

The Brain's Tendency to Bind Audiovisual Signals Is Stable but Not General

Psychological Science
2016, Vol. 27(4) 583–591
© The Author(s) 2016
Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0956797616628860
pss.sagepub.com


Brian Odegaard¹ and Ladan Shams^{1,2,3}

¹Department of Psychology, ²Department of Bioengineering, and ³Neuroscience Interdepartmental Program, University of California, Los Angeles

Abstract

Previous studies have shown a surprising amount of between-subjects variability in the strength of interactions between sensory modalities. For the same set of stimuli, some subjects exhibit strong interactions, whereas others exhibit weak interactions. To date, little is known about what underlies this variability. Sensory integration in the brain could be governed by a global mechanism or by task-specific mechanisms that could be either stable or variable across time. We used a rigorous quantitative tool (Bayesian causal inference) to investigate whether integration (i.e., binding) tendencies generalize across tasks and are stable across time. We report for the first time that individuals' binding tendencies are stable across time but are task-specific. These results provide evidence against the hypothesis that sensory integration is governed by a single, global parameter in the brain.

Keywords

binding tendency, multisensory integration, individual differences, spatial localization, temporal numerosity judgments, Bayesian causal inference, Bayesian priors

Received 9/17/15; Revision accepted 1/4/16

As we experience the surrounding world, our brains effortlessly process sights and sounds, inferring which visual sources produced which auditory signals (Parise, Spence, & Ernst, 2012; Shams & Beierholm, 2010; Shams, Ma, & Beierholm, 2005; Stein, 2012). At first, it might seem reasonable to assume that one's subjective experience of sights and sounds is similar to that of other individuals, but many studies of multisensory processing indicate otherwise: Even for the same set of stimuli, some individuals frequently integrate audiovisual stimuli, whereas other individuals may not integrate at all. This type of variability has been shown in several domains, including speech perception (Mallick, Magnotti, & Beauchamp, 2015), temporal-numerosity perception (Shams, Kamitani, & Shimojo, 2000; Stevenson, Zemtsov, & Wallace, 2012), and spatial perception (Hairston et al., 2003; Wozny, Beierholm, & Shams, 2010), which raises important questions: Why are different brains so variable in their interpretations of our sensory world? And what characterizes the elusive, idiosyncratic mechanism that binds the senses together inside our heads?

If multisensory integration is dependent on a hard-wired, global mechanism (e.g., neuroanatomical connectivity), it

could be stable across time and generalize across tasks involving the same modalities (e.g., across all audiovisual tasks) in a given individual. Indeed, one previous study has shown strong within-subjects correlations for illusory multisensory percepts using an audiovisual speech task and an audiovisual temporal-numerosity-judgment task (Tremblay, Champoux, Bacon, & Theoret, 2007), which would suggest at least a partially shared substrate of multisensory integration across tasks. Alternatively, sensory integration could be based on local mechanisms, which could differ across tasks involving the same modalities. These local mechanisms might be stable over time or might be unstable and volatile because of factors such as learning, arousal, mood, and cognitive variables.

In principle, any variability across subjects in the strength of audiovisual interactions could be driven by two distinct factors: differences in the relative reliabilities

Corresponding Author:

Brian Odegaard, Department of Psychology, University of California, Los Angeles, 1285 Franz Hall, Box 951563, Los Angeles, CA 90095-1563

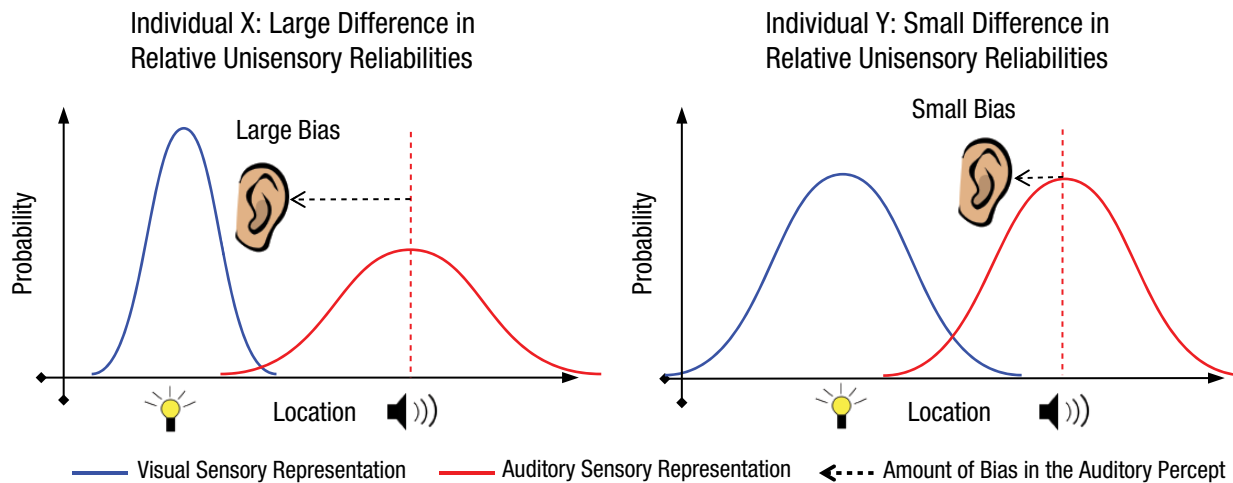
E-mail: odegaard.brian@gmail.com

of unisensory representations (e.g., in the visual and auditory modalities, Fig. 1a) or differences in the *tendency to integrate* or *bind* stimuli (henceforth, *binding tendency*) between the two given modalities (Fig. 1b). Recent investigations have revealed that although cross-modal interactions vary across individuals in any given task, such interactions can correlate strongly across audiovisual tasks within individuals (Stevenson et al., 2012; Stevenson & Wallace, 2013); however, no previous study has used analytic methods capable of teasing apart these two possible sources of variability.

