Evidence accumulation across the senses in the face of causal uncertainty

Jochem Beurskens*
Donders Centre for Neuroscience (DCN), Radboud University
Heyendaalseweg 135, Nijmegen, The Netherlands
jochem.beurskens@donders.ru.nl
Agata Wlaszczyk
Donders Centre for Neuroscience (DCN), Radboud University
Heyendaalseweg 135, Nijmegen, The Netherlands
agata.wlaszczyk@donders.ru.nl
Uta Noppeney
Donders Centre for Neuroscience (DCN), Radboud University
Heyendaalseweg 135, Nijmegen, The Netherlands
uta.noppeney@donders.ru.nl

Abstract

Humans need to make accurate and timely decisions based on constant influx of noisy sensory signals. They should integrate signals from common causes, but segregate those from separate causes. Previous research has shown that the brain arbitrates between integration and segregation consistent with Bayesian Causal Inference models (BCI). However, these static models ignored the dynamics of perceptual decision making and therefore could not account for response times. Using psychophysics, we show that the influence of spatially disparate visual signals on observers' perceived sound location declines with longer response times. This pattern is best captured by a dynamic BCI model that accumulates evidence jointly about the signals' locations and their causal structure (i.e. common vs. independent causes) over time in a forgetful fashion, until a decisional threshold is reached. By accounting for both response choices and times these dynamic BCI models advance our understanding of how observers dynamically combine signals from multiple sensory modalities in the face of causal uncertainty. They provide a novel perspective on previous neuroimaging results showing a progression from fusion to BCI multisensory interactions along cortical pathways.

Keywords: Bayesian causal inference; Evidence accumulation; Multisensory integration;

Introduction

In our natural environment observers need to make accurate and timely decisions based on a constant influx of noisy signals. A wealth of work in unisensory perception has shown that observers accumulate multiple sensory samples until a decisional threshold is reached consistent with normative sequential sampling models (Gold & Shadlen, 2007). Further, observers have been shown to integrate sensory signals from common causes weighted by their momentary reliabilities and accumulate this integrated evidence over time (Drugowitsch et al., 2014). However, in natural environment observers do not a priori know whether signals come from common sources and should hence be integrated. They need to infer the signals' causal structure from noisy cross-sensory correspondence cues such as signals happening at the same time or location (Körding et al., 2007; Noppeney, 2021). In the face of this causal uncertainty observers need to concurrently accumulate evidence about environmental properties such as an object's location and the signals' causal structure leading to complex non-linearities in the decision process.

This study combines psychophysics in human observers and Bayesian modelling to investigate how the brain accumulates evidence about the location of brief (50 ms) and long (until response with a maximum of 2000 ms) stimuli from auditory and visual senses under causal uncertainty into timely and accurate perceptual decisions.

Methods

Experimental procedure

In a spatial ventriloquist paradigm 6 observers were presented with synchronous, spatially congruent and disparate audiovisual signals, sampled independently from 4 equally spaced locations, ranging from -10.5° to 10.5° along the azimuth (stimulus duration: 50 ms or until response with a maximum of 2000 ms). In blocks observers reported either their perceived A or V location via 4-alternative keypress.

Generative and Recognition models

We built two dynamic Bayesian Causal Inference models (figure 1A, 1B). Each BCI model assumed that common and independent causal structures (C=1 or C=2) are sampled from a binomial distribution, defined by a causal prior p_{common} (Körding et al., 2007). In the evidence accumulation model, A and V stimulus locations (sA_t, sV_t) are constant, sampled jointly (C=1) or independently (C=2) from a spatial prior distribution $N(\mu_p, \sigma_p)$. In the forgetfulness model, (sA_t, sV_t) follow a joint or two independent random walks (e.g. $sA_t \; N(sA_{t-1}, \sigma^2_{noise})$. On each trial, these stimulus locations generate a series of conditionally independent noisy sensory observations $(xA_{\leq t}, xV_{\leq t})$. Given these noisy observations, the observer computes the posterior distribution over the causal structure C and the A location (similarly for V location)(Yu et al., 2009): $P(sA, C|xA_{\leq t}, xV_{\leq t}) = \frac{P(xA_t, xV_t|sA,C)P(sA,C|xA_{\leq t-1}, xV_{\leq t-1})}{\sum\limits_C \int P(xA_t, xV_t|sA,C)P(sA,C|xA_{\leq t-1}, xV_{\leq t-1})\, dsA}$

Observers' spatial responses are obtained:

$P(sA|xA_{\leq t}, xV_{\leq t}) = \sum_C P(sA|C, xA_{\leq t}, xV_{\leq t}) P(C|xA_{\leq t}, xV_{\leq t})$ with:

$P(sA|C=1, xA_{\leq t}, xV_{\leq t}) = N\left(\mu_{AV,t}, \sigma^2_{AV,t}\right)$;

$\mathrm{P}(sA|C=2, xA_{\leq t}, xV_{\leq t}) = N\left(\mu_{A,t}, \sigma^2_{A,t}\right)$

In the evidence accumulation model, the mean and variance are recursively updated:

$\mu_{AV,t} = \sigma^2_{AV,t}\left(\frac{xV_t}{\sigma^2_{xV}} + \frac{xA_t}{\sigma^2_{xA}} + \frac{\mu_{AV,t-1}}{\sigma^2_{AV,t-1}}\right)$; $\sigma^2_{AV,t} = \frac{1}{\frac{1}{\sigma^2_{xV}} + \frac{1}{\sigma^2_{xA}} + \frac{1}{\sigma^2_{AV,t-1}}}$, and $\mu_{A,t} = \sigma^2_{A,t}\left(\frac{xA_t}{\sigma^2_{xA}} + \frac{\mu_{A,t-1}}{\sigma^2_{A,t-1}}\right)$; $\sigma^2_{A,t} = \frac{1}{\frac{1}{\sigma^2_{xA}} + \frac{1}{\sigma^2_{A,t-1}}}$

In the forgetfulness model, the posteriors are perturbed by process noise $\sigma^2_{AV,t} = \sigma^2_{AV,t-1} + \sigma^2_{noise}$.

Model Responses

The observers decide whether to respond or continue sampling on the basis of the rate of change of either the maximum value, or the normalised Shannon entropy of the evolving posterior distribution(Li & Ma, 2020). They report the mean or the MAP of the evolving posterior distribution as the final A or V location estimate ($\hat{s}A$ or $\hat{s}V$) (figure 1B). The continuous location read outs are mapped onto the closest button responses for comparison with observers' discrete response data.

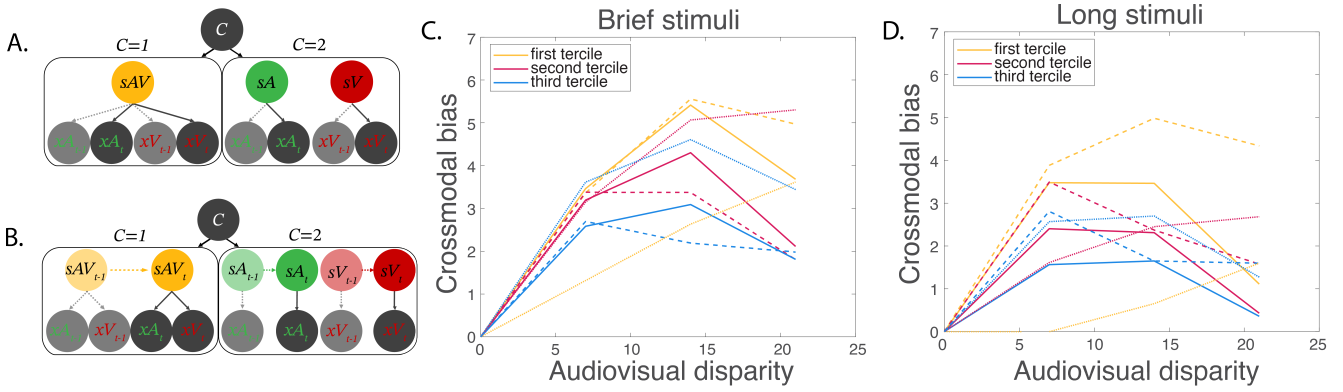

Figure 1: (A.) The generative model for dynamic Bayesian causal inference, and (B.) for the forgetfulness Bayesian Causal Inference model. The crossmodal bias for participant responses (continuous line) and the responses for evidence accumulation model (dotted line) and the forgetfulness model (dashed line) for the brief stimuli (C.), and the long stimuli (D.).

Results

Behaviour

The crossmodal bias (CMB) (e.g. for A report: $(Areport_{sA=x,sV=y} - Areport_{sA=sV=y})$ quantifies the influence of the task-irrelevant sensory (e.g. visual) signal on observers' (e.g. auditory) spatial report. CMB decreases with i. spatial disparity and ii. response times for both brief and long stimuli (figure 1C, 1D).

Modelling

Both dynamic BCI models account for observers' CMB profile across RT and disparity by spatial uncertainty principles. Initially, the models face large spatial and hence causal uncertainty, so that the bimodal posterior and hence the spatial estimate resembles a vision-dominated fusion estimate. With increasing number of samples this spatial uncertainty gradually resolves and the bimodal posterior has more probability mass over the true auditory location (figure 2). Thus, the progression from fusion to segregation and hence the reduction in CMB for longer RT naturally arises from a reduction in causal uncertainty. Crucially, because of process noise the forgetfulness model maintains a broader posterior even for later response times, being able to flexible adapt to new sensory inputs.

