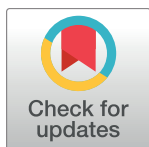# Causal Inference in Multisensory Heading Estimation

Ksander N. de Winkel*, Mikhail Katliar, Heinrich H. Bülthoff

Department of Human Perception, Cognition, and Action, Max Planck Institute for Biological Cybernetics, Tübingen, Baden-Württemberg, Germany

* ksander.dewinkel@tuebingen.mpg.de

Abstract

A large body of research shows that the Central Nervous System (CNS) integrates multisensory information. However, this strategy should only apply to multisensory signals that have a common cause; independent signals should be segregated. Causal Inference (CI) models account for this notion. Surprisingly, previous findings suggested that visual and inertial cues on heading of self-motion are integrated regardless of discrepancy. We hypothesized that CI does occur, but that characteristics of the motion profiles affect multisensory processing. Participants estimated heading of visual-inertial motion stimuli with several different motion profiles and a range of intersensory discrepancies. The results support the hypothesis that judgments of signal causality are included in the heading estimation process. Moreover, the data suggest a decreasing tolerance for discrepancies and an increasing reliance on visual cues for longer duration motions.

Introduction

A large body of research shows that the Central Nervous System (CNS) processes multiple sensory signals on any particular environmental property in a fashion that is consistent with Bayesian inference, a strategy commonly referred to as Multisensory Integration (MI) [1–6]. According to this strategy, the CNS computes as the final estimate of a property the single value that is most likely given the sensory signals that represent it and any prior beliefs an individual may hold about the property. However, as our senses continuously process stimuli emanating from a multitude of events in the surrounding environment, the case where multiple signals actually represent a common property can be considered relatively rare; the majority of the signals should not be integrated. Consequently, it is plausible that the CNS somehow incorporates assessments of signal causality in the estimation of environmental properties [7, 8].

The idea that the CNS assesses signal causality is explicitly addressed in Causal Inference (CI) models of multisensory perception [9–11]. In CI models, a final estimate is constructed from two intermediate estimates. The intermediate estimates arise from different interpretations of the causality of the sensory signals. The first interpretation is that multiple signals share a common cause, where the MI strategy dominates the final estimate; the alternative is an interpretation of independent causes, where an individual channel determines the final estimate

(Sensory Segregation; SS). A number of variants of the CI model have been proposed [11], that each construct a final estimate in a slightly different way: either by weighting and summing the intermediate estimates, or by applying a decision rule, where the weight or the probability of a specific choice is proportional to the probability of the respective causal structures.

In the majority of experimental studies on multisensory heading estimation, participants are presented with visual and inertial stimuli with minor discrepancies, assumed to be below the threshold of detection, or no discrepancies at all. Under these conditions, the general finding is that multisensory heading estimation is indeed consistent with the MI strategy (e.g., [12]). However, some findings violate the predictions of MI. For example, in [13], participants relied solely on one of the signals to estimate their heading. Other studies report multisensory integration for some, but not all, participants [14–16].

In previous work [17], we postulated that the dichotomy in experimental findings results from CI: whereas inertial information on self-motion is almost exclusively caused by self-motion (for exceptions, see e.g., galvanic stimulation, [18]), visually perceived motion can be caused by self-motion, object motion, or a combination thereof. Consequently, reports that are inconsistent with MI could imply that particularities of the stimuli led to an interpretation of independent causes. However, when we assessed the tenability of a CI model of multisensory heading estimation experimentally, the obtained data reflected MI regardless of the size of introduced discrepancies. This finding suggested that humans are essentially oblivious to discrepancies between visual and inertial heading. Although this conclusion is consistent with the finding that MI in heading estimation is robust for differences between the motion profiles of the visual and the inertial stimulus [19], it contrasts with earlier findings inconsistent with or partially supporting MI [13–16].

The most striking methodological difference between the reports consistent with MI and the aforementioned study by [13] concerns the characteristics of the motion profiles. Although the maximum acceleration was similar, at approximately $0.5 m/s^2$, the motions presented in the latter study had translations and durations that were about an order of magnitude larger than the motions used in the former studies. These differences amounted to, respectively, a translation of seven meters over the course of almost ten seconds versus translations of a few decimeters in one to two seconds. This inspired the present hypothesis that the CNS does perform CI in the heading estimation process, but that the multisensory heading estimation process is affected by the characteristics of the motion profiles.

Results

To test the hypothesis outline above, we performed two experiments (I and II). In both experiments, we used the Max Planck CyberMotion Simulator [20] to present participants with a large number (600-648) of unisensory visual-only and inertial-only, and multisensory visual-inertial motion stimuli, with headings ranging the full circle, and discrepancies up to ±90˚ (in the multisensory conditions). For each experimental trial, participants indicated the heading of self-motion using a pointer device. All motions followed a raised cosine bell velocity profile, and had durations of two, four, or six seconds. In experiment I (8 participants), maximum *velocity* was kept equal (0.30m/s) between motions of different duration; in experiment II (9 participants), maximum *acceleration* was kept equal ($0.50 m/s^2$). Differences between experiments were introduced to be able to dissociate potential effects of duration, velocity, and acceleration. Because motion profiles can be identified by their duration within experiments, we occasionally refer to a motion profile by its duration for convenience. For a complete overview of the task and stimuli, see section: Task and Stimuli. A video of simulator motion is available as supplementary material S1 Video.

From the responses obtained in the unisensory conditions, we determined whether patterns of bias and dispersion differed between motion profiles, and whether observations corresponded to the literature [17, 21, 22]; from the responses obtained in the multisensory conditions, we firstly determined whether either of a number of variants of a CI model could account for the data, and secondly whether there were differences in model parameters for different motion profiles.

To decide upon the best models of unisensory and multisensory heading estimation, we compared model Bayesian Information Criterion (BIC) scores [23]. This is a measure of model quality based on the model likelihood, number of observations, and the number of parameters. The model with the lowest BIC score is considered the best in an absolute sense. Differences in BIC scores of competing models, $\triangle$BIC, between 0-2; 2-6; 6-10 are considered negligible, positive, and strong evidence, and >10 is considered decisive evidence.

Unisensory Conditions

Models that allow for periodic patterns in bias (constant error) and dispersion (variable error) in heading estimates were fitted to the responses obtained in the unisensory conditions. Both the bias and the dispersion were constructed as having a constant part and a part that varied periodically with heading, with a 90° period. We assessed if and how bias and dispersion of visual and inertial heading estimates were affected by characteristics of the motion profiles. Specifically, we compared the tenability of models in which the parameters that describe bias (β_0, β_1) and dispersion (γ_0, γ_1) were either kept constant between profiles or allowed to vary. The evidence in favor of the best fitting unisensory model for individual participants was always decisive, as compared to a null-model that did not include bias or variability of dispersions ($\triangle$BIC = 72.1 − 770.0: see supplementary material S1 Table). As an illustration, Fig 1 shows for an individual participant the data and fits for a model where all parameters were allowed to vary.

In experiment I, a model in which the constant (γ_0) and heading-dependent (γ_1) parts of the dispersion of inertial heading estimates were allowed to vary was preferred for three out of eight participants, and a model in which only the constant part (γ_0) varied was preferred for three others. The strength of the evidence favoring these models ranged between positive and decisive ($\triangle$BIC = 4.06 − 18.02, median = 10.32). For the two remaining participants, there was no evidence for variability of bias or dispersion. For eight out of nine participants in experiment II, patterns of bias or dispersion were unaffected by motion duration. For the remaining participant, the constant part (γ_0) of the visual dispersion varied for different motion durations ($\triangle$BIC = 4.22). Findings on bias and total dispersion (see: S1 Appendix) of both experiments are visualized in Fig 2.

The finding of variability of dispersions for the majority of participants in experiment I suggested a systematic effect of motion characteristics on inertial heading estimation, and corresponded to participants' subjective reports that task difficulty increased with motion duration. An assessment of total dispersion confirmed that the size of inertial dispersion increases for longer motion durations ($\triangle$BIC = 2.708).

Apart from the observed effects of motion characteristics on dispersion in experiment I, the parameters that determine dispersion, γ_0 and γ_1, were generally positive and within the same order of magnitude, indicating that heading dependencies of variability in dispersion are similar in all conditions, with larger dispersions for headings further away from the cardinal axes.

