at work, after work and on the rocks!

Many scientists in and outside Munich have influenced my work. Foremost, I would like to thank Mike Shadlen, for a wonderful and inspiring research internship in Seattle, it has been, and still is, a pleasure to work with him and Klaas Enno Stephan for guiding me towards the next steps.

My deepest gratitude goes to my family and my partner Steph, who have given me their unequivocal support throughout, as always. Words cannot express just how decisive our bond has been. Danke für alles!

CURRICULUM VITAE

Frederike Hermi Petzschner
Date of birth 19.07.1986 in Bayreuth, Germany

EDUCATION

Ludwig-Maximilians-University PhD (candidate) Neuroscience	Munich expec. defense 01/2013
Julius-Maximilians-University Master of Science with Honors in Physics FOKUS Physics, Average Mark: Excellent (1.0)	Würzburg 09/2009
Julius-Maximilians-University Bachelor of Science in Physics FOKUS Physics, Average Mark: Excellent (1.4)	Würzburg 07/2008

RESEARCH EXPERIENCE

German Dizziness Center Graduate Researcher	Munich since 01/2011
Bernstein Center for Computational Neuroscience Graduate Researcher	Munich since 01/2009
Graduate School of Systemic Neurosciences (GSN), LMU PhD student Advisory Committee: Stefan Glasauer, Virginia Flanagin, Andreas Herz	Munich since 09/2009
University of Washington Research Internship; Advisor: Michael Shadlen	Seattle 01-04/2012
Research Center for Magnetic Resonance Bavaria e.V. Research Internship; Advisor: Felix Breuer	Würzburg 09-11/2007
Max-Planck-Institute Research Internship; Advisor: Jens Frahm, Dirk Voit	Göttingen 08-10/2007

HONORS AND AWARDS

Final Exam Award, Deutsche Physikalische Gesellschaft	2005
Publications Award, Graduate School of Systemic Neurosciences	2011
Poster Award, Conference Neural Coding, Frankfurt	2012
Poster Award, Graduate School of Systemic Neurosciences	2012

PUBLICATIONS:

Petzschner F.H., Maier P., Glasauer S. (2012). Combining symbolic cues with sensory input and prior experience in a Bayesian framework. *Frontiers in Integrative Neuroscience*. 6(58):1-12

Petzschner F.H., Krüger M. (2012). How to Reach: Movement Planning in the Posterior Parietal Cortex. *Journal of Neuroscience*. 32(14):4703-04

Petzschner F.H., Glasauer S. (2011). Iterative Bayesian estimation as an explanation for regression and range effects - a study on human path integration. *Journal of Neuroscience*. 31(47):17220-9

Petzschner F.H., Ponce I.P., Blaimer M., Jakob P.M., Breuer F.A. (2011). Fast MR parameter mapping using k-t principal component analysis. *Magnetic Resonance in Medicine*. 66(3):706-16

In Preparation:

Maier P., Petzschner F.H., Glasauer S. Human Navigation is Tuned to Predictably Changing Environments

Glasauer S., Maier P., Petzschner F.H. Learning from sensory experience requires behavioral relevance

Selected Conference Presentations:

Petzschner F.H., Maier P., Glasauer S. (2012) The influence of prior experience and symbolic cueing on human path integration. *Bernstein Conference*, Munich

Petzschner F.H., Glasauer S. (2012) Iterative Bayesian Estimation explains Regression and Range Effects in Human Path Integration *Conference Neural Coding, Decision-Making and Integration in Time*, Frankfurt

Petzschner F.H., Glasauer S. (2011) Beyond sensory input? Iterative Bayesian Estimation explains Regression and Range Effects in Human Path Integration. *Society for Neuroscience Conference*, Washington.

Petzschner F.H., Glasauer S. (2011) Optimal Navigation. *21. Oculomotor Meeting*, Munich

EIDESSTATTLICHE ERKLÄRUNG

Ich versichere hiermit an Eides statt, dass die vorgelegte kumulative Dissertation von mir, wo nicht anderweitig gekennzeichnet, selbständig und ohne unerlaubte Hilfe angefertigt ist. Die Beiträge zu den Manuskripten waren wie folgt:

Die Erstellung der Veröffentlichung
FH Petzschner and S Glasauer: Iterative Bayesian Estimation as an Explanation for Range and Regression Effects: A Study on Human Path Integration. The Journal of Neuroscience; 31(47):17220-17229, 2011
wurde unter der Leitung von Stefan Glasauer durchgeführt;
F.H.P. und S.G. entwickelten das Forschungsprojekt; F.H.P. und S.G. führten die Forschung durch; F.H.P. führte die Experimente durch; S.G. stellte nicht publizierte analytische Methoden zur Verfügung; F.H.P. analysierte die Daten; F.H.P. und S.G. schrieben das Paper.

Die Erstellung der Veröffentlichung
FH Petzschner, P Maier and S Glasauer: Combining symbolic cues with sensory input and prior experience in an iterative Bayesian framework. Frontiers in Integrative Neuroscience; 6:58, 1-18, 2012
wurde unter der Leitung von Stefan Glasauer durchgeführt;
F.H.P. und S.G. entwickelten das Forschungsprojekt; F.H.P., P.M. und S.G. führten die Forschung durch; F.H.P. führte die Experimente durch; F.H.P. und P.M. analysierten die Daten; F.H.P., P.M. und S.G. schrieben das Paper.

München, 21.11.2012

Frederike Hermi Petzschner

Hiermit bestätigen die Mitautoren die von Frau Petzschner angegebenen Beiträge zu den einzelnen Publikationen.

München, 21.11.2012

Paul Maier   Stefan Glasauer