The Bayesian causal inference (BCI) model (Körding et al., 2007; Wozny et al., 2010) allows quantitative, simultaneous estimation of both of these factors: The reliability of unisensory processing is computationally modeled by sensory likelihoods, and the tendency to integrate is captured by a prior for integrating stimuli, which we call the *binding tendency*. The fact that differences in unisensory processing influence integration is already well established (Alais & Burr, 2004; Ernst & Banks, 2002; Rohe & Noppeney, 2015b); on the other hand, the binding tendency, which provides a true

a

Individual Differences in Unisensory Reliabilities



b

Individual Differences in Binding Tendencies

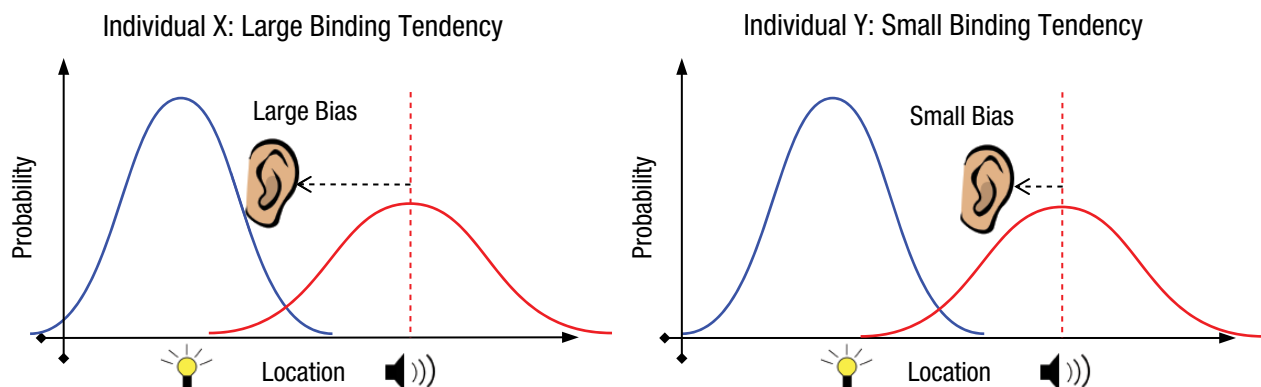


Fig. 1. Two possible mechanisms underlying between-subjects variability in cross-modal interactions. The lightbulb and the speaker represent the true locations of the visual and auditory stimuli in a spatial localization task; the neural representations (likelihood functions) of the stimuli are shown by the distributions for vision and audition. The ear icon designates the brain's final estimate of the perceived auditory location. One possible source of between-subjects variability in cross-modal interactions involves differences in relative unisensory reliabilities across subjects (a). If these differences are large (as shown by Individual X), estimates of percepts in the less reliable modality (audition in this example) will show a large bias; if these differences are small (as shown by Individual Y), the bias will be smaller. Another possible source of between-subjects variability in cross-modal interactions involves differences in the *binding tendency*, that is, the prior bias for perceiving a common cause and integrating the signals (b). As shown in the figure, Individual X's large bias could be due to a large binding tendency, and Individual Y's small bias could be due to a small binding tendency.

measure of the brain's tendency or capacity for integration, has not been systematically investigated to date. Therefore, we used the BCI model to quantitatively and rigorously estimate this measure for individual observers for two very distinct audiovisual tasks (one involving perception of space, the other involving perception of time) at two time points (1 week apart) to determine whether the binding tendency is stable across time and whether it generalizes across tasks.

Method

Subjects

Fifty-nine subjects (25 men, 34 women; age range = 18–30 years) completed two experimental sessions, spaced one week apart. The power analysis indicated that a sample size of 59 subjects was required for the study to have 80% power to detect a medium-sized effect ($p = .35$) at an α level of .05; thus, we stopped collecting data when we reached this number of subjects. Subjects were paid \$10 per hour for their participation in this study. Seven subjects missed their scheduled appointment for Session 2 and were rescheduled to complete it within 3 days of the missed appointment. Because their data were not anomalous in any way compared with data from the subjects who completed the sessions exactly 1 week apart, these subjects were included in the sample.

On a given day, each subject participated in two tasks: a temporal-numerosity judgment task (Shams et al., 2000, 2005) and a spatial localization task (Körding et al., 2007; Wozny et al., 2010). The tasks were performed successively with a 7-min break in between, and the order of the two tasks was counterbalanced across subjects.