The constant part of the bias (β_0) amounted to 3.67° in experiment I, and 2.98° in experiment II. This parameter was included to account for factors such as the orientation of the participant relative to the simulator, and was kept equal in all conditions. For the visual
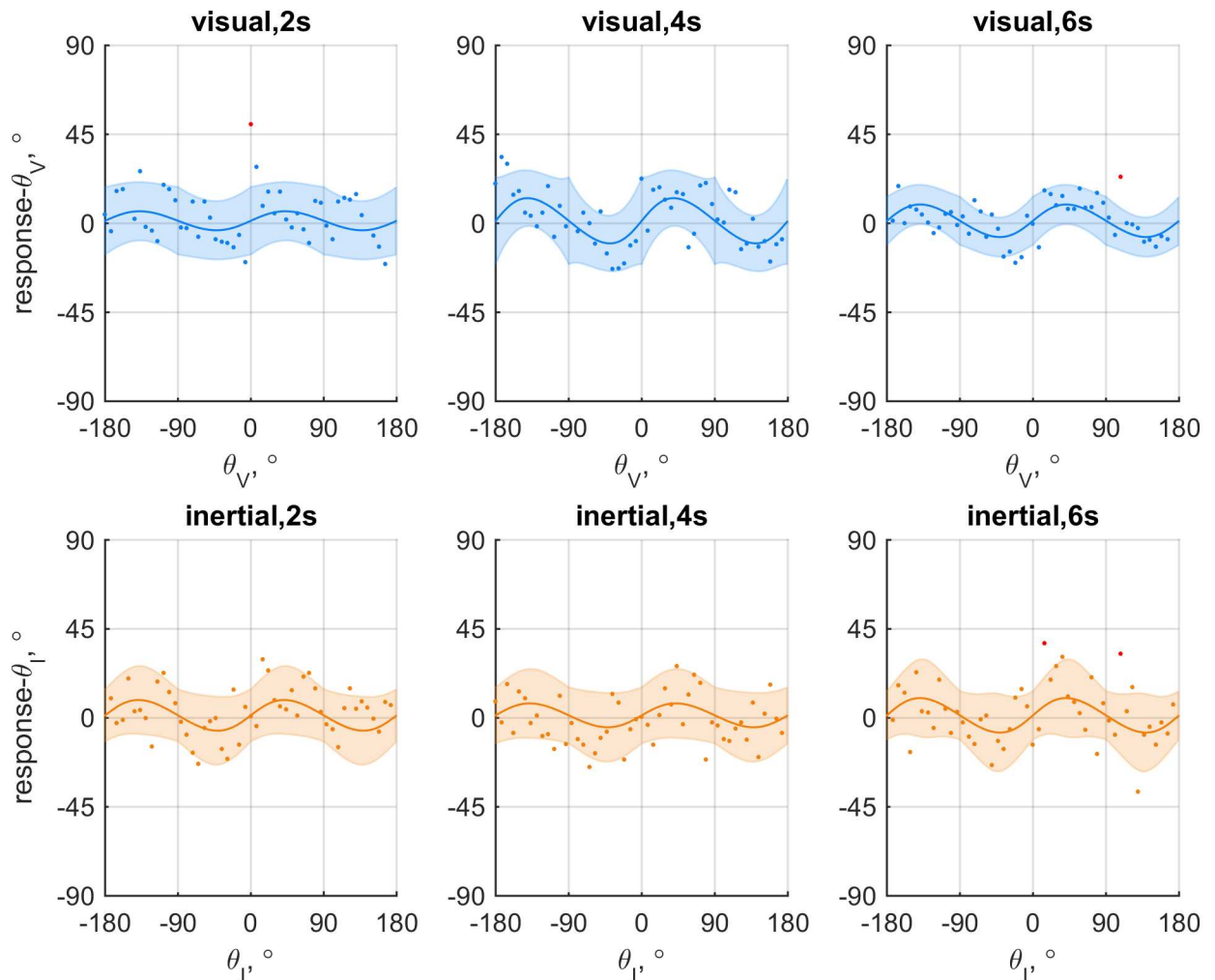
Fig 1. Data and fits for an exemplary participant (6) from experiment II split by experimental condition. The upper row shows the data for visual-only conditions; the bottom row shows the data for inertial-only conditions. Columns correspond to motion duration. Individual dots each represent the signed difference between the response and stimulus heading: red dots mark outliers (i.e., observation is more than three standard-deviations from $\mu(\theta)$). The colored lines show fit of the model where the parameters were allowed to vary between conditions. Shaded areas represent 95% confidence intervals.

doi:10.1371/journal.pone.0169676.g001

conditions of experiment I and II, the median estimates of parameter β_1 were 1.19 and 1.06; these averages were 1.02 and 0.88 in the inertial conditions. The corresponding peak sizes of the heading dependent part of the bias were 4.98° and 1.67° in the visual conditions, and 0.57° and 3.66° in the inertial conditions. There was considerable variability in the estimates of β_1. In the visual conditions, the value of ranged between $0.85 - 1.33$, corresponding to peak biases of 4.65° and 8.14°. In the inertial conditions, the estimates ranged between 0.67 an 1.76, corresponding to peak biases of 11.40° and 15.98°. Note that positive values of β_1 indicate bias away from the cardinal axes, and negative values indicate bias towards the cardinal axes.

The estimated coefficients for all participants are provided in supplementary material S2 Table.
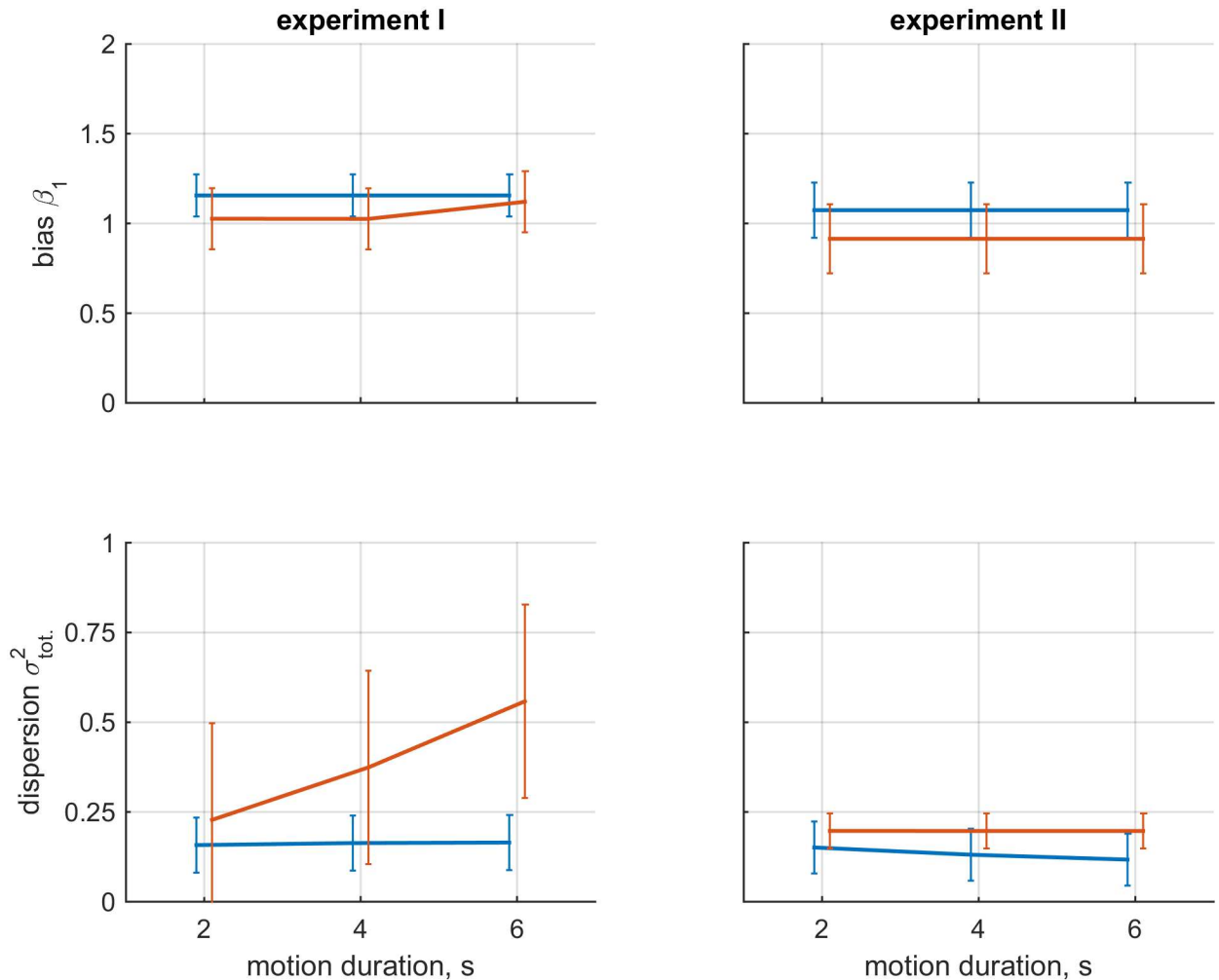
Fig 2. Visualization of findings on bias and dispersion aggregated over participants. The left column shows the data for experiment I; the right column shows the data for experiment II. The upper row shows the mean and standard deviation of parameter β_1 for the different motion conditions; the bottom row shows the total dispersion. Blue and orange lines represent the visual-only and inertial-only conditions, respectively. The errorbars represent standard deviation.

Multisensory Conditions

The analysis of multisensory data consisted of a comparison of the tenability of a Multisensory Integration (MI) model, a Sensory Segregation (SS) model, and three variants of Causal Inference (CI) models: Model Averaging (MA), Model Selection (MS), and Probability Matching (PM) [11]. In the following we provide a concise description of the models. A more detailed description of the models can be found in section: Analyses of Multisensory Data and in [17].

The MI model is a circular analogue of the model proposed by [4]. According to this model, the CNS determines as a heading estimate the single heading angle that is most likely to have caused both the visual and inertial signals.

