



# Older adults sacrifice response speed to preserve multisensory integration performance



Samuel A. Jones<sup>a,b,\*</sup>, Ulrik Beierholm<sup>c</sup>, David Meijer<sup>a</sup>, Uta Noppeney<sup>a</sup>

<sup>a</sup> Computational Cognitive Neuroimaging Laboratory, Computational Neuroscience and Cognitive Robotics Centre, University of Birmingham, Birmingham, UK

<sup>b</sup> The Staffordshire Centre for Psychological Research, Staffordshire University, Stoke-on-Trent, UK

<sup>c</sup> Department of Psychology, University of Durham, Durham, UK

## ARTICLE INFO

### Article history:

Received 19 November 2018

Received in revised form 10 July 2019

Accepted 17 August 2019

Available online 26 August 2019

### Keywords:

Aging

Multisensory integration

Multisensory perception

Decision making

Ventriloquist effect

Bayesian model

Bayesian causal inference

Evidence accumulation

## ABSTRACT

Aging has been shown to impact multisensory perception, but the underlying computational mechanisms are unclear. For effective interactions with the environment, observers should integrate signals that share a common source, weighted by their reliabilities, and segregate those from separate sources. Observers are thought to accumulate evidence about the world's causal structure over time until a decisional threshold is reached. Combining psychophysics and Bayesian modeling, we investigated how aging affects audiovisual perception of spatial signals. Older and younger adults were comparable in their final localization and common-source judgment responses under both speeded and unspeeded conditions, but were disproportionately slower for audiovisually incongruent trials. Bayesian modeling showed that aging did not affect the ability to arbitrate between integration and segregation under either unspeeded or speeded conditions. However, modeling the within-trial dynamics of evidence accumulation under speeded conditions revealed that older observers accumulate noisier auditory representations for longer, set higher decisional thresholds, and have impaired motor speed. Older observers preserve audiovisual localization performance, despite noisier sensory representations, by sacrificing response speed.

© 2019 Elsevier Inc. All rights reserved.

## 1. Introduction

Throughout life, we are continually exposed to a barrage of sensory signals. Our ability to effectively navigate through and respond to the world requires us to merge information from multiple sensory modalities into a coherent percept. We may, for example, more easily locate a predator in thick foliage by combining the sight of its movement with the sound of footsteps.

Accumulating evidence suggests that aging affects how observers integrate sensory signals into perceptual decisions. In speeded target detection paradigms, older adults show greater multisensory response facilitation (i.e., redundant target effect; Laurienti et al., 2006; Mahoney et al., 2011). Furthermore, older participants have been shown to integrate multisensory stimuli differently in illusory settings such as the sound-induced flash

illusion (DeLoss et al., 2013; McGovern et al., 2014; Setti et al., 2011) and the McGurk-MacDonald effect (Sekiyama et al., 2014; Setti et al., 2013). Yet, the computational mechanisms underlying these age differences in multisensory integration remain unclear.

Two key mechanisms need to be distinguished: First, aging is known to reduce the reliability of auditory and visual representations (Dobrevá et al., 2011; Lindenberger and Baltes, 1994; Otte et al., 2013; Salthouse et al., 1996). Differences in the reliability of sensory representations may in turn alter the weights that are assigned to the sensory signals during the integration process, thereby changing the final percept. Furthermore, less reliable sensory representations will also reduce observers' ability to determine whether sensory signals come from a common source and thereby influence how they arbitrate between sensory integration and segregation. In short, age-related increases in noise in the unisensory representations may alter the perceptual outcome of multisensory integration, even if the integration processes are intact.

Second, aging may genuinely impact how observers arbitrate between sensory integration and segregation depending on temporal, spatial, or higher-order statistical correspondence cues or how they weight sensory signals in the integration process. As a

\* Corresponding author at: Computational Cognitive Neuroimaging Laboratory, Computational Neuroscience and Cognitive Robotics Centre, University of Birmingham, B15 2TT Birmingham, UK. Tel.: +44 (0)1782 294678; fax: +44 (0)121 414 3971.

E-mail address: [samjones.saj@gmail.com](mailto:samjones.saj@gmail.com) (S.A. Jones).

consequence, even if unisensory processing were preserved, we would observe differences in multisensory perception.

In short, both changes in unisensory representations and multisensory integration can alter perceptual outcomes in a similar fashion. We thus need to apply models that allow us to dissociate between those 2 mechanisms.

In the laboratory, the computational principles of multisensory integration have been studied extensively in spatial ventriloquist paradigms where observers need to report their perceived sound (or visual) location when presented with synchronous, yet spatially disparate, auditory, and visual signals. For small spatial disparities, observers' perceived sound location is shifted (or biased) toward the location of the visual signal and vice versa depending on the relative auditory and visual reliabilities—a phenomenon known as the spatial ventriloquist effect. Yet, for large audiovisual spatial disparities where it is unlikely that signals come from a common source, audiovisual interactions and crossmodal biases are attenuated. Recent psychophysics and neuroimaging studies have shown that younger observers arbitrate between sensory integration and segregation in a way that is consistent with the predictions of hierarchical Bayesian Causal Inference (BCI; [Aller and Noppeney, 2019](#); [Körding et al., 2007](#); [Rohe et al., 2019](#); [Rohe and Noppeney, 2015a,b](#); [Shams and Beierholm, 2010](#); [Wozny et al., 2010](#)). BCI enables arbitration between sensory integration and segregation by explicitly modeling the 2 causal structures (i.e., common or independent causes) that could have generated the sensory signals. If signals are caused by the same source they are integrated, they are weighted in proportion to their relative sensory reliabilities; if they are caused by different sources, they are treated separately. To account for observers' uncertainty about the world's causal structure, a final estimate (e.g., an object's location) is obtained by averaging the estimates under the assumptions of common and independent sources weighted by their respective posterior probabilities, a decision strategy referred to as model averaging (for other decision functions see [Wozny et al., 2010](#)). Spatial ventriloquism, together with BCI, may thus allow us to tease apart whether aging affects only sensory reliabilities (i.e., sensory variance) or also observers' multisensory binding (as quantified by the model's causal prior), and to test whether older adults still respond in a way that is consistent with the predictions of BCI.

However, current models of BCI do not account for temporal constraints imposed by our natural world and the dynamics of observers' perceptual inference; BCI enables predictions only for an observer's response choices (e.g., spatial localization) but not for his or her response times. In our natural environment, we often need to trade off accuracy for speed: a faster, less accurate estimate of the location of a predator may prove far more useful than a highly accurate but slow one. Indeed, recent studies have shown that putatively suboptimal multisensory behavior can be considered optimal when the dynamics of perceptual decision-making, based on both response choices and times, are taken into account ([Drugowitsch et al., 2014](#)). Considering response choices and times together is particularly relevant for understanding the impact of aging on multisensory integration, as older adults have previously been shown to favor accuracy over speed to a greater degree than younger observers ([Smith and Brewer, 1995](#); [Starns and Ratcliff, 2010](#)).

Combining psychophysics and computational modeling, the present study was thus designed to investigate how aging impacts the computational parameters governing multisensory decision-making in both unsped and speeded contexts ([Körding et al., 2007](#); [Rohe and Noppeney, 2015a,b](#); [Wozny et al., 2010](#)).

First, in an unsped spatial ventriloquist paradigm, younger and older observers located the source of a sound (which implicitly relies on causal inference; see aforementioned) or judged whether the auditory and visual signal originated from the same source

(which explicitly requires the observer to infer the causal structure underlying the audiovisual signals). We assessed how aging affects observers' auditory and visual reliabilities (i.e., sensory noise), spatial prior (i.e., spatial expectations), and causal prior (i.e., multisensory binding tendency), as key parameters of the BCI model.