Temporal task

The temporal task in this study consisted of counting the number of flashes and beeps that were presented, which ranged from one to four. Visual stimuli consisted of brief flashes of a white disk (1.5° of visual angle in diameter) presented on a CRT monitor for one frame (~10 ms); auditory stimuli consisted of brief beeps (3.5-kHz carrier frequency at a 68-dB sound-pressure level) played from speakers on the side of the monitor for 10 ms each; if multiple stimuli were presented within a single modality, the stimulus onset asynchrony (i.e., the time between the onset of two consecutive pulses) was 60 ms. For unisensory trials, one to four flashes or beeps occurred. For bisensory trials, the centers of the visual and auditory trains were aligned. For instance, if equal numbers of beeps and flashes were presented, the beeps and flashes were perfectly synchronized. If two beeps and one flash were presented, the flash fell halfway between the onset

of the two beeps; if three beeps and one flash were presented, the onset of the flash was synchronous with the onset of the second beep, and so on. The 24 possible experimental conditions consisted of pseudorandomly interleaved unisensory visual, unisensory auditory, and bisensory presentations. Fifteen trials per condition were presented, for a total of 360 trials. Therefore, 75% of the bisensory trials presented incongruent stimuli. The 360 trials were divided into four blocks, with 90 trials in each block. Subjects were allowed 1 to 2 min to rest between blocks. Before subjects began the experiment, they took part in a 16-trial practice phase consisting of both unisensory and bisensory stimuli. No feedback was provided during the practice phase.

Each trial in the experiment began with subjects fixating a centrally presented cross. Once fixation was established by a ViewPoint EyeTracker (Arrington Research, Scottsdale, AZ), visual stimuli were displayed at 7° of visual angle below fixation, auditory stimuli were played from the speakers, and subjects were prompted with a sentence on the screen to report the perceived number of stimuli (1–4) in each modality. Responses were given by pressing a number key on a wireless keyboard. In unisensory visual trials, subjects were presented with the instruction, "Report the number of flashes you see." In unisensory auditory trials, subjects were presented with the instruction, "Report the number of beeps you hear." On bisensory trials, observers were prompted to report their percept in each of the two modalities. The order of responses was always the same for a given subject (flash-beep or beep-flash); however, this order was counterbalanced across subjects. On bisensory trials, the number of flashes and beeps presented could be the same (congruent trials) or different (incongruent trials).

Spatial task

The spatial task in this study consisted of localization of visual, auditory, and audiovisual stimuli along a display axis, with five possible stimulus positions (–13°, –6.5°, 0°, 6.5°, and 13°). Visual stimuli were presented by a ceiling-mounted projector (with a resolution of 1,280 × 1,024 pixels and a refresh rate of 75 Hz) onto an acoustically transparent black cloth subtending much of the visual field (134° width × 60° height), located 52 cm in front of the observers. Auditory stimuli were played from free-field speakers (5 × 8 cm; extended range, paper cone) behind the cloth in positions that coincided with the locations of the projected visual stimuli. Visual stimuli consisted of a white disk (0.41 cd/m²) masked with a Gaussian envelope of 1.5°, and auditory stimuli were ramped white-noise bursts with a sound-pressure level of 59 dB at a distance of 52 cm from the speaker. Both auditory and visual stimuli were presented for only 35 ms.

The black cloth was draped over the large wooden frame that housed the speakers and covered the entire range of the subjects' field of view when their chins rested on the chinrest, even though only a limited spatial range directly in front of the observer (spanning 26° of the display axis) was tested. Conditions included five unisensory visual stimuli (from -13° to $+13^\circ$), five unisensory auditory stimuli (from -13° to $+13^\circ$), and 25 spatial combinations of bisensory stimuli. Therefore, 80% of the bisensory trials presented spatially incongruent stimuli. Fifteen trials per condition were presented, for a total of 525 trials. Each trial began with a fixation cross, and an eye tracker was used to ensure that subjects were fixating properly before presenting stimuli.

Once subjects were fixating within 3.0° of the fixation cross, stimuli were presented 7° below fixation along the display axis. After their presentation, a cursor appeared at a random location on a screen. Subjects were asked to quickly and accurately localize all stimuli that were displayed; this could be a flash of light, a burst of sound, or both, and stimuli could be either spatially congruent or spatially incongruent. The response cursor was displayed on the screen by the ceiling-mounted projector and could move continuously along the display axis; subjects could move the cursor either to the left or right with the trackball on the mouse, and they pressed a mouse button to record their responses. The cursor could move beyond the location of the most eccentric stimulus positions (-13° and $+13^\circ$), so the response range was not constrained. For bisensory stimuli, the auditory and visual signals were always temporally synchronous, and subjects always localized the light and sound in the same order, but the order of response (light-sound or sound-light) was counterbalanced across subjects.

Using the BCI model, the binding tendency was quantitatively estimated for each individual in each task on each day. If sensory integration in the brain is based on a *global mechanism*, then the estimated binding tendency should be correlated across tasks; if it is based on a *task-specific mechanism*, then the binding tendency should show no correlation across tasks. In addition, if integration processes are stable in the brain, binding tendencies should show consistency across time; alternatively, if the binding tendency is volatile and influenced by transient factors, estimated values for the binding tendency in a given task across two sessions should show little or no correlation.