The SS model states that heading estimates are based upon either the visual or the inertial signal; the probability that heading estimates are based upon the visual signal is expressed by parameter ξ. ξ is treated as a free parameter, and takes on values between zero and one. Values

of zero and one respectively correspond to complete dominance of the inertial and visual system.

In the CI models, heading estimates resulting from MI and SS are treated as intermediate estimates, and combined according to the probability that the visual and inertial signal share a common cause ($P(C \mid x_V, x_I)$). This probability is proportional to the product of the likelihood of observing the specific pair of visual and inertial signals given that they have a common cause ($P(x_V, x_I|C)$), and a prior $P(C)$ that reflects a tolerance for discrepancies independent of the individual sensory signals or any discrepancy between them. $P(C)$ is treated as a free parameter that takes on a value between zero and one. When $P(C)$ equals zero, there is no tolerance for discrepancies, and multisensory signals will always be segregated (i.e., the CI model behaves like the SS model). In contrast, when $P(C)$ equals one, multisensory signals will always be attributed to a single cause (i.e., the CI model behaves like the MI model).

The variants of the CI model apply $P(C|x_V, x_I)$ in different ways. In the MA variant, $P(C|x_V, x_I)$ and its reciprocal serve as weights for the intermediate estimates according to MI and SS, respectively; in the MS variant, the final heading estimate is based upon the MI estimate when $P(C|x_V, x_I)$ is greater than 0.5, and based upon the SS estimate otherwise. Finally, in the PM variant, the final heading estimate is chosen at random from MI or SS, where $P(C|x_V, x_I)$ determines the probability of choosing MI.

On a more intuitive level, according to the MA model variant, the final heading estimate is a weighted average of the intermediate estimates according to the MI and SS strategies, with weights that correspond to the probability that the visual and inertial signal have a common, or independent causes. According to both the MS and PM variants, the final heading estimate is equal to *either* the MI or the SS estimate. The difference between these two variants is in how the choice is made. In the MS variant, the estimate with the most probable causal structure is chosen as the final estimate, whereas in the PM model, the final estimate is chosen at random between the two alternatives, with a probability of deciding on either the MI or the SS estimate that is equal to the probability of the corresponding causal structures.

Because the models and their variants differ in the way responses are constructed, there are also differences between the density functions that represent the likelihood of specific responses according to each of these models. We can evaluate which model corresponds best to the data by comparing these likelihoods. As an illustration of the behavior of the main models, Fig 3 shows the densities according to the MI, SS, and MA model for hypothetical experimental trials with a small, intermediate, and large visual-inertial discrepancy.

We assessed whether the characteristics of multisensory processing depended on characteristics of the motion profiles by determining if the overall model quality improved between fitting either a single pair of ξ and $P(C)$ parameters for all motion profiles, or fitting three pairs of ξ and $P(C)$ parameters; one pair for each profile. In the latter case, the models' and parameters' association to specific conditions is indicated by the subscript 'T', which represents the duration of the motion profiles.

Note that allowing the prior $P_T(C)$ to vary between motion duration conditions implies that it cannot be interpreted as an a-priori belief about signal causality at stimulus onset: participants were unaware of the motion characteristics at stimulus onset because stimuli were presented in random order, and beliefs could therefore not be adapted to the conditions at that exact moment. Instead, we interpret $P_T(C)$ as reflections of the tolerance for discrepancies after $T = 2, 4$ and 6s of motion, which is independent of the size of the discrepancy between sensory signals.

Due to their three dimensional nature, the raw data do not lend themselves well for visualization. Plots of the response bias (relative to visual heading angle) as a function of the size of
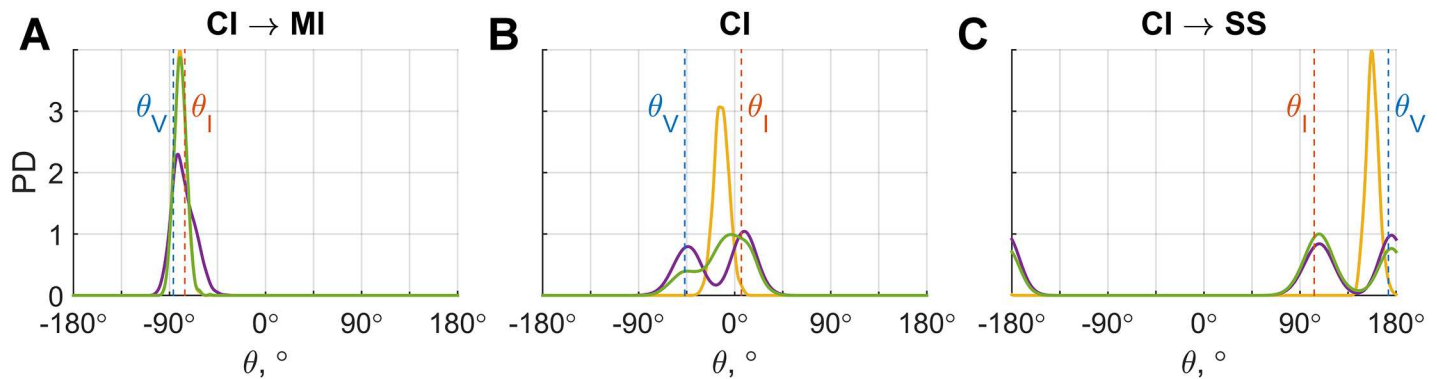
Fig 3. Visualization of probability densities of the response for hypothetical experimental trials, according to the MI(yellow), SS(purple), and CI-MA(green) models. The dashed vertical line in blue marked θ_V represents visual heading angle; the orange line marked θ_I represents inertial heading angle. The left panel illustrates that the CI-MA model behaves very similar to the MI model for small discrepancies; the middle panel shows that the CI-MA model averages between the MI and SS model for an 'intermediate' discrepancy; and the right panel shows that the CI-MA model behaves very similar to the SS model for large discrepancies. The likelihood of a response according to either model corresponds to the height of the associated density function at the response angle.

doi:10.1371/journal.pone.0169676.g003

the discrepancy between the visual and inertial heading angle are available as supplementary material S1 Fig.

Model Comparisons. Model comparisons for experiment I favored the CI-MA model for six out of eight participants. Compared to the best fitting alternative model (i.e., MI or SS), the evidence favoring a CI model was decisive in five out of these six cases ($\triangle$BIC = 12.8 – 125.4), and strong in the remaining case (pp. 8; $\triangle$BIC = 9.9). Data for the two remaining participants were best explained by the *SS* model.

Model comparisons in experiment II favored a CI model for eight out of nine participants. Five of these eight cases were in favor of the MA model; three in favor of the PM model. The evidence favoring CI models was decisive in six out of eight cases ($\triangle$BIC = 10.7 – 116.8), strong in one case ($\triangle$BIC = 8.2), and positive in the remaining case ($\triangle$BIC = 2.3). Data for the remaining participant were best explained by the *SS* model.

Combined over all participants, the MA model provided the best explanation of the data. The evidence was decisive, compared to the MI and SS model ($\triangle$BIC = 21869.8, and $\triangle$BIC = 440.6, respectively). All BIC scores can be found in supplementary material S3 Table; all obtained parameter estimates are presented in supplementary material S4 Table.

Multisensory effects of motion characteristics. For three of the participants from experiment I for whom the CI model provided the best fit, overall model fit was further improved by fitting the model separately for the different motion duration conditions. For these participants, the median estimates for $P_T(C)$ were 0.96, 0.61, and 0.11, for $T = 2$, 4, and 6s, respectively; the median values of parameter ξ_T were 0.32, 0.60, and 0.68.

In experiment II, the evidence favored parameter variability in two cases. For these cases, the median estimates of $P_T(C)$ were 0.95, 0.93, and 0.39, for $T = 2$, 4, and 6s, respectively; the median values of parameter ξ_T were 0.67, 0.99, and 1.00.

Combined over experiments, the value of parameter $P_T(C)$ for the $T = 6$s condition was generally lower than in the other two motion duration conditions, suggesting a decreasing tolerance for discrepancies independent of unisensory effects (Fig 4, left panel). Estimates of parameter ξ_T consistently increased between the $T = 2$s and $T = 6$s conditions, suggesting increasing reliance on the visual heading angle (Fig 4, right panel). These effects also hold
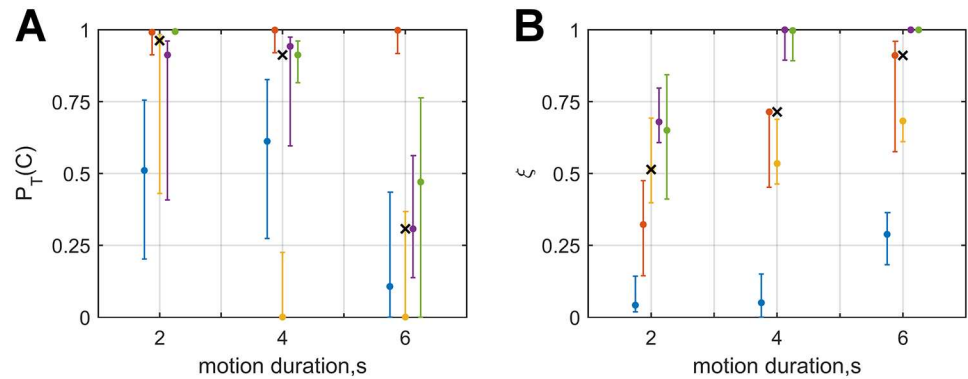
Fig 4. Estimates of the tolerance for discrepancies prior $P_T(C)$ (left panel) and reliance on the visual estimate ξ_T (right panel) for participants of both experiments for whom there was evidence that the parameters varied with motion duration, plotted against motion duration. Data points in blue, orange, yellow, purple, and green correspond to experiment-participant I3, I7, I8, II4, and II5, respectively. Errorbars represent the parameters' 95% confidence intervals, and the x-markers represent the median estimate for each $P_T(C)$ and ξ_T.

doi:10.1371/journal.pone.0169676.g004

when parameter estimates for the other participants are included. See supplementary material S4 Table for an overview of all the parameter estimates for multisensory models.