Second, in a speeded spatial ventriloquist paradigm, observers were presented with spatially congruent or incongruent audiovisual signals and rapidly discriminated whether the auditory (or visual) stimulus was presented in their left or right hemifield. We used a modified version of the Bayesian compatibility bias model ([Noppeney et al., 2010](#); [Yu et al., 2009](#)) to characterize how observers accumulate evidence concurrently about signal location and audiovisual spatial congruency (i.e., causal structure) and to make predictions jointly for response choices and times. The age groups were compared in terms of auditory and visual reliabilities, prior binding tendency, and final response threshold.

If older observers differ from younger observers only in sensory reliabilities in unsped and speeded contexts, age-related changes in perceptual outcomes are a consequence of their noisier sensory representations. However, if older observers' behavior is inconsistent with principles of BCI or explained by increases or decreases in their multisensory binding tendencies (as quantified by the causal prior), then aging genuinely impacts multisensory interactions.

## 2. Methods

### 2.1. Participants

Twenty-three younger adults (11 male, mean age = 19.5,  $SD = 1.6$ , range = 18–26 years) and 23 older adults (7 male, mean age = 72,  $SD = 5.2$ , range = 63–80 years) were included in the study. One older adult was excluded before testing was completed as she was unable to perform unisensory auditory localization (approximately the same response was given to all auditory stimuli, regardless of source location). The younger adults were undergraduate psychology students at the University of Birmingham and were compensated in cash or course credits for their time. Older adults were recruited to the study from a database of local participants maintained by the University of Birmingham's School of Psychology and were compensated in cash. These community-living older adults had a diverse range of backgrounds; 39% reported education at degree level or above. All participants reported normal hearing and normal or corrected-to-normal vision and were screened for basic auditory and visual localization ability using a forced left/right discrimination task (see [Supplementary S1](#)). Participants gave informed consent prior to the commencement of testing. The research was approved by the University of Birmingham Ethical Review Committee.

### 2.2. Experimental setup

Participants were seated at a chin rest 130 cm from a sound-transparent projector screen. Behind the screen, at the vertical center, a shelf held an array of 9 studio monitors (Fostex PM04n) spaced horizontally by 7° of visual angle, including a speaker in the middle of the screen. Auditory stimuli were presented via these speakers at approximately 75 dB SPL. The locations of the speakers were not known to participants. Images were displayed using a BENQ MP782ST multimedia projector at a total resolution of 1280 × 800. All stimuli were presented using the Psychophysics Toolbox 3 ([Kleiner et al., 2007](#)) in MATLAB R2010b running on a Windows 7 PC.

Responses were made using a two-button response pad or optical mouse, and in all cases, this was effectively self-speeded; the next trial would not begin until a valid response was made. However, for the speeded ventriloquist task, it was emphasized to

participants that they should respond as quickly as possible while maintaining accuracy. See Fig. 1A for an outline of the setup.

### 2.3. Stimuli

Visual stimuli consisted of a 50 ms flash of 15 white ( $88 \text{ cd/m}^2$ ) dots, each  $0.44^\circ$  of visual angle in diameter, against a dark gray ( $4 \text{ cd/m}^2$ ) background. Dot locations were sampled uniquely for each trial from a bivariate Gaussian distribution, with a constant vertical standard deviation of  $5.4^\circ$ . The horizontal standard deviation of this dot cloud was varied to manipulate the reliability of spatial information, with a wider cloud (expressed in degrees of visual angle) resulting in less reliable stimuli (Rohe and Noppeney, 2015). We define the specific horizontal standard deviations used for each paradigm below.

The auditory stimulus was a burst of white noise (duration: 50 ms) played from 1 speaker in the array in synchrony with the visual stimulus. Sounds were generated individually for each trial and ramped on/off over 5 ms. Across all tasks, participants fixated a central cross ( $0.22^\circ$  radius) that was constantly presented throughout the entire experiment.

### 2.4. Unspeeded audiovisual spatial ventriloquist paradigm

#### 2.4.1. Design and procedure

In a spatial ventriloquist paradigm, observers were presented with synchronous auditory and visual stimuli at variable audiovisual spatial disparities and performed implicit or explicit causal

inference tasks in separate blocks. First, in an auditory selective attention task, observers reported their perceived sound location. As highlighted in the introduction, spatial localization implicitly relies on solving the causal inference problem. Second, they explicitly inferred and reported the causal structure (i.e., common vs. independent sources) that could have generated the audiovisual signals in common source judgments.

Irrespective of task context, on each trial, auditory and visual stimuli were independently sampled from 5 possible locations ( $-14^\circ, -7^\circ, 0, 7, \text{ or } 14^\circ$ ), and could therefore be spatially congruent or incongruent with varying degrees of disparity ( $0^\circ, 7^\circ, 14^\circ, 21^\circ, \text{ or } 28^\circ$ ). Visual stimuli had 3 levels of reliability (horizontal  $SD$  of  $2^\circ, 6^\circ, \text{ or } 16^\circ$ ) (n.b. a fourth level of visual reliability was excluded from the analysis because the dots were erroneously sampled). The paradigm thus conformed to a 5 (A locations)  $\times$  5 (V locations)  $\times$  3 (V reliabilities) factorial design.

In the sound localization, task participants reported the perceived sound location as accurately as possible, after a 500 ms poststimulus delay, by moving a mouse-controlled cursor (white, subtending  $9^\circ$  in height and  $0.5^\circ$  wide) whose movement was constrained to the horizontal plane. The next trial was started one second after observers had indicated their perceived auditory location by clicking the mouse button. Trials were presented randomly in 200-trial blocks. In total, participants completed 600 trials (8 [repetitions]  $\times$  5 [A locations]  $\times$  5 [V locations]  $\times$  3 [V reliabilities]) of this task.

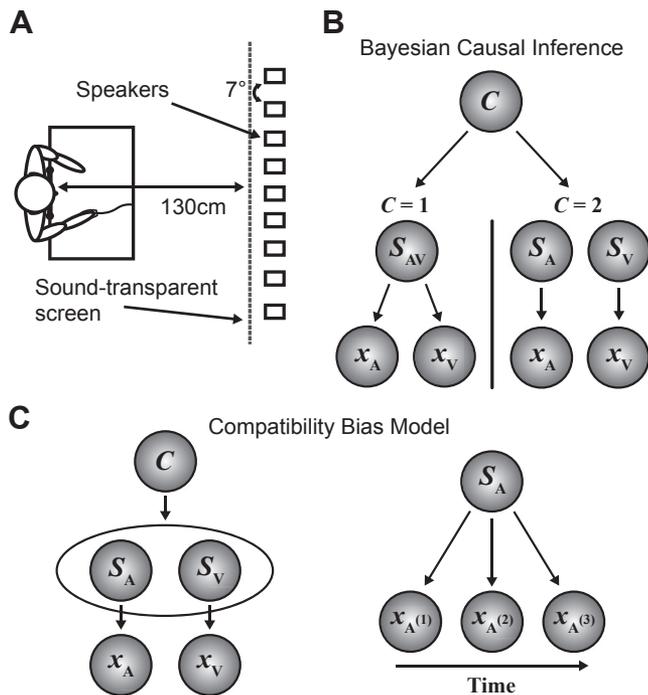
In the common-source judgment, task participants reported whether they perceived the auditory and visual signals to have originated from the same location. 500 ms after the presentation of the flash and beep, the words “same” and “different” appeared, respectively, above and below the fixation cross. Participants indicated with a button press whether the sound and flash were generated by a common source. Participants again completed 600 trials (8 [repetitions]  $\times$  5 [A locations]  $\times$  5 [V locations]  $\times$  3 [V reliabilities]) of this task, delivered in 3 blocks of 200 trials.

Unisensory auditory or visual localization blocks were also included to improve estimation of sensory reliabilities. In unisensory auditory blocks, observers were presented with sounds randomly at 1 of the 5 locations and indicated their perceived sound location with the mouse cursor, as aforementioned. 80 trials of this task (16 per location) were completed in 1 block. In unisensory visual blocks, stimuli from the 3 reliability levels indicated previously (horizontal  $SD$  of  $2^\circ, 6^\circ$  or  $16^\circ$ ) were presented randomly in 1 of the 5 locations and participants instructed to locate the center of the dot cloud with the mouse cursor. 120 trials of this task (8 per location, per reliability level) were completed in 1 block.