Computational modeling methods

Three factors affect the strength of cross-modal interactions: the degree of consistency, similarity, or congruence of the stimuli (factor a); the relative reliability and precision of representation of the unisensory signals (factor b;

Fig. 1a); and the strength of the brain's tendency to bind sensory information across the relevant modalities (factor c; Fig. 1b). As in many previous studies of multisensory integration, all of our participants were presented with the same set of stimuli, effectively controlling for factor a. Thus, the variability in cross-modal interactions across individuals could be due to either factor b or factor c (see Fig. 1). Because it is already well established that observers vary in their unisensory abilities, factor b was not of interest in this study. To investigate factor c, we used a BCI model to characterize and quantify the binding tendency in each individual observer in a fashion that was not confounded with the precision of unisensory encoding.

We used a variant of a BCI model with four free parameters to model each subject's data (for details, see Wozny et al., 2010) from each task. The data from each task and session were kept separate in the modeling work, which resulted in four sets of parameter fits for each subject. The free parameters in the model included the standard deviation of the visual likelihood, the standard deviation of the auditory likelihood, the standard deviation of a central prior over space, and, most importantly, the prior probability of a common cause, which we call the "binding tendency." Three possible perceptual-decision strategies with 10 different sets of initial seeds for each strategy were used in optimizing the model parameters for each subject (30 initial seeds total) for each subject's data from Session 1 and Session 2. Parameters from the best-fitting decision strategy and seed were used for the final analysis.

Results

Our results indicate that although subjects' binding tendencies were quite stable across time within a domain, they showed little evidence of generalizing across tasks (Fig. 2). Specifically, in the spatial task, subjects were extremely consistent from one session to the next; the estimated binding tendency from Session 1 was quite similar to the estimated binding tendency from Session 2 (differences between sessions: $M = -0.005$, $SD = 0.18$), $r = .86$, $p < .00001$ (Fig. 2a). In the numerosity-judgment task, subjects were slightly more variable in their binding tendencies but still showed a strong consistency from one session to the next (differences between sessions: $M = 0.03$, $SD = 0.25$), $r = .64$, $p < .00001$ (Fig. 2b). Comparisons across tasks within a session, however, revealed no correlation between the binding tendency for the spatial task and the binding tendency for the temporal task—Session 1: $r = .11$, $p = .37$ (Fig. 2c); Session 2: $r = .08$, $p = .54$ (Fig. 2d).

For the stability analyses, we used strict criteria to evaluate the consistency of subjects' binding tendencies