Fig 5 provides a visualization of how the probability that any pair of visual and inertial heading signals will be attributed to a common cause, $P(C|x_V, x_I)$, varies as a function of the different values of $P_T(C)$ obtained in the experiment, taking into account unisensory effects. The decreasing width of the band for which the probability of a common cause is relatively high is indicative of a declining tolerance for discrepancies as stimulus duration increases; the variability of the width of the band for different heading angles results from heading dependencies of unisensory biases and dispersions.

Discussion

In previous work [17], we evaluated the hypothesis that the Central Nervous System (CNS) performs Causal Inference (CI) in multisensory heading estimation. The obtained results, however, did not support the hypothesis, as the data could be explained equally well by the simpler Multisensory Integration (MI) model. Here we hypothesized that the lack of support for the CI model was due to the characteristics of the motion profiles. We assessed how multisensory heading stimuli were processed for motions with different durations, maximum velocities, and accelerations. In the following sections, we discuss the observations made on unisensory and multisensory heading estimation.

Unisensory Heading Estimation

In Experiment I, we observed that the dispersion of inertial heading estimates increased for longer duration motions, whereas this effect was absent in Experiment II. This result can be explained by considering that the otoliths of the vestibular system, which are the main contributor to perception of translational motion [24, 25], respond to acceleration: in Experiment I, maximum velocity was kept constant between motions of different durations. Consequently, the motions' maximum acceleration was lower for longer duration motions, and therefore the stimulus intensity was lower as well. In terms of Signal Detection Theory [26], low intensity signals are more difficult to distinguish from background noise than high intensity signals.
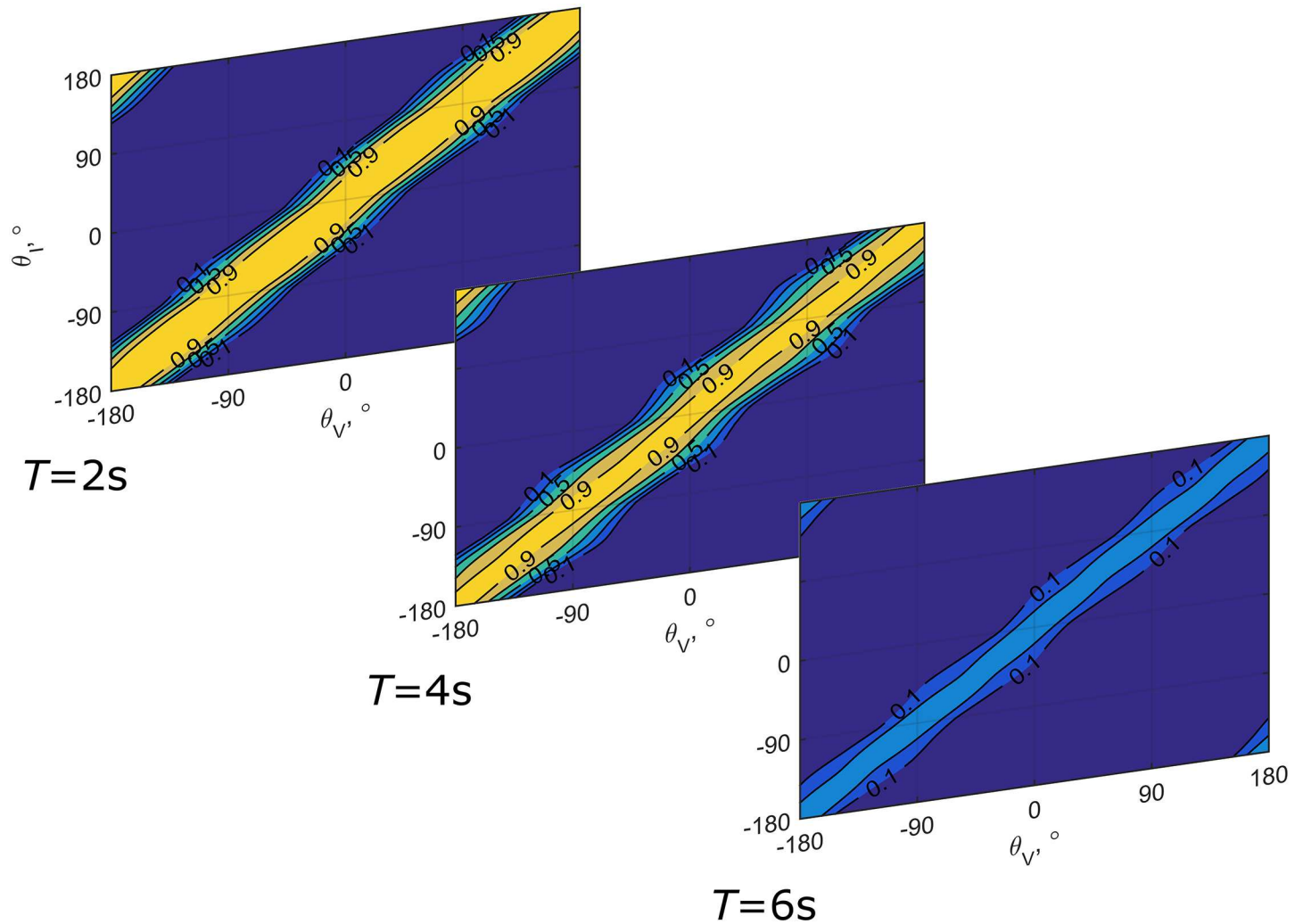
Fig 5. Visualization of the posterior probability that visual and inertial heading stimuli have a common cause $P_T(C|x_V, x_I)$, for each condition separately. The figure was created using the median values of $P_T(C)$ from Fig 4. Yellow represents a high probability; blue a low probability. The variability of the band's width results from bias in unisensory heading estimates and heading dependency of the dispersions.

This implies that misinterpretations of heading were more likely for the longer duration motions, resulting in larger dispersions. In contrast, in Experiment II, where maximum acceleration was kept constant, dispersion of heading estimates was found to be constant for all motion durations.

Apart from the variability of dispersion observed in experiment I, the findings on dispersion are consistent with the literature: inertial dispersion is larger than visual dispersion, and patterns in dispersion show that the smallest values occur around the cardinal axes, with values increasing as the heading angle deviates away from these axes [17, 21, 22, 27].

In a number of recent studies, both visual and inertial heading estimates were found to be biased away from the cardinal axes [21, 22, 27]. The present data show that visual heading estimates are on average also biased away from the fore-aft axis, in both experiments; inertial heading estimates were biased away from the cardinal axes in experiment I and biased towards the cardinal axes in experiment II. However, individual results show that there is considerable

variability in the direction and size of the bias between participants. Such variability was also observed in our previous study, using the same visual environment [17]. Biases in heading estimation are thought to result from non-uniformity of preferred directions of neurons at the otolith epithelium [28] and in area MSTd of the visual cortex [29]. The observed variability in biases could therefore suggest that the distribution of preferred directions in these neural populations varies between individuals.

Interestingly, the absence of a decline in the size of sensory bias for longer duration motions supports the notion that biases in heading estimation result from non-uniformity in the distribution of preferred directions rather than a-priori preferences for particular heading directions [22]; if bias in heading estimation was due to a-priori beliefs favoring particular directions, the relative impact of these beliefs on the resultant estimate should decline for longer duration motions, as accumulating sensory evidence devoid of bias would eventually outweigh any prior beliefs.

Multisensory Heading Estimation

The present analysis of multisensory data can be divided into two parts: in the first part, we assess whether the responses obtained in the multisensory conditions conform best to MI, SS, or CI; in the second part of the analysis, we assess whether characteristics of the motion profiles affect multisensory processing.

