COMMENTARY

The influence of early audiovisual experience on multisensory integration and causal inference (commentary on Smyre et al., 2021)

Uta Noppeney [ID]

Donders Institute for Brain, Cognition and Behaviour, Nijmegen, The Netherlands

Correspondence
Uta Noppeney, Donders Institute for Brain, Cognition and Behaviour, Nijmegen, The Netherlands.
Email: u.noppeney@donders.ru.nl

In everyday life, our senses are flooded with a plethora of sensory inputs. To interact effectively, the brain should combine signals when they come from common events and process signals independently when they arise from different events. Integrating information across the senses is crucial for an animal's survival, enabling faster and more accurate responses. A critical question is how the ability to integrate signals across the senses depends on sensory experience during neurodevelopment.

Previous elegant neurophysiological work has shown that early audiovisual experience is critical for the alignment of spatial receptive fields across visual (eye-centred) and auditory (head-centred) reference frames (Hyde & Knudsen, 2002; Stein et al., 2014a). Neurons in the superior colliculus of dark-reared cats failed to show a nonlinear response amplification for multisensory relative to unisensory stimuli, typically observed in normally reared cats (Stein et al., 2014b; Wallace et al., 2004). Likewise, selective exposure to independently sampled auditory and visual signals during neurodevelopment reduces the fraction of audiovisual neurons with superadditive responses to audiovisual stimuli (Xu et al., 2012). By contrast, exposure to collocated audiovisual signals appears to sharpen the spatio-temporal window in which superadditive responses can be observed.

A recent study in EJN (Smyre et al., 2021) started to probe the functional relevance of those neurophysiological changes. As one example of abnormal audiovisual statistics, it investigated how dark rearing influences an animal's audiovisual localization behaviour. The study presented cats, reared normally or in darkness, with auditory, visual or collocated audiovisual stimuli from six locations along the azimuth. Further, catch trials were included, on which no stimuli were presented. The animal had to approach the location of the stimulus or refrain from making a response on catch trials. Therefore, the task combined stimulus detection and spatial localization components. For stimulus detection, a multisensory benefit was observed predominantly for normally reared animals. Dark-reared animals were worse at detecting audiovisual stimuli than normally reared animals despite comparable or even better detection of unisensory stimuli.

A more complex picture arose for spatial localization, as here dark rearing profoundly affected localization accuracy even under unisensory stimulation. To enable a comparison between normally and dark-reared animals, the study developed predictions for audiovisual spatial localization using a statistical facilitation model. Inspired by the classical race model typically used for response time analyses (Miller, 1982; Otto & Mamassian, 2012), the authors applied the statistical facilitation model to response choices, thereby enabling predictions for audiovisual localization accuracy under independent processing assumptions. Specifically, the facilitation model sampled auditory and visual estimates independently and selected the more accurate one for making an overt response. Because in

reality the brain cannot assess the accuracy of its own auditory and visual estimates, the model's predictions form an extreme upper bound for audiovisual localization accuracy under the assumption of independent processing. Nevertheless, the localization accuracy of normally reared animals exceeded the statistical facilitation predictions. By contrast, the localization performance of dark-reared animals did not pass this upper bound, suggesting that the development of audiovisual integration abilities relies on exposure to audiovisual inputs during sensitive periods of neurodevelopment.

Further insights into how exposure to abnormal audiovisual spatial statistics affects localization behaviour can be provided from a normative Bayesian perspective (Ernst & Banks, 2002; Körding et al., 2007; Noppeney, 2021). A Bayesian observer sets a benchmark of optimal performance against which observers' behaviour can be compared. According to normative Bayesian inference, observers should combine signals that are known to come from one source weighted by their relative reliabilities (i.e. inverse of noise or variance) into more precise perceptual estimates (Ernst & Banks, 2002; Fetsch et al., 2012, 2013). Indeed, accumulating research has shown that human observers integrate spatial signals from vision and audition near-optimally into more reliable spatial estimates (Alais & Burr, 2004) though modest deviations from optimality have been noted (Meijer et al., 2019).