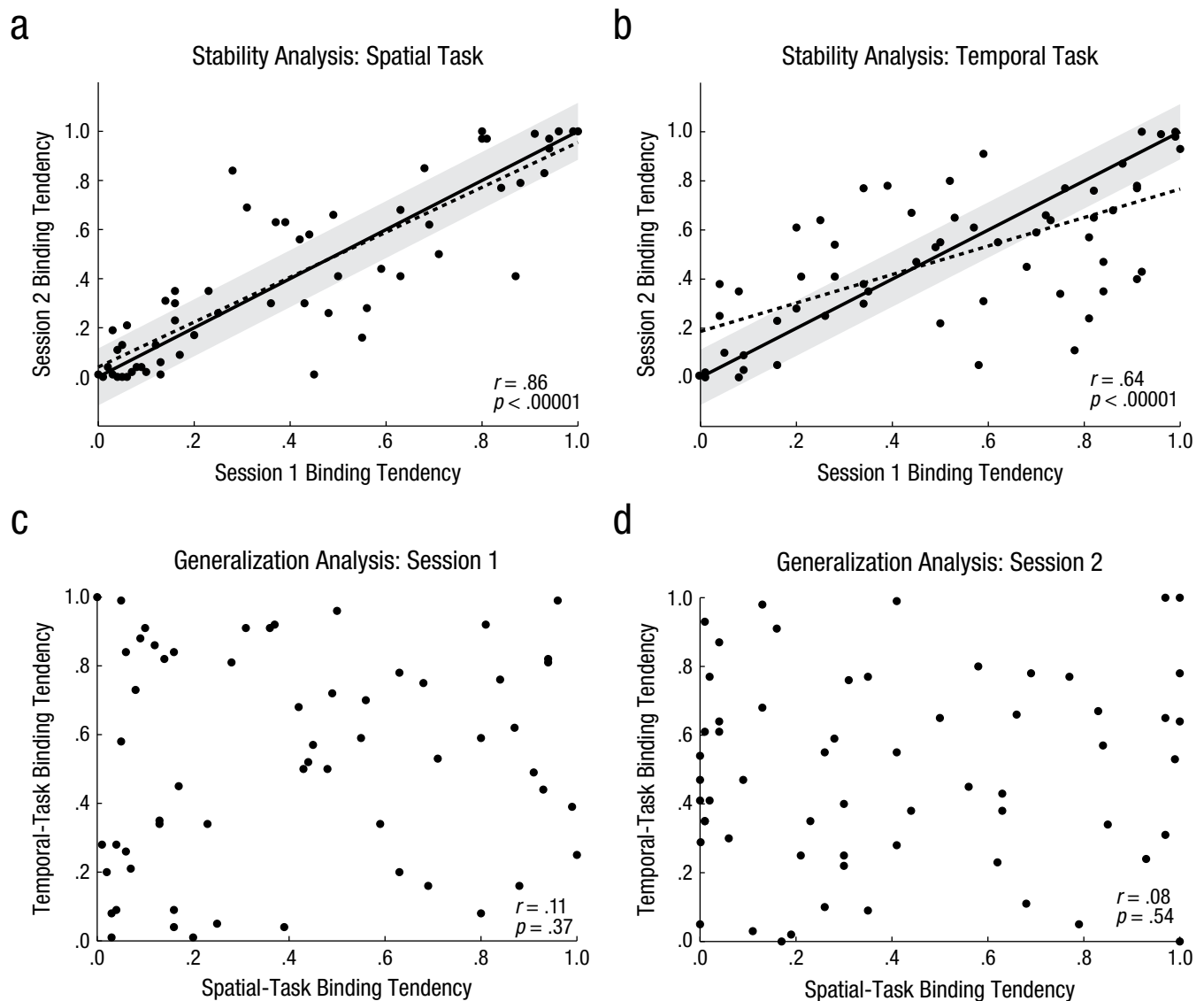


Fig. 2. Stability and generalization of the binding tendency: the relationship between the binding tendency in Session 1 and Session 2 for the (a) spatial and (b) temporal tasks and the relationship between the binding tendency in the temporal task and the spatial task in (c) Session 1 and (d) Session 2. The black dots represent the best model fit from 10 initial seeds. In (a) and (b), the black lines are the identity lines, which represent perfect stability, and the dashed lines are the best-fitting regression lines; the shaded areas show the standard deviation of optimal parameter fits (for the same data set, but using different seeds for fitting), which gives an indication of how much noise could be expected as a result of variability in the parameter-fitting procedure.

from one session to the next. Consistent binding tendencies should fall near the identity line in Figures 2a and 2b. It is important to note that we determined the binding tendencies by applying a parameter optimization procedure using 10 sets of initial seeds. The same exact data set might result in different estimates of binding tendency depending on the initial state of the optimization procedure (i.e., the parameter seeds used). To get a measure of the reliability of the parameter-optimization process (or, in other words, to get a sense of the degree to which variability in the estimation process can be expected to affect variability in the estimates of binding tendency), we

calculated the standard deviation of the binding tendency estimate across different parameter-estimation runs. Therefore, even if the binding tendency for an individual was perfectly constant over time, we would still expect the binding tendency estimates to deviate somewhat from the unity line within this range some of the time.

To obtain the standard deviation of the estimate of the binding tendency parameter, we obtained optimal parameter fits five times for each subject and each session, each time using a distinct set of 10 initial seeds. We then calculated the standard deviation of the five optimal binding-tendency estimates from each set and took the

average of these standard deviations across all data sets (all sessions and all participants). This provided a reliable estimate of variability (or noise) in the estimation of the binding-tendency parameter (see Figs. 2a and 2b). In the spatial task, 33 of 59 subjects fell within 1 standard deviation of the estimate, and in the temporal task, 29 of 59 subjects fell in this region, which indicates that about half of subjects were extremely stable, and the values for most of the remaining subjects did not drastically differ from those of the subjects who fell within 1 standard deviation of the estimate of variability.

The mean deviation of the binding tendency, which is the average change from Session 1 to Session 2 across all subjects, was minimal for both tasks ($-.005$ for the spatial task, $.03$ for the temporal task), but the temporal task showed somewhat more deviation. Likewise, the variability of deviations from zero was larger for the temporal task ($.25$) than for the spatial task ($.18$). Therefore, overall, the spatial task showed a stronger stability than the temporal task.

Our criterion for generalization was more lax than the criterion for stability. Given that the tasks are quite different, the required binding tendency should not be the same across tasks. However, if the binding tendency is driven by a common factor (e.g., connectivity between auditory and visual regions), then there should be at least a correlation between the binding tendencies across the two tasks. For example, if a subject has a large binding tendency in one task, that subject should show a large (although not necessarily identical) binding tendency in the other task as well, and the same would be true for small tendencies. Thus, in these analyses, we examined whether any correlation existed between binding tendencies across the two tasks. Figures 2c and 2d show that there was no relationship between the binding tendency from the spatial task and the binding tendency from the temporal task. In other words, the tendencies across tasks were neither identical nor correlated.