The first part of the analysis is nearly identical to [17]. However, whereas the conclusion in that study was that multisensory responses were best explained by the MI model (given an overall $\triangle$BIC = 47.76, compared to CI), the present results provide decisive evidence in favor of the CI-MA model ($\triangle$BIC = 21869.8 vs. MI, and $\triangle$BIC = 440.6 vs. SS). The CI model is an alternative to multisensory interpretations (e.g., [9, 30–34]) of the robust cue integration model that was originally proposed to model intrasensory cue combination in visual perception of depth (e.g., [35–38]); hereafter referred to as interaction-prior (IP) model [9, 33]. IP models also allow responses to range between sensory segregation and forced fusion, but do so by making use of multidimensional priors that represent the strength of coupling between sensory systems as a function of discrepancy between the cues (e.g., [31, 39, 40]). Although it is possible to cast the CI model as an IP model [32], there are conceptual differences: first, IP models do not assume that there are alternative causal structures, and do not make predictions on judgments of causality; and second, the IP used in these models must be estimated from the data or designed by the experimenters. In the present case, the IP should essentially take the shape of the probability maps presented in Fig 5 [32, 33]. In a comparison of the performance of a CI model and IP models proposed by Roach et al. [39] and Bresciani et al. [31] for audio-visual perceptions, Beierholm et al. [32] found that a CI model performs slightly better than the IP model alternatives. Moreover, Rohe and Noppeney [41] recently identified a cortical hierarchy that appears to perform CI: segregated internal representations of multisensory stimuli exist in unisensory cortical areas; forced fusion occurs in the posterior intraparietal sulcus; and a combination of the intermediate estimates according to the probability of the respective causal structures is constructed in the anterior intraparietal sulcus. Recent literature thus tends to support the notion of CI, and we interpret the present evidence as support for the notion that the CNS incorporates assessments of causality in multisensory heading estimation.

In the second part of the analysis, we split the data up by motion profile and fitted the models separately. This was done to determine whether multisensory processing is affected by characteristics of the motion profiles, apropos the hypothesis. A comparison of the quality of the fits provided evidence that splitting up the data improved the fit for 5 out of 14

participants. There was no distinct difference in the number of participants for whom the evidence favored separate models between experiments I and II (3/6, resp. 2/8).

The fact that CI was favored regardless of evident differences between motion profiles suggests that the experimental manipulations were more powerful than in the previous study. This finding might reflect the fact that there was a higher relative incidence of stimuli with large discrepancies in the present study than in [17], because intersensory discrepancies were drawn from a uniform distribution rather than from a normal distribution, with the same effective extremes of ±90˚.

For the participants for whom the evidence favored separate fits, the strength of the evidence was strong in two cases, and decisive in three. Consequently, even though explicit support for differences in multisensory processing was only found for a minority of participants, the strength of the evidence warrants further consideration of the results. The obtained estimates of the $P_T(C)$ parameters are shown in Fig 4. We interpret these parameters as reflections of the tolerance for discrepancies acting on the heading estimation process at time T. The effect of the value of this parameter on the posterior belief that a visual and an inertial heading have a common cause $P_T(C|x_V, x_I)$ is visualized in Fig 5. As the figure shows, the width and height of the band of combinations for which the strength of this belief is relatively high decreases for increasing motion durations due to the value of $P_T(C)$. The value of this parameter decreases with motion duration. A decreasing tolerance for discrepancies could reflect accumulation of evidence on signal causality over time. Such a mechanism in latent decision making would be consistent with findings on evidence accumulation in human and animal overt decision making (e.g., [42–45]). Unfortunately, the current data are not suitable to determine the characteristics of such an hypothetical updating mechanism, as we do not know the rate at which this evidence accumulates, nor do we have access to intermediate internal representations.

For six second motions, the belief that visual and inertial heading stimuli have a common cause can be considered weak even for stimulus combinations that yield very similar internal heading representations. The effect is apparent in Fig 5: for the six second duration conditions, the probability of an interpretation of common causes is low even when visual and inertial heading angles are identical. This suggests that visual stimuli may not have been considered as reflecting self-motion, which also resembles the findings of [13], who reported that participants segregated congruent visual and inertial heading cues completely for even longer duration motions (9.3s).

An alternative explanation of the findings on $P_T(C)$ and ξ_T that is consistent with this finding might be found in the dynamics that have been reported for the visual and inertial perception of self-motion. It has been suggested that visually induced perceptions of self-motion (i.e., 'vection', see e.g.: [46–48]) are consistent with predictions based on *low-pass* frequency filtering of the visual input([49]). In other words, the CNS segregates visually perceived motion based on frequency content: the low frequency component is attributed to self-motion, and the high frequency component to object motion. In contrast, perceptions of translation resulting from inertial stimulation are consistent with segregation based on *high-pass* frequency filtering of otolith afferents (e.g., [49–54]): the high frequency component of the gravito-inertial acceleration is attributed to translation, whereas the low-frequency component is attributed to body tilt relative to gravity. The cut-off frequencies of the inertial high-pass filter [51, 53–57], and the visual low-pass filter [49] are estimated to be approximately 0.5Hz –although the exact value of the estimates varies between studies and individuals, from approximately 0.1 to 1Hz. The cut-off frequency indicates a boundary above or below which the output of a filter begins to be attenuated (for a low-pass and high-pass filter, respectively). Consequently, the range of motion frequencies for which both the visual and inertial systems contribute perceptions of self-motion may only partially overlap; near the cut-off frequency. This limited overlap could

explain the observed decrease in the tendency to integrate multisensory information with motion duration, as well as the dichotomy in the literature: if we interpret the duration of the motions as the reciprocal of frequency, observations appear to support MI when motion frequencies are such that both systems are thought to yield perceptions of self-motion (i.e., close to the cut-off), whereas observations appear to oppose MI for lower frequencies, such as in [13] and the 6s condition of the present experiment. It could also explain the data of the three participants for whom the SS model was favored: the dynamics of the individual sensory systems may have been such that there was no overlap for the range of durations/frequencies tested.

Finally, it is interesting to note that in the study by De Winkel et al. [13], where participants appeared to segregate congruent visual-inertial cues, the visual stimulus was presented monoscopically, and consisted of motion through a black virtual environment populated with white circles; in the present study, the visual stimuli were presented stereoscopically, and consisted of limited lifetime random dots in an environment with a blue sky, and a grass like ground plane, with the horizon approximately at eye height. The present stimuli thus provided scaling information that allowed participants to relate the velocity specified in the visual flow magnitude to the inertial velocity, which was not possible in the earlier study. Consequently, the observed segregation of congruent stimuli in that study may also have been caused by perceptions of intersensory discrepancy in motion amplitude, which would imply that CI is also performed on the basis of discrepancies in other stimulus dimensions. This notion is consistent with findings from literature on vehicle simulation, which report on specific deterministic tolerances for discrepancies between the amplitude of visual and inertial motions (e.g., [58–60]).

We conclude that the current results provide supporting evidence for the hypothesis that the CNS includes judgments of signal causality in heading estimation. Moreover, specific findings in a subset of participants suggest that multisensory heading estimation for the purpose of self-motion perception might result from a complicated interplay between the dynamics of unisensory processing and multisensory strategies.

Materials and Methods

The experimental paradigm, equipment and analyses are adapted from those used in [17]. Descriptions of the changes implemented in the present analysis to account for effects of motion characteristics are provided below. For a more detailed account on the derivation of the models, please refer to the aforementioned study.

Ethics statement

The experiment was conducted in accordance with the Declaration of Helsinki. All participants gave their written informed consent prior to participation. The experimental protocol and consent forms were approved by the ethical commission of the medical faculty of the Eberhard-Karls University in Tübingen, Germany.

Participants

Eight participants (5F; aged 19-34, mean 27.3) were recruited to take part in experiment I; nine participants (6F; aged 19-34, mean 25.8) were recruited to take part in experiment II. Two participants participated in both experiments: participants 1 and 7 in Experiment I are the same participants as 3 and 2 in Experiment II. Except for these two, participants were naïve to the purposes of the study. The participants all reported minimal susceptibility to motion sickness and claustrophobia, and no history of any intestinal, neurological, or

Table 1. Motion duration _T_, translation _x_, maximum velocity _v__{_max_}, and maximum acceleration _a__{_max_}, for experiment I and II.

	experiment I			experiment II		
_T_	_x_, m	_v__{_max_}, m/s	_a__{_max_}, m/s^2	_x_, m	_v__{_max_}, m/s	_a__{_max_}, m/s^2
2	0.30	0.30	0.47	0.32	0.32	0.50
4	0.60	0.30	0.24	1.27	0.64	0.50
6	0.90	0.30	0.16	2.87	0.96	0.50

doi:10.1371/journal.pone.0169676.t001

vestibular illnesses. Due to safety regulations of the motion simulator, participation was only allowed for people between 18-65 years old, weighing up to 90kg, and measuring at most 1.95m.

Task and Stimuli

In both experiments, stimuli consisted of presentations of horizontal linear translational motions with various headings, in a visual-only, an inertial-only and a combined visual-inertial condition. The stimuli were presented using the Max Planck Institute CyberMotion Simulator facility [20]. Visual stimuli showed movements through a three dimensional environment with a ground plane and limited-lifetime particles, presented stereoscopically using a dual projection system in conjunction with stereo glasses (Infitec®GmbH, Ulm, Germany, model INFITEC®Premium Glasses). Inertial stimulation was achieved by translating participants seated in the simulator cabin along the simulator's linear track. To achieve different headings, the simulator cabin was oriented at different angles relative to the linear track. A video of simulator motion is available as supplementary material S1 Video.