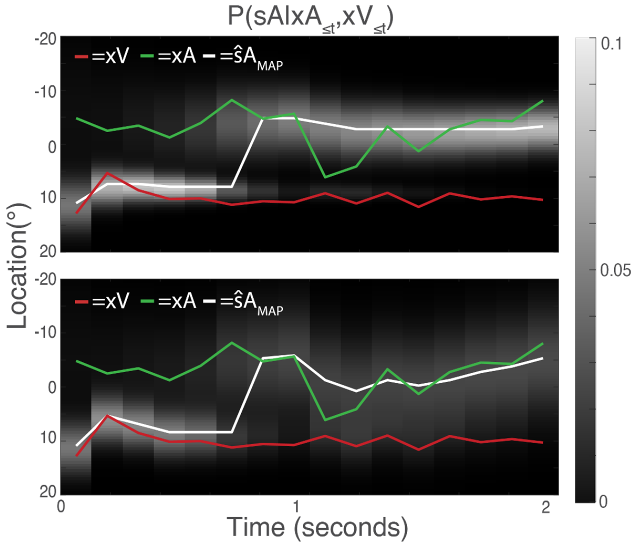

Figure 2: Simulated evolving posteriors, and MAP readout for both the evidence accumulation (top), and the forgetfulness model (bottom), based on noisy observations $(xA_{\leq t}, xV_{\leq t})$.

Factorial model comparison at the group level reveals superior performance of the forgetfulness model for both brief and long stimuli $(P_{exp,b} = 0.9750,\ P_{exp,l} = 0.9749)$(Acerbi et al., 2018).

Discussion

Previous static Bayesian Causal Inference models explained how the brain arbitrates between sensory integration and segregation(Körding et al., 2007). By ignoring the dynamics of perceptual decision making they provided predictions only for response choices but not for response times. Dynamic BCI can explain how observers' response choices and crossmodal biases depend on their response times. Accumulating sensory evidence about the signals' locations progressively resolves observers' spatial uncertainty, and thereby their uncertainty about the signals' causal structure. This decline in causal uncertainty in turn reduces the influence of task-irrelevant spatially disparate inputs from other sensory modalities. Further, even though the true stimulus locations are constant within a trial, observers accumulated evidence in a forgetful fashion, i.e. using a strategy that is optimal in a dynamic world with time-varying stimulus locations. From a neural perspective, dynamic BCI can also explain the progression from earlier sensory fusion estimates in posterior parietal cortices to Bayesian Causal inference estimates in anterior parietal cortices (Aller & Noppeney, 2019; Rohe et al., 2019; Cao et al., 2019).

Acknowledgements

This research was funded by the ERC Advanced Grant 'MakingSense' (101096659).

References

Acerbi, L., Dokka, K., Angelaki, D. E., & Ma, W. J. (2018). Bayesian comparison of explicit and implicit causal inference strategies in multisensory heading perception. *PLoS computational biology*, *14*(7), e1006110.

Aller, M., & Noppeney, U. (2019). To integrate or not to integrate: Temporal dynamics of hierarchical bayesian causal inference. *PLoS biology*, *17*(4), e3000210.

Cao, Y., Summerfield, C., Park, H., Giordano, B. L., & Kayser, C. (2019). Causal Inference in the Multisensory Brain. *Neuron*, *102*(5), 1076–1087.e8.

Drugowitsch, J., DeAngelis, G. C., Klier, E. M., Angelaki, D. E., & Pouget, A. (2014). Optimal multisensory decision-making in a reaction-time task. *Elife*, *3*, e03005.

Gold, J. I., & Shadlen, M. N. (2007). The neural basis of decision making. *Annu. Rev. Neurosci.*, *30*, 535–574.

Körding, K. P., Beierholm, U., Ma, W. J., Quartz, S., Tenenbaum, J. B., & Shams, L. (2007). Causal Inference in Multisensory Perception. *PLoS ONE*, *2*(9), e943.

Li, H.-H., & Ma, W. J. (2020). Confidence reports in decision-making with multiple alternatives violate the bayesian confidence hypothesis. *Nature communications*, *11*(1), 2004.

Noppeney, U. (2021). Perceptual Inference, Learning, and Attention in a Multisensory World. *Annual Review of Neuroscience*, *44*(1), 449–473.

Rohe, T., Ehlis, A.-C., & Noppeney, U. (2019). The neural dynamics of hierarchical bayesian causal inference in multisensory perception. *Nature communications*, *10*(1), 1907.

Yu, A. J., Dayan, P., & Cohen, J. D. (2009). Dynamics of attentional selection under conflict: Toward a rational Bayesian account. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(3), 700–717.