#### 2.4.2. Bayesian Causal Inference model

We use BCI (Aller and Noppeney, 2019; Körding et al., 2007; Rohe et al., 2019; Rohe and Noppeney, 2015a,b; Shams and Beierholm, 2010; Wozny et al., 2010) to investigate how younger and older observers arbitrate between sensory integration and segregation. In the following, we briefly describe the BCI model; for further details, see the study of Körding et al. (2007).

The BCI generative model assumes that common ( $C = 1$ ) or independent ( $C = 2$ ) sources are determined by sampling from a binomial distribution with the causal prior  $P(C = 1) = p_{\text{common}}$ . For a common source, the “true” location  $S_{AV}$  is drawn from the spatial prior distribution  $N(\mu_p, \sigma_p)$ . For 2 independent causes, the “true” auditory ( $S_A$ ) and visual ( $S_V$ ) locations are drawn independently from this spatial prior distribution. For the spatial prior distribution, we assumed a central bias (i.e.,  $\mu_p = 0$ ). We introduced sensory noise by drawing  $x_A$  and  $x_V$  independently from normal distributions centered on the true auditory (respectively visual) locations with parameters  $\sigma_A$  (respectively  $\sigma_V$  for each visual reliability level).



**Fig. 1.** Experimental setup and generative models. (A) Participants were presented with visual stimuli on a sound-transparent projector screen. Sounds were produced by individual speakers concealed behind this screen, which were separated by  $7^\circ$  of visual angle. Responses were given via a mouse or a two-button response pad. (B) Bayesian Causal Inference (BCI) model, based on the study by Körding et al. (2007). Auditory ( $x_A$ ) and visual ( $x_V$ ) signals may be generated by one common ( $C = 1$ ) audiovisual source ( $S_{AV}$ ), or by separate ( $C = 2$ ) auditory ( $S_A$ ) and visual ( $S_V$ ) sources. (C) Compatibility bias model, adapted from the study by Yu et al. (2009). Left: auditory ( $S_A$ ) and visual ( $S_V$ ) sources can either be congruent ( $C = 1$ , i.e., in same hemifield) or incongruent ( $C = 2$ , i.e., in opposite hemifields). Right: across time, the auditory source generates a series of auditory inputs, and the visual source (not shown) a series of visual inputs, in an independent and identical fashion.

Thus, the generative model included the following free parameters: the causal prior  $p_{common}$ , the spatial prior standard deviation  $\sigma_B$ , the auditory standard deviation  $\sigma_A$ , and visual standard deviations corresponding to the 3 visual reliability levels  $\sigma_{V1}$ ,  $\sigma_{V2}$ , and  $\sigma_{V3}$ .

During perceptual inference, the observer is assumed to invert this generative model. The probability of the underlying causal structure can be inferred by combining the causal prior with the sensory evidence according to Bayes' rule:

$$p(C = 1|x_A, x_V) = \frac{p(x_A, x_V|C = 1)p_{common}}{p(x_A, x_V)} \quad (1)$$

We assumed that subjects report “common source” (i.e., explicit causal inference) when the posterior probability of a common source is greater than the threshold of 0.5:

$$\hat{C} = \begin{cases} 1 & \text{if } p(C = 1|x_A, x_V) > 0.5 \\ 2 & \text{if } p(C = 1|x_A, x_V) \leq 0.5 \end{cases} \quad (2)$$

In the case of a common source ( $C = 1$ ; Fig. 1B left), the maximum a posteriori probability estimate of the auditory location is a reliability-weighted average of the auditory and visual estimates and the prior.

$$\hat{S}_{A,C=1} = \frac{\frac{x_A}{\sigma_A^2} + \frac{x_V}{\sigma_V^2} + \frac{\mu_p}{\sigma_p^2}}{\frac{1}{\sigma_A^2} + \frac{1}{\sigma_V^2} + \frac{1}{\sigma_p^2}} \quad (3)$$

In the case of a separate-source inference ( $C = 2$ ; Fig. 1B right), the estimate of the auditory signal location is independent from the visual spatial signal.

$$\hat{S}_{A,C=2} = \frac{\frac{x_A}{\sigma_A^2} + \frac{\mu_p}{\sigma_p^2}}{\frac{1}{\sigma_A^2} + \frac{1}{\sigma_p^2}} \quad (4)$$

Given the decisional strategy of model averaging (for other decisional strategies, see the study by Wozny et al., 2010), the observer will compute a final auditory localization estimate by averaging the spatial estimates under common and independent source assumptions, weighted in proportion to their posterior probabilities (i.e., implicit causal inference).

$$\hat{S}_A = p(C = 1|x_A, x_V)\hat{S}_{A,C=1} + (1 - p(C = 1|x_A, x_V))\hat{S}_{A,C=2} \quad (5)$$

The predicted distributions of the auditory spatial estimates,  $p(\hat{S}_A|S_A, S_V)$ , and the common source estimates,  $p(\hat{C}|S_A, S_V)$ , were obtained by marginalizing over the internal variables  $x_A$  and  $x_V$ . For the unisensory auditory and visual localization tasks, we used the predicted distributions  $p(\hat{S}_{A,C=2}|S_A)$  for auditory blocks and  $p(\hat{S}_{V,C=2}|S_V)$ , respectively.

These distributions were generated by simulating  $x_A$  and  $x_V$  10,000 times for each of the conditions and inferring  $\hat{S}_A$ ,  $\hat{S}_{A,C=2}$ ,  $\hat{S}_{V,C=2}$ , and  $\hat{C}$  from the equations aforementioned. Based on these predicted distributions (given an additional noise kernel with a fixed  $\sigma_{motor} = 1$ ), we computed the log-likelihood of participants' auditory localization and common-source judgment responses.

We fitted the BCI model jointly to observers' localization responses in the audiovisual and the unisensory visual and auditory stimulation conditions. We modeled the sensory noise and spatial prior parameters separately for unisensory and bisensory trials, as this was found to fit the data best overall (see Supplementary S5 for a formal comparison with models that did not separate parameters based on unisensory or audiovisual context). Therefore, a total of 11 free parameters was fitted for each participant: the causal prior  $p_{common}$ , the spatial prior standard deviations  $\sigma_{P uni}$  and  $\sigma_{P bi}$ , the

auditory standard deviations  $\sigma_{A uni}$  and  $\sigma_{A bi}$ , and visual standard deviations corresponding to the 3 visual reliability levels  $\sigma_{V1 uni}$ ,  $\sigma_{V2 uni}$ ,  $\sigma_{V3 uni}$ ,  $\sigma_{V1 bi}$ ,  $\sigma_{V2 bi}$ ,  $\sigma_{V3 bi}$  (indices *uni* and *bi* correspond to unisensory and bisensory trials, respectively). Assuming independence of conditions and responses, we summed the log-likelihoods across conditions and across localization and common-source judgment responses to obtain a single log-likelihood for each subject. To obtain maximum likelihood estimates for each subject's model parameters, we used a Bayesian adaptive search algorithm (Acerbi and Ma, 2017) with the parameters for initialization determined by a prior grid search.

The parameters (causal prior, spatial prior[s], and sensory variances) obtained from the winning model were compared between age groups using separate nonparametric Mann-Whitney *U* tests. We also calculated Bayes factors using the Bayesian Mann-Whitney test as implemented in JASP (JASP Team, 2019; van Doorn et al., 2017) using the default Cauchy prior (scale = 0.707).