For all motions, the motion profile followed a raised cosine bell in velocity, specified as

$$v(t) = \frac{v_{max}}{2}\left(1 - \cos\frac{2\pi t}{T}\right) \tag{1}$$

where _t_ is time and v_{max} is maximum velocity. _T_ represents the period of the motion profile, which took values of 2, 4 and 6s. The translations, velocities, and accelerations of the motions used in experiment I and II are presented in Table 1.

In the unisensory visual-only and inertial-only conditions of experiment I(cf. experiment II), 72(50) stimuli were presented for each of the three motion durations, with headings evenly spaced between ±180°, in 5°(7.2°) steps.

In the combined visual-inertial conditions, the stimuli were 72(100) visual motions with headings evenly spaced between ±180°, in 5°(3.6°) steps, paired with inertial motions with headings that had a discrepancy from the visual heading randomly drawn from a uniform distribution ranging between ±90°.

Participants were instructed to estimate the heading of self-motion on every experimental trial using a pointer device. This device consisted of a 20cm stainless steel rod mounted to a potentiometer. One end of the rod was covered by 5cm of black heat shrink tubing; the other end was to be interpreted as an arrow's head and was to be pointed in the direction of the motion. The pointer device was free of discontinuities, and provided a <0.1° resolution. Because visual-only stimuli do not necessarily induce a sensation of self-motion, participants received the additional instruction to indicate the heading suggested by the visual stimulus for visual-only trials when no self-motion was perceived. For each participant, a total of 648(600) stimuli were presented in randomized order. Stimulus presentations were divided in blocks of 30 minutes, with compulsory breaks of at least 15 minutes in between. The experiment consisted of 3(2) 3(4)-hour sessions, which were completed over the course of 3(2) days.

Data Analyses

The model fitting procedures were performed for each participant individually. To assess general effects of motion profile characteristics on unisensory and multisensory heading estimation, we tested for trends in the observations made for individual participants.

Analyses of Unisensory Data. Unisensory visual and inertial heading estimation were described using a model that can account for periodic patterns of bias and dispersion as these are reported in the literature [17, 21, 22, 29].

The output of each sensory system is modeled as a realization x of a von Mises distributed random variable with mean μ and concentration parameter κ:

$$R = x \sim \mathcal{M}(\mu, \kappa) \tag{2}$$

with

$$\mu = \beta_0 + \mathrm{atan}(\beta_{1mT} \sin\theta, \ \cos\theta) \tag{2a}$$

$$\kappa = \gamma_{0mT} - \gamma_{1mT} |\sin 2\theta| \tag{2b}$$

here θ is the stimulus' heading, β_0, β_{1mT}, γ_{0mT}, and γ_{1mT} are free parameters, and atan represents the four-quadrant inverse tangent function. The subscript m refers to the sensory modality, which was visual or inertial, and the subscript T refers motion duration, which was 2, 4 or 6s. The final responses to the unisensory stimuli, R, are assumed to be equal to x.

Parameter β_0 was included to account for any small constant error that could arise due to factors such as a small offset in the position of a participant with respect to the simulator. This parameter was assumed to be constant for all conditions.

By either keeping the other six parameters (β_1, γ_0, γ_1 times two sensory modalities) constant or allowing them to vary for the different profiles, we end up with a total of $2^6 = 64$ different models. Each model was fitted to the data using the method of Maximum Likelihood. We exhaustively compared model fits by means of the Bayesian Information Criterion (BIC) scores [23]. For each participant, we chose as the final unisensory model the model with the lowest BIC score, for which the evidence was at least positive compared to alternative models with fewer parameters. Data points more than three standard deviations away from the mean (i.e., $\mu(\theta)$) were considered outliers in the unisensory conditions and excluded from the analysis.

To assess whether the characteristics of the motion profile affected the bias or dispersion of unisensory heading estimates, we looked for general trends in the parameters obtained for all the participants. Trends in the bias were assessed by testing whether parameter β_1 differed between profiles. Because dispersion is characterized by the combination of γ_0 and γ_1, and varies with heading angle, we needed a condensed measure that takes both parameters into account. Therefore, we calculated a total variance $\sigma^2_{tot.}$ for the different profiles (see supplementary material S1 Appendix). The values of $\sigma^2_{tot.}$ obtained were subsequently tested for differences between profiles.

Analyses of Multisensory Data. Heading estimation for multisensory stimuli was modeled according to a Multisensory Integration (MI) model, a model that reflects Sensory Segregation (SS), and three variants of a Causal Inference (CI) model [17]. The MI and SS models form extremes of a spectrum, for which multisensory signals are either completely integrated or segregated; the CI models come in between these extremes. Responses to multisensory stimuli were compared to model predictions to determine which model could best account for the observations. The CI models tested here are based upon the model originally proposed by

[9], and the adaptations of that model that account for different decision strategies by [11]; the present versions are further modified to account for the circular nature of heading (see: [17]).

In each multisensory model, final heading estimates are constructed from the internal representations x_V, x_I of the visual and inertial stimulus on the one hand, and knowledge of the size of their respective noises κ_V, κ_I (i.e., dispersions) on the other. Because internal representations cannot be directly measured, we used the fitted unisensory models (see section: Analyses of Unisensory Data) to simulate them. Consistent with the unisensory models, x_V and x_I are realizations of von Mises distributed random variables with a heading dependent bias and noise. For each experimental trial, 1000 x_V, x_I pairs were generated, from which final estimates were constructed according to each multisensory model. The likelihood of actual responses was subsequently derived from kernel density estimation [61] on the simulated data. Data points that differed more than ±90˚ from either the visual or inertial stimulus were considered outliers in the multisensory conditions.

We assume that the CNS is unaware of the relationship between the size of the noise and the stimulus' heading angle. This assumption may be justified by the fact that if this relationship *was* known to the CNS, this would provide information which could be used to avoid bias. However, as bias in heading estimation is a common experimental finding [17, 21, 22, 62–65], it seems reasonable to assume that this relationship is unknown to the CNS. Consequently, the relationship between the κ_V, κ_I and physical stimuli is not included in the multisensory models. For each experimental trial, the values of κ_V, κ_I given by the unisensory models were passed to the multisensory models. Because the values of κ_V, κ_I were equal for the 1000 simulations per experimental trial, they only enter the equations for the CI models (below) implicitly.

The response according to the MI model R_{MI} is the Maximum-A-Posteriori (MAP) heading estimate given the sensory signals and prior beliefs about the probability of different heading angles. Assuming a uniform prior on heading angle, R_{MI} can be expressed as

$$R_{MI} = \text{Arg}\left(\kappa_V e^{ix_V} + \kappa_I e^{ix_I}\right), \tag{3}$$

where the function Arg refers to the argument of a complex number, yielding an angle.

According to the SS model, the response R_{SS} is either the estimate carried by the visual or the inertial channel, chosen at random:

$$R_{SS} = \begin{cases} x_V, & \text{with probability } \xi \\ x_I, & \text{with probability } 1 - \xi \end{cases} \tag{4}$$

where ξ represents the probability that the visual channel dominates. ξ was treated as a free parameter.

According to the CI models, the CNS constructs a final estimate from R_{MI} and R_{SS}, while taking into account the posterior probability P(C|x_V, x_I) that the unisensory signals have a common cause given the actual sensory signals. P(C|x_V, x_I) is calculated as

$$P(C \mid x_V, x_I) = \frac{P(x_V, x_I \mid C)P(C)}{P(x_V, x_I)} \tag{5}$$

The likelihood P(x_V, x_I|C) represents the probability of observing the pair of visual and inertial signals x_V, x_I given that the signals have a common cause. The prior P(C) is interpreted as a tolerance for discrepancies, isolated from the sensory information. The probability of the observations P(x_V, x_I) is a normalizing factor that ensures that the posterior is a proper

probability density. It is calculated as

$$P(x_V, x_I) = P(C)P(x_V, x_I \mid C) + (1 - P(C))P(x_V, x_I \mid \bar{C}) \tag{6}$$

In each version of the CI model, knowledge of probability $P(C|x_V, x_I)$ is applied in a different way. In the Model Averaging (MA) variant, intermediate estimates R_{MI} and R_{SS} are weighted according to $P(C|x_V, x_I)$ and the complement of this probability $P(\bar{C} \mid x_V, x_I)$, respectively:

$$R_{\mathrm{MA}} = \mathrm{Arg}\left(P(C \mid x_V, x_I)e^{iR_{MI}} + P(\bar{C} \mid x_V, x_I)e^{iR_{SS}}\right) \tag{7}$$

In the Model Selection variant (MS), the final estimate equals R_{MI} or R_{SS}, depending on which causal structure is more likely: $P(C|x_V, x_I)$ or $P(\bar{C} \mid x_V, x_I)$, respectively:

$$R_{\mathrm{MS}} = \begin{cases} R_{MI}, & \text{if } P(C \mid x_V, x_I) \geq \frac{1}{2} \\[2mm] R_{SS}, & \text{if } P(C \mid x_V, x_I) < \frac{1}{2} \end{cases} \tag{8}$$

According to the Probability Matching (PM) variant, the final estimate also equals R_{MI} or R_{SS}, but is chosen at random with probabilities matching $P(C|x_V, x_I)$ and $P(\bar{C} \mid x_V, x_I)$.

$$R_{\mathrm{PM}} = \begin{cases} R_{MI}, & \text{with probability } P(C \mid x_V, x_I) \\[2mm] R_{SS}, & \text{with probability } P(\bar{C} \mid x_V, x_I) \end{cases} \tag{9}$$

In the MA, MS, and PM models, parameters $P(C)$ and ξ are equal for the different motion profile conditions. Consequently, these models cannot account for differences in multisensory responses between conditions that are not explained by characteristics of unisensory processing. To assess whether there are differences in multisensory processing, we assessed if the overall fit of the MA, MS, and PM models improved when the models were fitted to the different conditions individually; thus fitting three pairs of $P(C)$ and ξ parameters rather than one. We indicate the condition associated with a parameter estimate by adding a subscript T, $P_T(C)$ and ξ_T, where the subscript relates to the motion profile's duration T, with $T = 2, 4$ and $6s$. We interpret $P_T(C)$ as reflections of the tolerance for discrepancies after $T = 2, 4$ and $6s$ of motion, independent of information on discrepancy carried by the sensory signals themselves.