Given that binding tendency is an important factor driving cross-modal interactions, one would expect it to correlate with cross-modal interactions observed in subjects' behavior. A commonly used behavioral measure of cross-modal interactions is cross-modal bias, which refers to the amount of bias in estimates of the less reliable modality as a result of influences of the more reliable modality. For example, in the spatial task, auditory estimates are often pulled toward the location of visual stimuli, and the degree to which this occurs is known as *auditory bias*. Figures 3a and 3b show that the correlation between binding tendency and cross-modal interaction, as measured by the amount of auditory bias ($A_{\text{response}} - A_{\text{location}}/V_{\text{location}} - A_{\text{location}}$) in the spatial task, was quite strong: $r = .84$ for Session 1 and $r = .89$ for Session 2, $p < .00001$ for both sessions. In the temporal-numerosity

task, the estimated number of visual events is often biased by the number of auditory events and can be quantified by *visual bias* ($V_{\text{response}} - V_{\text{number}}/A_{\text{number}} - V_{\text{number}}$). Indeed, as shown in Figures 3c and 3d, the correlation between binding tendency and visual bias in the numerosity task was $.57$ for Session 1 and $.46$ for Session 2, $p < .001$ for both days. These results show how the binding tendency parameter, which provides a clean, unconfounded measure of the tendency to bind sensory information, drove the strength of the interactions between the auditory and visual modalities in the two tasks.

Discussion

In this study, we used a BCI model to characterize and quantify the binding tendency in each individual observer in a manner that did not confound binding tendency with the precision of unisensory encoding. Our results regarding the binding tendency provide strong evidence that spatial and temporal integration processes are not governed by a single, universal parameter in the brain. Instead, integration processes within these domains are governed by distinct perceptual biases that are, nonetheless, stable over time.

The binding tendency's observed stability over time makes it more tractable and allows future studies to determine how this tendency may be modulated in various domains. Given that research shows that a deficit in multisensory integration is associated with several disorders, such as autism (e.g., Foxe et al., 2015; Stevenson, Siemann, Schneider, et al., 2014; Stevenson, Siemann, Woynarowski, et al., 2014; Wallace & Stevenson, 2014), dyslexia (e.g., Hahn, Foxe, & Molholm, 2014; Harrar et al., 2014), and schizophrenia (e.g., Stekelenburg, Maes, Van Gool, Sitskoorn, & Vroomen, 2013; Szycik et al., 2009; Williams, Light, Braff, & Ramachandran, 2010), the elucidation of the characteristics of the binding tendency could have important clinical and educational implications. In addition, the stability and task specificity of the binding tendency observed in the present study is consistent with the stability and task specificity reported elsewhere for synesthesia, which could be interpreted as an extreme, unique manifestation of the binding tendency (Ghazanfar & Schroeder, 2006; Newell & Mitchell, 2015; Parise & Spence, 2009).

Recent research has revealed that cross-modal interactions start at early stages of sensory processing, including primary cortical areas that were traditionally believed to be strictly unisensory (e.g., Foxe & Schroeder, 2005; Ghazanfar & Schroeder, 2006). Although the BCI model used here for the analysis of behavioral data is a computational model (as opposed to a neural model) and therefore does not make explicit predictions about the