## 2.5. Speeded ventriloquist paradigm

### 2.5.1. Design and procedure

To assess participants' audiovisual integration of spatial cues under speeded conditions, taking into account both final responses and reaction times, we used a simpler 2 (auditory location: left vs. right)  $\times$  2 (visual location: left vs. right)  $\times$  2 (relevant and reported sensory modality: auditory vs. visual) ventriloquist paradigm. On each trial, a visual stimulus with horizontal  $SD = 5.4^\circ$  was displayed simultaneously with a burst of white noise. The center of the visual cloud and the white noise were presented at  $14^\circ$  either left or right of a central fixation cross. These audiovisual stimuli were spatially congruent on half of the trials and incongruent on the other half. In an auditory or visual selective attention paradigm, participants indicated either the location of the sound (respond-auditory task) or the cloud (respond-visual task) as quickly and accurately as possible via a two-choice key press, while ignoring the other modality. The task was self-speeded in this way (i.e., no response deadline) as any imposed incentives or timing criteria may have affected the groups differently; we rely on the compatibility bias model (Yu et al., 2009; described in the following) to separate age differences in motor speed and speed/accuracy trade-off from potential differences in sensory reliability/evidence accumulation. The tasks were performed in 2 blocks of 160 trials. The order of these tasks was counterbalanced between participants. In total the experiment included 320 trials: 40 (repetitions)  $\times$  2 (visual location)  $\times$  2 (auditory location)  $\times$  2 (reported sensory modality).

### 2.5.2. Compatibility bias model

To assess age differences in responses to multisensory stimuli under temporal constraints, we analyzed the respond-auditory data by adapting the “compatibility bias” model to an audiovisual context (Noppeney et al., 2010; Yu et al., 2009). This models the within-trial dynamics of audiovisual evidence accumulation, leading to predictions for both response choice and response times.

See the study by Yu et al. (2009) for full details about the compatibility bias model. Briefly, this generative model assumes that congruent ( $C = 1$ ) or incongruent ( $C = 2$ ) sources are determined by sampling from a binomial distribution with the compatibility or congruency prior  $P(C = 1) = p_{congruency}$ . The visual  $S_V$  and auditory  $S_A$  sources can either be left ( $-1$ ) or right ( $+1$ ). For a congruent trial, the auditory and visual locations are identical, that is,  $S_A = S_V$  ( $S_A$  and  $S_V$  are either both left or both right). For an incongruent trial, the auditory and visual locations are in opposite hemifields, that is,  $S_A = -S_V$  (2 possibilities:  $S_A = -1$  and  $S_V = 1$ , or  $S_A = 1$  and  $S_V = -1$ ). Hence, we obtain a total of 4 possible stimulus

combinations. We then sample noisy sensory inputs successively for each time point within a trial by drawing  $x_t = [x_A(t) \ x_V(t)]$  independently from normal distributions centered on  $S_A$  (or  $S_V$ ) with parameters  $\sigma_A$  (or  $\sigma_V$ , respectively). This thereby models that the brain receives progressively more information about the location of the auditory and visual sources and thus, indirectly, about whether or not they are congruent (n.b. although in our experiment auditory and visual inputs are brief, we model evidence accumulation via feedback loops as a series of sensory inputs). Based on a stream of audiovisual inputs  $X_t = [x_1, x_2, x_3 \dots x_t]$ , the observer is then assumed to compute the posterior probability over congruency  $C$  and auditory (or visual) source location iteratively according to Bayes' rule (initialized with the prior  $P(C) = \beta$ ):

$$P(S_A, C | \mathbf{X}_t) = \frac{p(\mathbf{x}_t | S_A, C) P(S_A, C | \mathbf{X}_{t-1})}{\sum_{C, S'_A} p(\mathbf{x}_t | S'_A, C') P(S'_A, C' | \mathbf{X}_{t-1})} \quad (6)$$

A left/right decision is then made when the evolving trajectory of the marginal

$$P(S_A = 1 | \mathbf{X}_t) = P(S_A = 1, C = 1 | \mathbf{X}_t) + P(S_A = 1, C = 2 | \mathbf{X}_t) \quad (7)$$

reaches a threshold  $q$ .

Thus, incongruent visual information should be most influential on perceived auditory location at the onset of the trial, when the initial compatibility prior dominates, but this influence decreases as information about the location of each stimulus is accumulated. The process is terminated when sufficient evidence is accumulated about the location of the auditory stimulus for a decisional threshold to be reached, after which a left/right spatial response is made. To accommodate that older adults have slower motor speed than younger adults (as confirmed by a separate finger tapping task reported in [Supplementary S2](#)), we included an additional nondecision time parameter  $t_{nd}$  to account for motor delays.

The model therefore has 5 free parameters in total: the compatibility prior (i.e., prior probability of audiovisual signals coming from a common cause)  $\beta$ ; the standard deviations of the auditory and visual signals,  $\sigma_A$  and  $\sigma_V$ , respectively; the response threshold  $q$ ; and a nondecision time parameter  $t_{nd}$  that allows for a variable motor delay between the threshold being reached and a response being given.

As in the BCI model, we obtained the predicted distributions of the auditory spatial estimates,  $P(\hat{S}_A | S_A, S_V)$ , and response times,  $P(\hat{RT}_A | S_A, S_V)$ , by marginalizing over the internal variables  $x_A$  and  $x_V$ . These distributions were generated by simulating  $x_A$  and  $x_V$  for 300 time steps (of 10 ms length) 10,000 times for each of the conditions. For each simulated trial with a series of 300  $x_A$  and  $x_V$ , we then computed the response time and choice when  $P(S_A = -1 | \mathbf{X}_t)$  first crossed the decisional threshold  $q$  using Equations 5 and 6 aforementioned. Based on these predicted response choice and response time distributions, we computed the log-likelihood of participants' auditory (or visual) localization responses and the response times (after adding the nondecision time  $t_{nd}$ ). Assuming independence of conditions as well as independence of the log-likelihoods for response times and choices, we summed the log-likelihoods across conditions and across response times and choices for a particular subject. To obtain maximum likelihood estimates for the model parameters for each subject ( $\beta, \sigma_A, \sigma_V, q, t_{nd}$ ), we used a Bayesian Adaptive Search optimization algorithm ([Acerbi and Ma, 2017](#)) with parameters initialized based on a grid search.

To investigate whether any of the parameters of these 2 Bayesian models were significantly different between older and younger adults, the fitted parameters were entered into separate nonparametric Mann-Whitney  $U$  tests. We also calculated Bayes factors

using the Bayesian Mann-Whitney test as implemented in JASP ([JASP Team, 2019](#); [van Doorn et al., 2017](#)) using the default Cauchy prior (scale = 0.707).

### 3. Results

#### 3.1. Unisensory screening tests and the Montreal Cognitive Assessment

Before the main unspeeded and speeded ventriloquist experiments, all observers were screened for basic auditory and visual localization ability with a binary left/right forced-choice spatial classification task. Individuals were characterized in terms of the slope and threshold of psychometric functions fitted to these responses. Older and younger adults were closely matched: no significant age differences in threshold or bias were observed for auditory or visual spatial processing, suggesting that sensory spatial reliability was approximately similar between age groups. No participants were excluded as a result of poor performance on this task. See [Supplementary S1](#) for full details.

Older participants were also screened using the Montreal Cognitive Assessment with a cutoff score of 23 ([Coen et al., 2011](#); [Luis et al., 2009](#); [Roalf et al., 2013](#)); none of our older participants scored below 25.