Supporting Information

S1 Fig. Response bias as a function of discrepancy. Plots of the bias in responses relative to visual heading angle (dots) as a function of discrepancy between visual and inertial heading angle. Dots with different colors represent data from different participants. The horizontal blue line indicates a perfect theoretical correspondence between responses and the visual heading angle; the diagonal orange line indicates a perfect theoretical correspondence between responses and the inertial heading angle. Discrepancy is defined as the angular difference between visual and inertial heading angle; bias is defined as the angular difference between the visual heading angle and the response. Respectively, the corresponding calculations are: $\mathrm{Arg}\left(e^{i\theta_V}/e^{i\theta_I}\right)$, and $\mathrm{Arg}\left(e^{i\theta_V}/e^{i\theta_R}\right)$.
(PDF)

S1 Table. Unisensory BIC scores. BIC scores for the null model (assuming no bias//0, and for the final model of unisensory perception.
(XLSX)

S2 Table. Unisensory model parameter estimates. Estimates of all unisensory model parameters β_0, β_{1mT}, γ_{0mT}, and γ_{1mT}, for modalities m visual and inertial, and T representing motion profile duration: 2, 4, or 6s.
(XLSX)

S3 Table. Multisensory BIC scores. BIC scores for the MI, SS, and CI models, fitted either with invariant parameters for different motion profiles, or with different parameters for different conditions. The latter fits are marked with the subscript T.
(XLSX)

S4 Table. Multisensory model parameters. Estimates of all multisensory model parameters.
(XLSX)

S1 Video. Video of simulator motion.
(AVI)

S1 Appendix. Calculation of total variance.
(PDF)

S1 Data. Experimental data.
(ZIP)

Acknowledgments

Author Contributions

Conceptualization: KNdW.

Data curation: KNdW.

Formal analysis: KNdW MK.

Funding acquisition: HHB.

Investigation: KNdW.

Methodology: KNdW.

Project administration: KNdW.

Resources: HHB.

Software: KNdW.

Supervision: KNdW.

Validation: KNdW.

Visualization: KNdW.

Writing – original draft: KNdW.

Writing – review & editing: KNdW MK HHB.

References

1. van Beers RJ, Sittig AC, van der Gon Denier JJ. How humans combine simultaneous proprioceptive and visual position information. Experimental Brain Research. 1996; 111(2):253–261. doi: 10.1007/BF00227302 PMID: 8891655

2. Ghahramani Z, Wolpert DM, Jordan MI. Computational models of sensorimotor integration. Advances in Psychology. 1997; 119:117–147. doi: 10.1016/S0166-4115(97)80006-4

3. van Beers RJ, Sittig AC, van der Gon JJD. Integration of proprioceptive and visual position-information: An experimentally supported model. Journal of Neurophysiology. 1999; 81(3):1355–1364. PMID: 10085361

4. Ernst MO, Banks MS. Humans integrate visual and haptic information in a statistically optimal fashion. Nature. 2002; 415(6870):429–433. doi: 10.1038/415429a PMID: 11807554

5. Ernst MO, Bülthoff HH. Merging the senses into a robust percept. Trends in Cognitive Sciences. 2004; 8(4):162–169. doi: 10.1016/j.tics.2004.02.002 PMID: 15050512

6. Alais D, Burr D. The ventriloquist effect results from near-optimal bimodal integration. Current biology. 2004; 14(3):257–262. doi: 10.1016/j.cub.2004.01.029 PMID: 14761661

7. Welch RB, Warren DH. Immediate perceptual response to intersensory discrepancy. Psychological bulletin. 1980; 88(3):638. doi: 10.1037/0033-2909.88.3.638 PMID: 7003641

8. Lewald J, Guski R. Cross-modal perceptual integration of spatially and temporally disparate auditory and visual stimuli. Cognitive brain research. 2003; 16(3):468–478. doi: 10.1016/S0926-6410(03)00074-0 PMID: 12706226

9. Körding KP, Beierholm U, Ma WJ, Quartz S, Tenenbaum JB, Shams L. Causal inference in multisensory perception. PLoS ONE. 2007; 2(9):e943. doi: 10.1371/journal.pone.0000943 PMID: 17895984

10. Sato Y, Toyoizumi T, Aihara K. Bayesian inference explains perception of unity and ventriloquism aftereffect: identification of common sources of audiovisual stimuli. Neural computation. 2007; 19(12):3335–3355. doi: 10.1162/neco.2007.19.12.3335 PMID: 17970656

11. Wozny DR, Beierholm UR, Shams L. Probability matching as a computational strategy used in perception. PLoS Computational Biology. 2010; 6(8):e1000871. doi: 10.1371/journal.pcbi.1000871 PMID: 20700493

12. Angelaki DE, Gu Y, DeAngelis GC. Visual and vestibular cue integration for heading perception in extrastriate visual cortex. The Journal of Physiology. 2011; 589(4):825–833. doi: 10.1113/jphysiol.2010.194720 PMID: 20679353

13. de Winkel KN, Weesie J, Werkhoven PJ, Groen EL. Integration of visual and inertial cues in perceived heading of self-motion. Journal of Vision. 2010; 10(12):1. doi: 10.1167/10.12.1 PMID: 21047733

14. Fetsch CR, Turner AH, DeAngelis GC, Angelaki DE. Dynamic reweighting of visual and vestibular cues during self-motion perception. The Journal of Neuroscience. 2009; 29(49):15601–15612. doi: 10.1523/JNEUROSCI.2574-09.2009 PMID: 20007484

15. Butler JS, Smith ST, Campos JL, Bülthoff HH. Bayesian integration of visual and vestibular signals for heading. Journal of Vision. 2010; 10(11):23. doi: 10.1167/10.11.23 PMID: 20884518

16. Butler JS, Campos JL, Bülthoff HH, Smith ST. The Role of Stereo Vision in Visual–Vestibular Integration. Seeing and Perceiving. 2011; 24(5):453–470. doi: 10.1163/187847511X588070 PMID: 21888763

17. de Winkel KN, Katliar M, Bülthoff HH. Forced Fusion in Multisensory Heading Estimation. PLoS ONE. 2015; 10(5):e0127104. doi: 10.1371/journal.pone.0127104 PMID: 25938235

18. Fitzpatrick RC, Day BL. Probing the human vestibular system with galvanic stimulation. Journal of Applied Physiology. 2004; 96(6):2301–2316. doi: 10.1152/japplphysiol.00008.2004 PMID: 15133017

19. Butler JS, Campos JL, Bülthoff HH. Optimal visual–vestibular integration under conditions of conflicting intersensory motion profiles. Experimental brain research. 2015; 233(2):587–597. doi: 10.1007/s00221-014-4136-1 PMID: 25361642

20. Nieuwenhuizen FM, Bülthoff HH. The MPI CyberMotion Simulator: A Novel Research Platform to Investigate Human Control Behavior. Journal of Computing Science and Engineering. 2013; 7(2):122–131. doi: 10.5626/JCSE.2013.7.2.122

21. Crane BT. Direction specific biases in human visual and vestibular heading perception. PLoS ONE. 2012; 7(12):e51383. doi: 10.1371/journal.pone.0051383 PMID: 23236490

22. Cuturi LF, MacNeilage PR. Systematic biases in human heading estimation. PLoS ONE. 2013; 8(2):e56862. doi: 10.1371/journal.pone.0056862 PMID: 23457631

23. Schwarz G. Estimating the Dimension of a Model. The Annals of Statistics. 1978; 6(2):461–464. doi: 10.1214/aos/1176344136

24. Walsh E. Role of the vestibular apparatus in the perception of motion on a parallel swing. The Journal of physiology. 1961; 155(3):506–513. doi: 10.1113/jphysiol.1961.sp006643 PMID: 13782902

25. Valko Y, Lewis RF, Priesol AJ, Merfeld DM. Vestibular labyrinth contributions to human whole-body motion discrimination. The Journal of Neuroscience. 2012; 32(39):13537–13542. doi: 10.1523/JNEUROSCI.2157-12.2012 PMID: 23015443

26. Green DM, Swets JA. Signal Detection Theory and Psychophysics. John Wiley and Sons; 1966. Available from: https://books.google.de/books?id = yYsQAQAAIAAJ

27. Hummel N, Cuturi LF, MacNeilage PR, Flanagin VL. The effect of supine body position on human heading perception. Journal of vision. 2016; 16(3):19–19. doi: 10.1167/16.3.19 PMID: 26891826

28. Fernandez C, Goldberg JM. Physiology of peripheral neurons innervating otolith organs of the squirrel monkey. I. Response to static tilts and to long-duration centrifugal force. Journal of neurophysiology. 1976; 39(5):970–984. PMID: 824412

29. Gu Y, Fetsch CR, Adeyemo B, DeAngelis GC, Angelaki DE. Decoding of MSTd population activity accounts for variations in the precision of heading perception. Neuron. 2010; 66(4):596–609. doi: 10.1016/j.neuron.2010.04.026 PMID: 20510863

30. Ernst MO. A Bayesian view on multimodal cue integration. Human body perception from the inside out. 2006; 131:105–131.