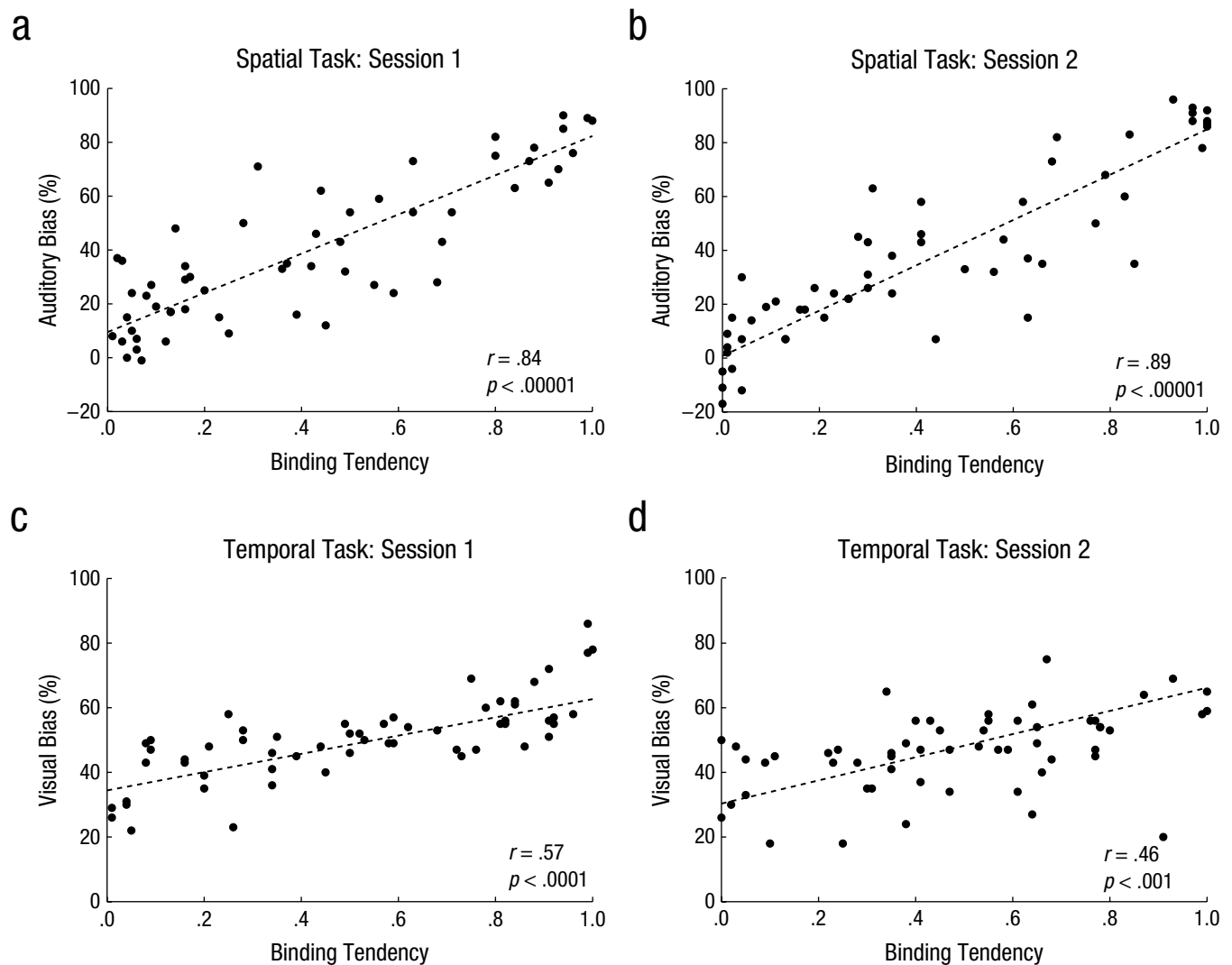


Fig. 3. Scatterplots (with best-fitting regression lines) showing the relationship between the average amount of auditory bias across all incongruent conditions and the binding-tendency estimates in the spatial task in (a) Session 1 and (b) Session 2, and the relationship between the average amount of visual bias across all incongruent conditions and binding-tendency estimates in the temporal task in (c) Session 1 and (d) Session 2. Auditory bias was computed as in Körding et al. (2007). For each spatially discrepant trial, auditory bias was computed by subtracting the actual auditory location from the subject's auditory spatial estimate and dividing that difference by the distance between the spatially discrepant visual and auditory stimuli. We computed this measure for all trials in each incongruent condition, and we computed the average bias across all conditions. The visual bias for a temporally discrepant trial was computed by subtracting the actual number of visual events that occurred from each subject's visual numerosity judgment and dividing that difference by the difference between the number of auditory beeps and number of visual flashes that were presented. We computed this measure for all trials in each incongruent condition and then computed the average bias across all conditions.

locations of unisensory and multisensory processing areas in the brain, one recent study using functional MRI and a spatial localization task made progress in this area by mapping estimates from this model onto different cortical areas (Rohe & Noppeney, 2015a). However, more research is needed to shed light on the neural, anatomical, and physiological correlates of the binding tendency in the nervous system.

Finally, it is important to note that in the current study, only one task was investigated in each domain (i.e.,

temporal and spatial); thus, it remains unclear whether the lack of generalization between the two tasks reflects task specificity or domain specificity of the binding tendency. Thus, future studies should map the boundary of the task specificity found here by exploring the binding tendency in different tasks within the same or similar perceptual domains.

Action Editor

Ralph Adolphs served as the action editor for this article.

Author Contributions

L. Shams developed the study concept. L. Shams and B. Odegaard contributed to the study design. Testing and data collection were performed by B. Odegaard. B. Odegaard performed the data analysis and computational modeling simulations under the supervision of L. Shams. B. Odegaard and L. Shams drafted the manuscript and approved the final version of the manuscript for submission.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Funding

This work was supported by a National Science Foundation grant (BCS-1057969) and a UCLA senate grant.