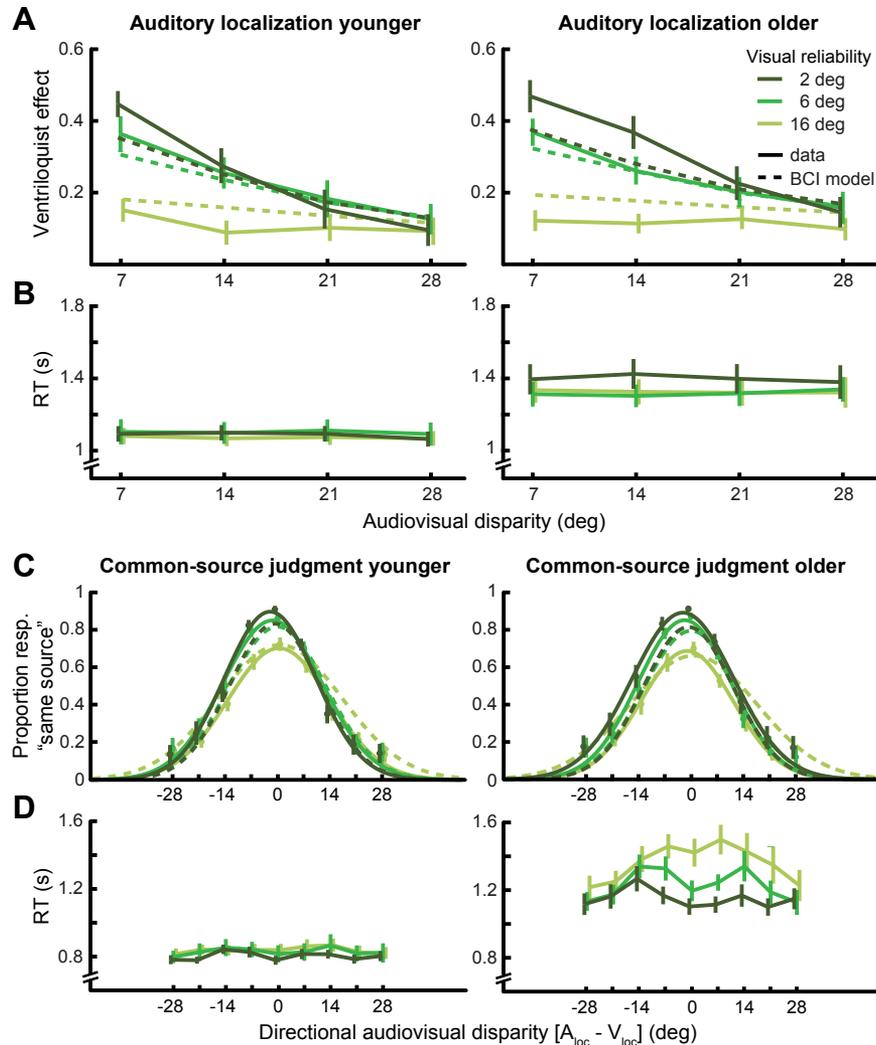
#### 3.2. Unspeeded ventriloquist paradigm: localization and common source judgment

##### 3.2.1. Descriptive and GLM-based analysis

An unspeeded spatial ventriloquist paradigm was used to compare younger and older adults' responses to audiovisual spatial stimuli in the absence of temporal constraints. [Fig. 2](#) shows participants' auditory localization (presented in terms of the magnitude of ventriloquist effect,  $VE = [A_{resp} - A_{loc}] / [V_{loc} - A_{loc}]$ ) and common-source judgment responses (characterized as the probability of responding "same-source") as a function of visual reliability level and audiovisual disparity. As predicted by BCI, the ventriloquist effect was strongest when visual reliability was high and the audiovisual disparity small. The age groups performed remarkably similarly on both measures, with standard GLM analyses revealing no significant effects of age on final response choices. However, older observers were significantly slower than younger adults when localizing sounds in the spatial ventriloquist paradigm. Furthermore, we observed significant age effects on the common-source judgment reaction times ([Fig. 2D](#)), including significant interactions between age, visual reliability, and audiovisual disparity. See [Supplementary S3](#) for full GLM analyses of these results.

##### 3.2.2. Bayesian modeling

[Table 1](#) summarizes the fitted parameters (within-group mean and  $SD$ ) of the BCI model for younger and older participants. [Table 1](#) also reports the results of the nonparametric tests comparing the parameters between the older and younger groups together with the Bayes factors associated with each statistical comparison. We observed small but significant group differences in auditory and visual variance parameters that were estimated based on unisensory localization tasks alone, suggesting that older adults were slightly less precise when locating both auditory and particularly unreliable visual stimuli. These group differences were not significant when the sensory variance parameters were estimated based on responses to audiovisual stimuli, probably because these parameters were less precisely estimated in this case: in the audiovisual context, the visual variance parameter is only estimated indirectly from auditory responses, and the auditory variance is



**Fig. 2.** Behavioral responses, reaction times, and BCI model predictions for younger and older adults. (A) Relative ventriloquist effect ( $VE = [A_{resp} - A_{loc}] / [V_{loc} - A_{loc}]$ ) for auditory localization, shown as a function of audiovisual disparity (x-axis, pooled over direction) and visual reliability (color coded). Behavioral data (mean across subjects, solid lines) and the predictions of the Bayesian Causal Inference model (dashed lines) are shown. (B) Reaction times in auditory localization task. (C) Proportion reported "same source" in common-source judgment task, as a function of audiovisual disparity and visual reliability. The panels show the Gaussians fitted to the behavioral response (mean across subjects, solid lines) and the predictions of the Bayesian Causal Inference model (dashed lines). (D) Reaction times (pooled over response; mean across subjects) in common-source judgment task. Error bars show  $\pm 1$  SEM. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

always estimated in the presence of interfering visual signals (and so may be influenced by factors other than peripheral sensory noise).

Crucially, however, no significant group differences were observed for the  $P_{common}$  or  $\sigma_P$  parameters. This suggests that the 2 age groups had similar central spatial priors and causal priors, suggesting that older and younger adults showed similar tendencies to bind audiovisual signals (in an unspeeded context) consistent with BCI.

To verify that these results were not confounded by possible age differences in motor noise (i.e., noisier mouse localization responses), we also fitted a version of the model that allowed the parameter  $\sigma_{motor}$  to vary freely ( $\sigma_{motor}$  was fixed at  $1^\circ$  for all participants in the main analysis). The pattern of results remained similar, although the group difference in  $\sigma_{A_{uni}}$  became marginally nonsignificant ( $p = 0.052$ ). Furthermore, there were no significant group differences in the  $\sigma_{motor}$  parameter ( $p > 0.05$ ,  $BF_{01} = 3.15$ ). See [Supplementary S6](#) for details.

In summary, age did not influence observer's implicit (auditory localization) or explicit (common-source judgment) causal inference in terms of response choices. Our Bayesian modeling analysis revealed that older adults had slightly noisier auditory and visual

representations when estimated separately for the unisensory conditions. Importantly, though, the comparable causal prior (and central prior), and similar mean response choices, indicate that older observers combined audiovisual spatial signals according to the same computational principles as younger adults.

Yet, aging was associated with complex changes in reaction times to multisensory stimuli. The profile of these age differences suggests that older adults took more time to respond when the causal structure of the stimuli was more ambiguous and the task therefore more challenging, such as when the visual stimulus was less reliable and/or the audiovisual disparity of intermediate size. These response time findings were followed up in a speeded ventriloquist task, where observers were explicitly instructed to respond as quickly as possible while maintaining accuracy.

### 3.3. Speeded ventriloquist paradigm

#### 3.3.1. Descriptive and GLM-based analysis

A simplified, speeded ventriloquist paradigm was used to assess younger and older adults' responses to audiovisual spatial