Recent modelling efforts have moved towards more complex situations, in which signals can come from common or separate causes. In these situations with causal uncertainty, observers need to infer the signals' causal structure by combining noisy correspondence cues (e.g. spatial disparity or temporal synchrony) with prior expectations about the signals' causal structure. Intriguingly, human and non-human observers have been shown to gracefully transition from sensory integration to segregation for increasing audiovisual spatial disparity or asynchrony consistent with normative models of Bayesian Causal Inference (Acerbi et al., 2018; Aller & Noppeney, 2019; Cao et al., 2019; Körding et al., 2007; Magnotti et al., 2013; Mohl et al., 2020; Rohe et al., 2019; Rohe & Noppeney, 2015a, 2015b; Wozny et al., 2010). Observers integrate audiovisual signals when they are close in time and space, yet keep them separate at large spatio-temporal conflicts.

A central aim for future research is to define the role of early audiovisual experience in shaping how animals perform causal inference and arbitrate between sensory integration and segregation. How does exposure to abnormal audiovisual spatial statistics affect the key computational ingredients of Bayesian causal inference? From the perspective of efficient coding (Wei & Stocker, 2015), the brain may mould receptive fields and tuning functions to optimize the sensory representations to the abnormal audiovisual input statistics—which may in turn alter Bayesian causal inference via changes in the likelihood function. For instance, less precise spatial representations would make the arbitration between sensory integration and segregation less reliable. Further, exposure to independently sampled auditory and visual signals may reduce binding of spatio-temporally coincident audiovisual signals. Critically, however, profound changes in sensory experience during neurodevelopment may not only affect key parameters of Bayesian causal inference but even result in multisensory processing that deviates from normative computational principles.

CONFLICT OF INTEREST
The author declare no conflict of interest.

PEER REVIEW
The peer review history for this article is available at https://publons.com/publon/10.1111/ejn.15576.

ORCID
Uta Noppeney https://orcid.org/0000-0002-7697-2290

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*, e1006110–e1006110. https://doi.org/10.1371/journal.pcbi.1006110

Alais, D., & Burr, D. (2004). Ventriloquist effect results from near-optimal bimodal integration. *Current Biology*, *14*, 257–262. https://doi.org/10.1016/j.cub.2004.01.029

Aller, M., & Noppeney, U. (2019). To integrate or not to integrate: Temporal dynamics of hierarchical Bayesian causal inference. *PLoS Biology*, *17*, e3000210–e3000210. https://doi.org/10.1371/journal.pbio.3000210

Cao, Y., Summerfield, C., Park, H., Giordano, B. L., & Kayser, C. (2019). Causal inference in the multisensory brain. *Neuron*, *102*, 1076–1087, e1078. https://doi.org/10.1016/j.neuron.2019.03.043

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*, 429–433. https://doi.org/10.1038/415429a

Fetsch, C. R., DeAngelis, G. C., & Angelaki, D. E. (2013). Bridging the gap between theories of sensory cue integration and the physiology of multisensory neurons. *Nature Reviews Neuroscience*, *14*, 429–442. https://doi.org/10.1038/nrn3503

Fetsch, C. R., Pouget, A., DeAngelis, G. C., & Angelaki, D. E. (2012). Neural correlates of reliability-based cue weighting during multisensory integration. *Nature Neuroscience*, *15*, 146–154. https://doi.org/10.1038/nn.2983

Hyde, P. S., & Knudsen, E. I. (2002). The optic tectum controls visually guided adaptive plasticity in the owl's auditory space map. *Nature*, *415*, 73–76. https://doi.org/10.1038/415073a