31. Bresciani JP, Dammeier F, Ernst MO. Vision and touch are automatically integrated for the perception of sequences of events. Journal of Vision. 2006; 6(5):2–2. doi: 10.1167/6.5.2 PMID: 16881788

32. Beierholm U, Shams L, Ma WJ, Koerding K. Comparing Bayesian models for multisensory cue combination without mandatory integration. In: Advances in neural information processing systems; 2007. p. 81–88.

33. Wei K, Kording KP. Causal Inference in Sensorimotor Learning. In: Trommershauser J, Kording K, Landy MS, editors. Sensory cue integration. Oxford: Oxford University Press; 2011. p. 30–45.

34. Shams L, Beierholm U. Humans' multisensory perception, from integration to segregation, follows Bayesian inference. In: Trommershauser J, Kording K, Landy MS, editors. Sensory cue integration. Oxford: Oxford University Press; 2011. p. 251–262.

35. Yuille AL, Bülthoff HH. Bayesian decision theory and psychophysics. In: Knill DC, Richards W, editors. Perception as Bayesian inference. Cambridge: Cambridge University Press; 1996. p. 123–161.

36. Knill DC. Mixture models and the probabilistic structure of depth cues. Vision research. 2003; 43 (7):831–854. doi: 10.1016/S0042-6989(03)00003-8 PMID: 12639607

37. Knill DC. Robust cue integration: A Bayesian model and evidence from cue-conflict studies with stereoscopic and figure cues to slant. Journal of Vision. 2007; 7(7):5–5. doi: 10.1167/7.7.5 PMID: 17685801

38. Landy MS, Banks MS, Knill DC. Ideal-observer models of cue integration. In: Trommershauser J, Kording K, Landy MS, editors. Sensory cue integration. Oxford: Oxford University Press; 2011. p. 5–29.

39. Roach NW, Heron J, McGraw PV. Resolving multisensory conflict: a strategy for balancing the costs and benefits of audio-visual integration. Proceedings of the Royal Society of London B: Biological Sciences. 2006; 273(1598):2159–2168. doi: 10.1098/rspb.2006.3578 PMID: 16901835

40. Shams L, Ma WJ, Beierholm U. Sound-induced flash illusion as an optimal percept. Neuroreport. 2005; 16(17):1923–1927. doi: 10.1097/01.wnr.0000187634.68504.bb PMID: 16272880

41. Rohe T, Noppeney U. Cortical hierarchies perform Bayesian causal inference in multisensory perception. PLoS Biol. 2015; 13(2):e1002073. doi: 10.1371/journal.pbio.1002073 PMID: 25710328

42. Drugowitsch J, DeAngelis GC, Klier EM, Angelaki DE, Pouget A. Optimal multisensory decision-making in a reaction-time task. Elife. 2014; 3:e03005. doi: 10.7554/eLife.03005 PMID: 24929965

43. Ratcliff R, McKoon G. The diffusion decision model: theory and data for two-choice decision tasks. Neural computation. 2008; 20(4):873–922. doi: 10.1162/neco.2008.12-06-420 PMID: 18085991

44. O'Connell RG, Dockree PM, Kelly SP. A supramodal accumulation-to-bound signal that determines perceptual decisions in humans. Nature neuroscience. 2012; 15(12):1729–1735. doi: 10.1038/nn.3248 PMID: 23103963

45. Kira S, Yang T, Shadlen MN. A Neural Implementation of Wald's Sequential Probability Ratio Test. Neuron. 2015; 85(4):861–873. doi: 10.1016/j.neuron.2015.01.007 PMID: 25661183

46. Bubka A, Bonato F, Palmisano S. Expanding and contracting optic-flow patterns and vection. Perception. 2008; 37(5):704–711. doi: 10.1068/p5781 PMID: 18605144

47. Brandt T, Dichgans J, Büchele W. Motion habituation: inverted self-motion perception and optokinetic after-nystagmus. Experimental Brain Research. 1974; 21(4):337–352. doi: 10.1007/BF00237897 PMID: 4442493

48. Cuturi LF, MacNeilage PR. Optic flow induces nonvisual self-motion aftereffects. Current Biology. 2014; 24(23):2817–2821. doi: 10.1016/j.cub.2014.10.015 PMID: 25454589

49. Bos JE, Bles W, Groen EL. A theory on visually induced motion sickness. Displays. 2008; 29(2):47–57. doi: 10.1016/j.displa.2007.09.002

50. Mayne R. A systems concept of the vestibular organs. In: Vestibular system part 2: psychophysics, applied aspects and general interpretations. Springer; 1974. p. 493–580. doi: 10.1007/978-3-642-65920-1_14

51. Bos JE, Bles W. Theoretical considerations on canal–otolith interaction and an observer model. Biological cybernetics. 2002; 86(3):191–207. doi: 10.1007/s00422-001-0289-7 PMID: 12068786

52. Bos JE, Correia Grácio BJ. Perceived radial translation during centrifugation. Journal of Vestibular Research. 2015; 25(3,4):119–124. doi: 10.3233/VES-150555 PMID: 26756127

53. Grácio BC, De Winkel K, Groen E, Wentink M, Bos J. The time constant of the somatogravic illusion. Experimental brain research. 2013; 224(3):313–321. doi: 10.1007/s00221-012-3313-3

54. Seidman S, Telford L, Paige G. Tilt perception during dynamic linear acceleration. Experimental brain research. 1998; 119(3):307–314. doi: 10.1007/s002210050346 PMID: 9551831

55. De Graaf B, Bos J, Tielemans W, Rameckers F, Rupert A, Guedry F. Otolith Contribution to Ocular Torsion and Spatial Orientation During Acceleration. DTIC Document; 1996.

56. Glasauer S. Linear acceleration perception: frequency dependence of the hilltop illusion. Acta Oto-Laryngologica. 1995; 115(sup520):37–40. doi: 10.3109/00016489509125184 PMID: 8749075

57. Merfeld D, Zupan L, Gifford C. Neural processing of gravito-inertial cues in humans. II. Influence of the semicircular canals during eccentric rotation. Journal of Neurophysiology. 2001; 85(4):1648–1660. PMID: 11287488

58. Correia Grácio BJ, Pais ARV, Van Paassen M, Mulder M, Kelly LC, Houck JA. Optimal and coherence zone comparison within and between flight simulators. Journal of Aircraft. 2013; 50(2):493–507. doi: 10.2514/1.C031870

59. Groen EL, V Valenti Clari MS, AW Hosman RJ. Evaluation of perceived motion during a simulated take-off run. Journal of Aircraft. 2001; 38(4):600–606. doi: 10.2514/2.2827

60. Valente Pais AR, Van Paassen M, Mulder M, Wentink M. Perception coherence zones in flight simulation. Journal of Aircraft. 2010; 47(6):2039–2048. doi: 10.2514/1.C000281

61. Atanasiu V. Kernel smoothing density estimate for circular data; 2011. Available from: http://de.mathworks.com/matlabcentral/fileexchange/32614-kernel-smoothing-density-estimate-for-circular-data/content/circ_ksdensity.m

62. D'Avossa G, Kersten D. Evidence in human subjects for independent coding of azimuth and elevation for direction of heading from optic flow. Vision Research. 1996; 36(18):2915–2924. doi: 10.1016/0042-6989(96)00010-7 PMID: 8917793

63. Johnston IR, White GR, Cumming RW. The role of optical expansion patterns in locomotor control. The American Journal of Psychology. 1973; p. 311–324. doi: 10.2307/1421439 PMID: 4775379

64. Li L, Peli E, Warren WH, et al. Heading perception in patients with advanced retinitis pigmentosa. Optometry & Vision Science. 2002; 79(9):581–589. doi: 10.1097/00006324-200209000-00009 PMID: 12322928

65. Llewellyn KR. Visual guidance of locomotion. Journal of Experimental Psychology. 1971; 91(2):245. doi: 10.1037/h0031788 PMID: 5134668