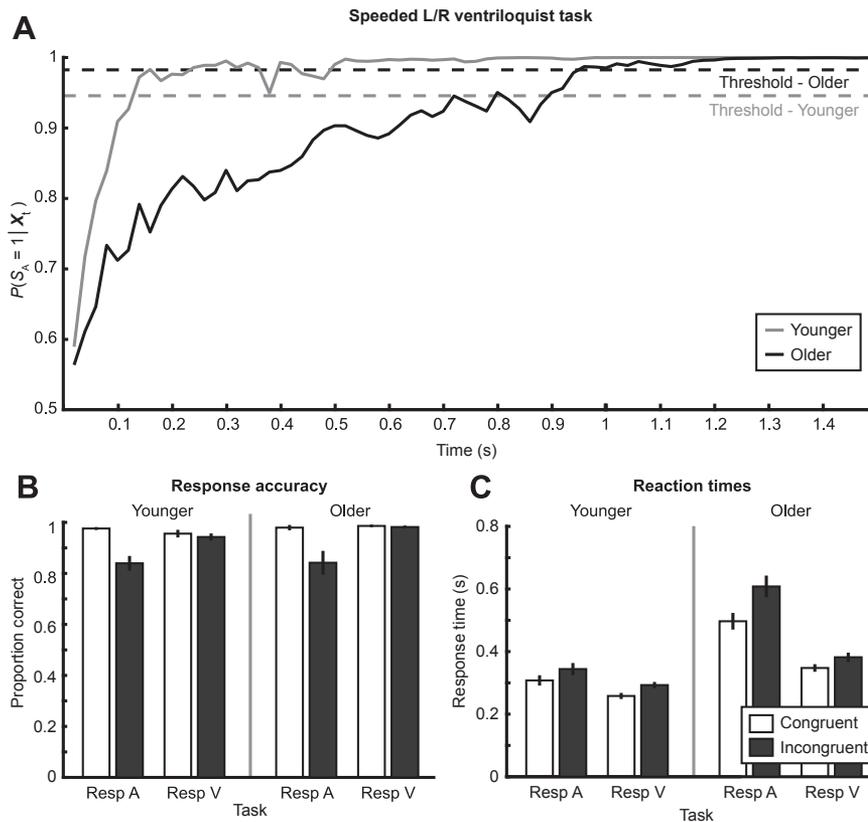
**Table 1**  
Bayesian Causal Inference parameters (across-participants mean, SD) for younger ( $n = 23$ ) and older ( $n = 22$ ) participants

Parameter	Younger		Older		Mann-Whitney $U$			Bayes factors	
	Mean	SD	Mean	SD	$W$	$p$	$\eta^2$	$BF_{10}$	$BF_{01}$
<b>Unisensory</b>									
$\sigma_P$ uni	37.20	35.69	24.79	28.63	299	0.305	0.02	0.46	2.16
$\sigma_A$ uni	5.27	1.96	6.79	2.76	155	<b>0.026</b>	0.11	2.19	0.46
$\sigma_{V1}$ uni	1.76	1.22	2.10	1.06	174	0.075	0.07	1.10	0.91
$\sigma_{V2}$ uni	2.32	0.76	2.89	1.52	198	0.218	0.04	0.54	1.84
$\sigma_{V3}$ uni	4.22	1.00	5.38	1.67	132	<b>0.005</b>	0.17	4.95	0.20
<b>Bisensory</b>									
$P_{common}$	0.42	0.13	0.43	0.13	245	0.866	<0.01	0.30	3.28
$\sigma_P$ bi	38.71	25.88	32.20	27.37	303	0.264	0.03	0.40	2.49
$\sigma_A$ bi	8.59	4.40	9.37	5.78	234	0.677	<0.01	0.35	2.84
$\sigma_{V1}$ bi	3.19	4.08	3.08	3.13	241	0.796	<0.01	0.30	3.31
$\sigma_{V2}$ bi	5.12	4.32	6.07	5.47	204	0.274	0.03	0.44	2.25
$\sigma_{V3}$ bi	12.79	9.72	20.61	26.13	209	0.327	0.02	0.48	2.09

Mann-Whitney  $U$  tests with Bayes factors comparing the BCI parameters between older and younger adults. The Bayesian Causal Inference model was fitted jointly to unisensory and audiovisual conditions allowing for separate parameters for the standard deviation of the spatial prior ( $\sigma_{P,uni}$ ,  $\sigma_{P,bi}$ ) and sensory noise ( $\sigma_{A,uni}$ ,  $\sigma_{A,bi}$ ,  $\sigma_{V1,uni}$ ,  $\sigma_{V1,bi}$ , ...,  $\sigma_{V3,uni}$ ,  $\sigma_{V3,bi}$ ).  $BF_{10}$  quantifies degree of support for the alternative hypothesis that the groups differ, relative to the null hypothesis;  $BF_{01}$  shows degree of support for the null hypothesis that there is no difference between groups, relative to the alternative hypothesis. Bold indicates  $p < 0.05$ .

stimuli under speed instructions. Fig. 3 summarizes response accuracy (panel B) and speed (panel C) for younger and older adults; trials are pooled over left and right to characterize them in terms of spatial (in) congruence. Standard GLM analysis of these results shows that older adults were significantly more accurate than younger adults in the respond-visual task. Older adults were also significantly slower overall and, importantly, age

interacted with congruence in the respond-auditory tasks (see Section 3.2). Mirroring the profile of the unspeeded common-source judgment responses, older adults again took disproportionately longer to respond under the most challenging conditions where they located the auditory signal in the presence of an incongruent visual distractor. See Supplementary S4 for full GLM analysis.



**Fig. 3.** Speeded left/right ventriloquist paradigm and compatibility bias model. (A) Accumulation of evidence traces for the compatibility bias model: for “respond auditory” trials the observer is thought to accumulate audiovisual evidence about whether the auditory source is left = -1 or right = 1 within a trial until a decisional threshold is reached and a response elicited. Solid lines show the posterior probability  $P(S_A = 1 | X_t)$  as a function of within-trial time with auditory and visual inputs arriving every 10 ms. Each trace represents the mean across 10 (incongruent, auditory right) simulated trials for a representative participant in each group, using these participant’s maximum likelihood parameters. Dashed lines indicate the participants’ fitted decisional thresholds. Older observers accumulate noisier evidence until a higher decisional threshold is reached. (B and C) Response accuracy and reaction times (across-participants mean  $\pm 1$  SEM) for respond-auditory and respond-visual tasks, separated by spatial congruence (i.e., pooled over left and right).

### 3.3.2. Compatibility bias model

The compatibility bias model was fitted to participants' auditory spatial responses and reaction times. This allowed us to characterize how younger and older observers accumulate audiovisual evidence about spatial location and audiovisual congruency until a decisional threshold is reached and a response given. Fitted parameters were compared using separate Mann-Whitney  $U$  tests and the Bayesian version of the Mann-Whitney test (JASP Team, 2019; van Doorn et al., 2017). See Table 2 for a summary of results.

Corroborating the findings of the BCI model, the age groups did not differ in their prior tendency to integrate multisensory stimuli, quantified in this case by the compatibility prior  $\beta$ . However, similar to the results from unspeeded localization, the auditory signal ( $\sigma_{\text{auditory}}$ ) was significantly noisier in older than younger adults, leading to a slower accumulation of evidence and thus (in combination with the motor slowing and higher decision threshold, see below) slower response times. This indicates that it takes older participants longer than their younger counterparts to reach any given level of evidence about the location of an auditory stimulus. The groups did not differ in the variance of the visual input  $\sigma_{\text{visual}}$ . However, the remaining 2 parameters were also significantly different between the groups. First, the nondecision time  $t_{\text{nd}}$ , which captures the time between a decision-making and the response, was significantly higher for the older age group. This is unsurprising; our older adults' impaired motor speed is confirmed by a separate finger-tapping task reported in Supplementary S2. Second, older adults also set their decision threshold  $q$  significantly higher, requiring more evidence before deciding on a response. See Fig. 3A for an illustration of the model. Taken as a whole, our Bayesian modeling analysis confirms that older adults show a similar multisensory binding tendency and combine signals to the same computational principles as younger adults. However, older adults have noisier unisensory auditory spatial representations. As a result of i. those noisier auditory spatial representations, ii. a different speed-accuracy trade off (i.e., decision threshold  $q$ ), and iii. slower motor speed (i.e., nondecision time  $t_{\text{nd}}$ ) they have slower response times.

## 4. Discussion

This study investigated the effects of aging on audiovisual integration for spatial localization under both speeded and unspeeded conditions. Our results demonstrate that aging does not fundamentally impact how observers integrate auditory and visual spatial signals into representations of space: older adults showed the same audiovisual binding tendency as the younger age group, and their behavior conformed similarly to the predictions of the Bayesian models. However, older adults showed noisier sensory, in particular auditory, representations. Moreover, they used a higher decisional threshold, trading off speed for accuracy. This suggests that older observers preserve audiovisual localization performance,

despite noisier sensory representations, by sacrificing response speed.