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

Magnotti, J. F., Ma, W. J., & Beauchamp, M. S. (2013). Causal inference of asynchronous audiovisual speech. *Frontiers in Psychology*, *4*, 798–798. https://doi.org/10.3389/fpsyg.2013.00798

Meijer, D., Veselič, S., Calafiore, C., & Noppeney, U. (2019). Integration of audiovisual spatial signals is not consistent with maximum likelihood estimation. *Cortex*, *119*, 74–88. https://doi.org/10.1016/j.cortex.2019.03.026

Miller, J. (1982). Divided attention: Evidence for coactivation with redundant signals. *Cognitive Psychology*, *14*, 247–279. https://doi.org/10.1016/0010-0285(82)90010-X

Mohl, J. T., Pearson, J. M., & Groh, J. M. (2020). Monkeys and humans implement causal inference to simultaneously localize auditory and visual stimuli. *Journal of Neurophysiology*, *124*, 715–727. https://doi.org/10.1152/jn.00046.2020

Noppeney, U. (2021). Perceptual inference, learning, and attention in a multisensory world. *Annual Review of Neuroscience*, *44*, 449–473. https://doi.org/10.1146/annurev-neuro-100120-085519

Otto, T. U., & Mamassian, P. (2012). Noise and correlations in parallel perceptual decision making. *Current Biology*, *22*, 1391–1396. https://doi.org/10.1016/j.cub.2012.05.031

Rohe, T., Ehlis, A.-C., & Noppeney, U. (2019). The neural dynamics of hierarchical Bayesian causal inference in multisensory perception. *Nature Communications*, *10*, 1907. https://doi.org/10.1038/s41467-019-09664-2

Rohe, T., & Noppeney, U. (2015a). Cortical hierarchies perform Bayesian causal inference in multisensory perception. *PLoS Biology*, *13*, e1002073–e1002073. https://doi.org/10.1371/journal.pbio.1002073

Rohe, T., & Noppeney, U. (2015b). Sensory reliability shapes perceptual inference via two mechanisms. *Journal of Vision*, *15*, 22–22. https://doi.org/10.1167/15.5.22

Smyre, S. A., Wang, Z., Stein, B. E., & Rowland, B. A. (2021). Multisensory enhancement of overt behavior requires multisensory experience. *The European Journal of Neuroscience*, *54*, 4514–4527. https://doi.org/10.1111/ejn.15315

Stein, B. E., Stanford, T. R., & Rowland, B. A. (2014a). Development of multisensory integration from the perspective of the individual neuron. *Nature Reviews. Neuroscience*, *15*, 520–535. https://doi.org/10.1038/nrn3742

Stein, B. E., Stanford, T. R., & Rowland, B. A. (2014b). Development of multisensory integration from the perspective of the individual neuron. *Nature Reviews. Neuroscience*, *15*, 520–535. https://doi.org/10.1038/nrn3742

Wallace, M. T., Perrault, T. J. Jr., Hairston, W. D., & Stein, B. E. (2004). Visual experience is necessary for the development of multisensory integration. *The Journal of Neuroscience*, *24*, 9580–9584. https://doi.org/10.1523/JNEUROSCI.2535-04.2004

Wei, X. X., & Stocker, A. A. (2015). A Bayesian observer model constrained by efficient coding can explain 'anti-Bayesian' percepts. *Nature Neuroscience*, *18*, 1509–1517. https://doi.org/10.1038/nn.4105

Wozny, D. R., Beierholm, U. R., & Shams, L. (2010). Probability matching as a computational strategy used in perception. *PLoS Computational Biology*, *6*, e1000871. https://doi.org/10.1371/journal.pcbi.1000871

Xu, J., Yu, L., Rowland, B. A., Stanford, T. R., & Stein, B. E. (2012). Incorporating cross-modal statistics in the development and maintenance of multisensory integration. *The Journal of Neuroscience*, *32*, 2287–2298. https://doi.org/10.1523/JNEUROSCI.4304-11.2012