References

- Alais, D., & Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Current Biology*, *14*, 257–262. doi:10.1016/j.cub.2004.01.029
- Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*, 429–433. doi:10.1038/415429a
- Foxe, J. J., Molholm, S., Bene, V. A. D., Frey, H.-P., Russo, N. N., Blanco, D., . . . Ross, L. A. (2015). Severe multisensory speech integration deficits in high-functioning school-aged children with autism spectrum disorder (ASD) and their resolution during early adolescence. *Cerebral Cortex*, *25*, 298–312. doi:10.1093/cercor/bht213
- Foxe, J. J., & Schroeder, C. E. (2005). The case for feedforward multisensory convergence during early cortical processing. *NeuroReport*, *16*, 419–423.
- Ghazanfar, A. A., & Schroeder, C. E. (2006). Is neocortex essentially multisensory? *Trends in Cognitive Sciences*, *10*, 278–285. doi:10.1016/j.tics.2006.04.008
- Hahn, N., Foxe, J. J., & Molholm, S. (2014). Impairments of multisensory integration and cross-sensory learning as pathways to dyslexia. *Neuroscience & Biobehavioral Reviews*, *47*, 384–392. doi:10.1016/j.neubiorev.2014.09.007
- Hairston, W. D., Wallace, M. T., Vaughan, J. W., Stein, B. E., Norris, J. L., & Schirillo, J. A. (2003). Visual localization ability influences cross-modal bias. *Journal of Cognitive Neuroscience*, *15*, 20–29. doi:10.1162/089892903321107792
- Harrar, V., Tammam, J., Pérez-Bellido, A., Pitt, A., Stein, J., & Spence, C. (2014). Multisensory integration and attention in developmental dyslexia. *Current Biology*, *24*, 531–535. doi:10.1016/j.cub.2014.01.029
- 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), Article e943. doi:10.1371/journal.pone.0000943
- Mallick, D. B., Magnotti, J. F., & Beauchamp, M. S. (2015). Variability and stability in the McGurk effect: Contributions of participants, stimuli, time, and response type. *Psychonomic Bulletin & Review*, *22*, 1–9. doi:10.3758/s13423-015-0817-4
- Newell, F. N., & Mitchell, K. J. (2015). Multisensory integration and cross-modal learning in synaesthesia: A unifying model. *Neuropsychologia*. Advance online publication. doi:10.1016/j.neuropsychologia.2015.07.026
- Parise, C. V., & Spence, C. (2009). “When birds of a feather flock together”: Synesthetic correspondences modulate audiovisual integration in non-synesthetes. *PLoS ONE*, *4*(5), Article e5664. doi:10.1371/journal.pone.0005664
- Parise, C. V., Spence, C., & Ernst, M. O. (2012). When correlation implies causation in multisensory integration. *Current Biology*, *22*, 46–49. doi:10.1016/j.cub.2011.11.039
- Rohe, T., & Noppeney, U. (2015a). Cortical hierarchies perform Bayesian causal inference in multisensory perception. *PLoS Biology*, *13*(2), Article e1002073. doi:10.1371/journal.pbio.1002073
- Rohe, T., & Noppeney, U. (2015b). Sensory reliability shapes perceptual inference via two mechanisms. *Journal of Vision*, *15*(5), Article 22. doi:10.1167/15.5.22
- Shams, L., & Beierholm, U. R. (2010). Causal inference in perception. *Trends in Cognitive Sciences*, *14*, 425–432. doi:10.1016/j.tics.2010.07.001
- Shams, L., Kamitani, Y., & Shimojo, S. (2000). Illusions: What you see is what you hear. *Nature*, *408*, 788. doi:10.1038/35048669
- Shams, L., Ma, W. J., & Beierholm, U. (2005). Sound-induced flash illusion as an optimal percept. *NeuroReport*, *16*, 1923–1927.
- Stein, B. E. (2012). *The new handbook of multisensory processes*. Cambridge, MA: MIT Press.
- Stekelenburg, J. J., Maes, J. P., Van Gool, A. R., Sitskoorn, M., & Vroomen, J. (2013). Deficient multisensory integration in schizophrenia: An event-related potential study. *Schizophrenia Research*, *147*, 253–261. doi:10.1016/j.schres.2013.04.038
- Stevenson, R. A., Siemann, J. K., Schneider, B. C., Eberly, H. E., Woynaroski, T. G., Camarata, S. M., & Wallace, M. T. (2014). Multisensory temporal integration in autism spectrum disorders. *The Journal of Neuroscience*, *34*, 691–697. doi:10.1523/JNEUROSCI.3615-13.2014
- Stevenson, R. A., Siemann, J. K., Woynaroski, T. G., Schneider, B. C., Eberly, H. E., Camarata, S. M., & Wallace, M. T. (2014). Evidence for diminished multisensory integration in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *12*, 3161–3167. doi:10.1007/s10803-014-2179-6
- Stevenson, R. A., & Wallace, M. T. (2013). Multisensory temporal integration: Task and stimulus dependencies. *Experimental Brain Research*, *227*, 249–261. doi:10.1007/s00221-013-3507-3
- Stevenson, R. A., Zemtsov, R. K., & Wallace, M. T. (2012). Individual differences in the multisensory temporal binding window predict susceptibility to audiovisual illusions. *Journal of Experimental Psychology: Human Perception and Performance*, *38*, 1517–1529. doi:10.1037/a0027339
- Szyck, G. R., Münte, T. F., Dillo, W., Mohammadi, B., Samii, A., Emrich, H. M., & Dietrich, D. E. (2009). Audiovisual integration of speech is disturbed in schizophrenia: An fMRI study. *Schizophrenia Research*, *110*, 111–118. doi:10.1016/j.schres.2009.03.003

- Tremblay, C., Champoux, F., Bacon, B. A., & Theoret, H. (2007). Evidence for a generic process underlying multisensory integration. *Open Behavioral Science Journal*, *1*, 1–4. doi:10.2174/1874230000701010001
- Wallace, M. T., & Stevenson, R. A. (2014). The construct of the multisensory temporal binding window and its dysregulation in developmental disabilities. *Neuropsychologia*, *64*, 105–123. doi:10.1016/j.neuropsychologia.2014.08.005
- Williams, L. E., Light, G. A., Braff, D. L., & Ramachandran, V. S. (2010). Reduced multisensory integration in patients with schizophrenia on a target detection task. *Neuropsychologia*, *48*, 3128–3136. doi:10.1016/j.neuropsychologia.2010.06.028
- Wozny, D. R., Beierholm, U. R., & Shams, L. (2010). Probability matching as a computational strategy used in perception. *PLoS Computational Biology*, *6*(8), Article e1000871. doi:10.1371/journal.pcbi.1000871