These results may initially seem surprising in light of accumulating research showing that aging alters multisensory integration. For example, older adults have been shown to be more susceptible to the sound-induced flash illusion (DeLoss et al., 2013; McGovern et al., 2014; Setti et al., 2011) and to respond differently to McGurk-MacDonald stimuli (Sekiyama et al., 2014; Setti et al., 2013). It is possible, however, for such effects to occur in the absence of age differences in the actual computational processes underlying multisensory perception. Any change that leads to an increase in sensory variances may make the arbitration between common and separate sources more challenging, and/or change the relative weighting of the sensory modalities in the final percept. Potentially, susceptibility to the sound-induced flash illusion is changed with age because it relies on precise representations of stimulus timing that have been shown to be impaired by aging (Chan et al., 2014; Mazelová et al., 2003). Ng and Recanzone (2018) provide a possible mechanism for this decline: a study of neural responses to simple stimuli in macaque primary auditory cortex found that aged monkeys showed firing patterns that were noisier (i.e., less temporally precise) and less selective than those seen in younger animals. Age-related differences in perception of McGurk-MacDonald stimuli may also be due in part to impaired temporal perception, as the fine temporal structure of speech signals is an important cue for comprehension (especially in the context of competing noise; Moore, 2008). In this case the effect is likely to be further compounded by reductions in speech comprehension, resulting from presbycusis that particularly affects higher sound frequencies (Pichora-Fuller and Souza, 2003). These mechanisms are notably unisensory and do not imply any change in the computational process of multisensory integration itself.

The argument that older adults' changed multisensory perception results primarily from differences in unisensory variances, and not from alterations in the computational mechanisms per se, can also explain why we did not find significant age differences in the final responses to our multisensory tasks: our unisensory results, and those of others (Dobrevá et al., 2011; Otte et al., 2013), demonstrate only limited age differences in localization ability. Based on screening tests involving binary left/right judgments, younger and older adults were similar in their ability to locate unisensory auditory and visual stimuli. The sensory variance parameters of a BCI model fitted to multisensory localization and common-source judgment responses also did not differ between age groups. However, the same parameters fitted using the more sensitive unisensory free-localization responses did reveal small but significant age differences in sensory variances, suggesting that older adults were less reliable in their localization of both auditory and (low-reliability) visual stimuli.

Existing literature is similarly ambiguous about age-related declines in (especially) auditory localization. Dobrevá et al. (2011)

**Table 2**  
Compatibility bias parameters (across-participants mean, SD) for younger ( $n = 23$ ) and older ( $n = 22$ ) participants

Parameter	Younger		Older		Mann-Whitney $U$			Bayes factors	
	Mean	SD	Mean	SD	$W$	$p$	$\eta^2$	$BF_{10}$	$BF_{01}$
$\sigma_A$	1.53	0.54	2.93	4.01	164	<b>0.044</b>	0.09	3.11	0.32
$\sigma_V$	1.85	3.50	0.87	0.97	283	0.507	0.01	0.44	2.27
$\beta$	0.75	0.12	0.78	0.13	192	0.169	0.04	0.56	1.79
$q$	0.93	0.05	0.95	0.07	141	<b>0.010</b>	0.14	2.68	0.37
$t_{\text{nd}}$	0.22	0.05	0.33	0.07	54	<b>&lt; 0.001</b>	0.45	5101.52	<0.01

Mann-Whitney  $U$  tests with Bayes factors comparing the compatibility bias parameters between older and younger adults: standard deviation of the auditory signal  $\sigma_A$ , standard deviation of the visual signal  $\sigma_V$ , compatibility prior  $\beta$ , response threshold  $q$ , and nondecision time  $t_{\text{nd}}$ .  $BF_{10}$  quantifies degree of support for the alternative hypothesis that the groups differ, relative to the null hypothesis;  $BF_{01}$  shows degree of support for the null hypothesis that there is no difference between groups, relative to the alternative hypothesis.

Bold indicates  $p < 0.05$ .

report limited but significant age differences in observers' ability to freely localize transient broadband stimuli along the azimuth, whereas *Otte et al. (2013)* found no such effects. It therefore seems that the effects of normal, healthy aging on auditory localization ability may be subtle and difficult to detect.

In terms of visual localization, we note that our older adults are likely to have had impaired accommodation responses compared with the younger age group (*Glasser and Campbell, 1998*). Depending on the corrective lenses worn (participants were instructed to wear their normal spectacles for testing), this may have led to the older group expending more effort to keep the visual stimuli in focus and/or the stimuli appearing less focused. The small but significant age differences we observed in unisensory visual localization may be, in part, a reflection of this reduced accommodation ability.

In light of these limited age differences in audiovisual localization performance, it would be interesting for future research to apply computational modeling to multisensory contexts where strong age differences have been shown previously. The sound-induced flash illusion is a strong candidate for this, as older adults are known to be significantly more susceptible (*DeLoss et al., 2013; McGovern et al., 2014; Setti et al., 2011*) and young observers' perception of the illusion has previously been successfully modeled using a BCI framework (*Shams et al., 2005*). Fitting the BCI model to younger and older observers' responses would allow us to distinguish whether age differences in perception of the sound-induced flash illusion result from changes in unisensory variances (i.e., noise) or in observers' multisensory binding itself.

Our discussion of age differences in multisensory integration has thus far addressed only final response choices, ignoring reaction times, but our natural environment does not afford us infinite time to react to multisensory stimuli. When we define and evaluate multisensory integration performance, it is therefore also important to consider the time taken to respond. In fact, GLM-based analyses of common-source judgment reaction times suggested that older adults took disproportionately longer to respond to audiovisual signals at intermediate levels of spatial disparity, where the underlying causal structure (i.e., common vs. independent sources) was less certain. Such findings imply the presence of differences in the groups' evidence accumulation and decision-making process, and/or in their speed/accuracy criteria, even in an unsped context.

We thus applied a simplified, speeded ventriloquist paradigm to directly address the question of age differences in response times to multisensory spatial stimuli. GLM analyses again showed that older adults were disproportionately slower in the most challenging condition, in this case locating a sound in the presence of an incongruent visual distractor. To characterize the computational processes underlying these differences, it is necessary to move beyond the static BCI model to a dynamical approach that can make predictions jointly about observers' spatial choices and response times. We thus applied the compatibility bias model (*Noppeney et al., 2010; Yu et al., 2009*) to participants' auditory judgment responses in this paradigm.

This model assumes that the observer accumulates auditory and visual evidence about the location of the reported stimulus, and about the causal structure of the signals, until a decisional threshold is reached and a response given. It thereby provides an important perspective on the dynamics of decision-making within a trial. Again in this case, the fundamental computations were not affected by healthy aging. Likewise, older adults' prior binding tendency was not significantly different from the younger group. However, the compatibility bias model also revealed that older adults responded more slowly than younger adults for 3 reasons. First, older adults have impaired motor speed, as indexed by the nondecision time

variable (and confirmed by a supplementary finger-tapping task; see *Supplementary S2*). Second, they use a higher response threshold, requiring a greater degree of certainty before a response is given. This is consistent with previous studies of age differences in speed/accuracy trade-off (*Smith and Brewer, 1995; Starns and Ratcliff, 2010*). Third, the compatibility bias model analysis suggests that the auditory representations are less reliable (i.e., greater auditory variance) in older participants, such that evidence accumulates more slowly (see *Fig. 3*). In other words, the initial auditory representation may be noisier and less reliable for older adults, but older observers can achieve equal performance levels (in terms of final response choices) to younger participants by accumulating this noisy evidence for longer via internal feedback loops.

It is important to note that the BCI model, and other approaches that consider only the observer's final response, may be less sensitive to these age-related changes in internal sensory noise (although the unisensory localization data do provide some evidence of small reliability differences). This illustrates how dynamical models that accommodate both reaction times and final response choices can provide critical new insights into evidence accumulation and perceptual decision-making.

In conclusion, our results demonstrate that multisensory causal inference is preserved in older adults. However, older observers only maintain this performance by accumulating noisier auditory information over a longer period. When combined with well-established changes in motor speed and speed/accuracy trade-off, this leads to significant and nonlinear age differences in reaction times to complex multisensory stimuli during spatial localization.

## Disclosure

Participants gave informed consent prior to the commencement of testing, and were compensated in cash or research credits for their time. The research was approved by the University of Birmingham Ethical Review Committee.

## Acknowledgements

This research was funded by the European Research Council, European Union (ERC-2012-StG\_2011109multisens) and the MRC-ARUK Center for Musculoskeletal Aging Research (CMAR).

All authors have reviewed the contents of the article being submitted, approved its contents, and validated the accuracy of the data.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.neurobiolaging.2019.08.017>.

## References

- Acerbi, L., Ma, W.J., 2017. Practical bayesian optimization for model fitting with bayesian adaptive direct search. In: Guyon, I., Luxburg, U.V., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., Garnett, R. (Eds.), *Advances in Neural Information Processing Systems 30*. Curran Associates, Inc., Red Hook, NY, pp. 1836–1846.
- Aller, M., Noppeney, U., 2019. To integrate or not to integrate: temporal dynamics of hierarchical Bayesian causal inference. *PLoS Biol.* 17, e3000210.
- Chan, Y.M., Pianta, M.J., McKendrick, A.M., 2014. Older age results in difficulties separating auditory and visual signals in time. *J. Vis.* 14 <https://doi.org/10.1167/14.11.13>.
- Coen, R.F., Cahill, R., Lawlor, B.A., 2011. Things to watch out for when using the Montreal cognitive assessment (MoCA). *Int. J. Geriatr. Psychiatry* 26, 107–108.
- DeLoss, D.J., Pierce, R.S., Andersen, G.J., 2013. Multisensory integration, aging, and the sound-induced flash illusion. *Psychol. Aging* 28, 802–812.
- Dobrev, M.S., O'Neill, W.E., Paige, G.D., 2011. Influence of aging on human sound localization. *J. Neurophysiol.* 105, 2471–2486.

- 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.
- Glasser, A., Campbell, M.C.W., 1998. Presbyopia and the optical changes in the human crystalline lens with age. *Vision Res.* 38, 209–229.
- JASP Team (2019). JASP (Version 0.9.1) [Computer software]. Kleiner, M., Brainard, D., Pelli, D., 2007. What's new in Psychtoolbox-3?. In: Paper presented at the 30th European Conference on Visual Perception. Arezzo, Italy.
- 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.
- Laurienti, P.J., Burdette, J.H., Maldjian, J.A., Wallace, M.T., 2006. Enhanced multisensory integration in older adults. *Neurobiol. Aging* 27, 1155–1163.
- Lindenberger, U., Baltes, P.B., 1994. Sensory functioning and intelligence in old age: a strong connection. *Psychol. Aging* 9, 339–355.
- Luis, C.A., Keegan, A.P., Mullan, M., 2009. Cross validation of the Montreal Cognitive Assessment in community dwelling older adults residing in the Southeastern US. *Int. J. Geriatr. Psychiatry* 24, 197–201.
- Mahoney, J.R., Li, P.C.C., Oh-Park, M., Verghese, J., Holtzer, R., 2011. Multisensory integration across the senses in young and old adults. *Brain Res.* 1426, 43–53.
- Mazelová, J., Popelar, J., Syka, J., 2003. Auditory function in presbycusis: peripheral vs. central changes. *Exp. Gerontol.* 38, 87–94.
- McGovern, D.P., Roudaia, E., Stapleton, J., McGinnity, T.M., Newell, F.N., 2014. The sound-induced flash illusion reveals dissociable age-related effects in multisensory integration. *Front Aging Neurosci.* 6, 250.
- Moore, B.C.J., 2008. The role of temporal fine structure processing in pitch perception, masking, and speech perception for normal-hearing and hearing-impaired people. *J. Assoc. Res. Otolaryngol.* 9, 399–406.
- Ng, C.-W., Recanzone, G.H., 2018. Age-related changes in temporal processing of rapidly-presented sound sequences in the macaque auditory cortex. *Cereb. Cortex* 28, 3775–3796.
- Noppeney, U., Ostwald, D., Werner, S., 2010. Perceptual decisions formed by accumulation of audiovisual evidence in prefrontal cortex. *J. Neurosci.* 30, 7434–7446.
- Otte, R.J., Agterberg, M.J.H., Van Wanrooij, M.M., Snik, A.F.M., Van Opstal, A.J., 2013. Age-related hearing loss and ear morphology affect vertical but not horizontal sound-localization performance. *J. Assoc. Res. Otolaryngol.* 14, 261–273.
- Pichora-Fuller, M.K., Souza, P.E., 2003. Effects of aging on auditory processing of speech. *Int. J. Audiol.* 42, 11–16.
- Roalf, D.R., Moberg, P.J., Xie, S.X., Wolk, D.A., Moelter, S.T., Arnold, S.E., 2013. Comparative accuracies of two common screening instruments for classification of Alzheimer's disease, mild cognitive impairment, and healthy aging. *Alzheimers Dement.* 9, 529–537.
- Rohe, T., Ehlis, A.-C., Noppeney, U., 2019. The neural dynamics of hierarchical Bayesian causal inference in multisensory perception. *Nat. Commun.* 10, 1907.
- Rohe, T., Noppeney, U., 2015. Sensory reliability shapes perceptual inference via two mechanisms. *J. Vis.* 15, 22.
- Rohe, T., Noppeney, U., 2015. Cortical hierarchies perform bayesian causal inference in multisensory perception. *PLoS Biol.* 13, e1002073.
- Salthouse, T.A., Hancock, H.E., Mein, E.J., Hambrick, D.Z., 1996. Interrelations of age, visual acuity, and cognitive functioning. *J. Gerontol. B Psychol. Sci. Soc. Sci.* 51B, P317–P330.
- Sekiyama, K., Soshi, T., Sakamoto, S., 2014. Enhanced audiovisual integration with aging in speech perception: a heightened McGurk effect in older adults. *Front Psychol.* 5, 323.
- Setti, A., Burke, K.E., Kenny, R.A., Newell, F.N., 2011. Is inefficient multisensory processing associated with falls in older people? *Exp. Brain Res.* 209, 375–384.
- Setti, A., Burke, K.E., Kenny, R., Newell, F.N., 2013. Susceptibility to a multisensory speech illusion in older persons is driven by perceptual processes. *Front Psychol.* 4. <https://doi.org/10.3389/fpsyg.2013.00575>.
- Shams, L., Beierholm, U.R., 2010. Causal inference in perception. *Trends Cogn. Sci.* 14, 425–432.
- Shams, L., Ma, W.J., Beierholm, U., 2005. Sound-induced flash illusion as an optimal percept. *Neuroreport* 16, 1923–1927.
- Smith, G.A., Brewer, N., 1995. Slowness and age: speed-accuracy mechanisms. *Psychol. Aging* 10, 238–247.
- Starns, J.J., Ratcliff, R., 2010. The effects of aging on the speed-accuracy compromise: boundary optimality in the diffusion model. *Psychol. Aging* 25, 377–390.
- van Doorn, J., Ly, A., Marsman, M., Wagenmakers, E.-J., 2017. Bayesian Latent-Normal Inference for the Rank Sum Test, the Signed Rank Test, and Spearman's Rho [arXiv:1712.06941 \[stat\]](https://arxiv.org/abs/1712.06941). <https://arxiv.org/abs/1712.06941>. (Accessed 29 March 2019).
- Wozny, D.R., Beierholm, U.R., Shams, L., 2010. Probability matching as a computational strategy used in perception. *PLoS Comput. Biol.* 6, e1000871.
- Yu, A.J., Dayan, P., Cohen, J.D., 2009. Dynamics of attentional selection under conflict: toward a rational bayesian account. *J. Exp. Psychol.* 35, 700–717.