Non-conscious Multisensory Integration in the Ventriloquist Effect

Elise Turkovich[1], Soorya Sankaran[1], Wei Dou[1], Jason Samaha[1*]
[1]Department of Psychology, University of California, Santa Cruz
[*]correspondence: jsamaha@ucsc.edu

Abstract
The degree to which information from distinct sensory modalities can interact in the absence of conscious awareness remains controversial. According to the Global Neuronal Workspace Theory (GNWT), unconscious sensory information remains relatively confined to sensory cortex and should not be capable of interacting with other modalities until it is broadcast into the (conscious) global workspace comprising late (>300ms) frontal-parietal activation. The ventriloquist effect is a classical multisensory integration phenomenon that refers to the misperception of sound location towards concurrent visual stimulation, such as perceiving the voice of a ventriloquist actor as coming from the moving dummy. Here we used meta-contrast masking to render a brief flash stimulus non-conscious while participants performed a sound localization task. We found that, despite being at chance performance in discriminating the flash location, participants were nevertheless biased to localize sounds towards the unconscious flash locations. The effect was present in virtually all participants, was nearly as large as the effect on conscious trials, and was robust to controls for individual differences in task performance. Decoding analyses of concurrently recorded EEG signals showed that the non-conscious flash location information was present up until around 230ms but not after; confirming that the visual influence on sound perception likely occurred before conscious broadcast. Our findings suggest that subjective perception is not required for the integration of signals originating in distinct sensory modalities prompting new questions about the role of subjective perception in multisensory integration.

Introduction

The role that conscious awareness plays in cognitive and perceptual functions has been a central topic in the development of theories of consciousness (Block, 1995; Cohen & Dennett, 2011; Ludwig, 2023; Samaha, 2015; Seth & Bayne, 2022). Most current neuroscientific theories of consciousness propose that conscious perception confers some functional advantage that non-conscious processing lacks (Lau & Rosenthal, 2011; Seth & Bayne, 2022). One such function common to many approaches is that consciousness has a role in integrating neural processes that would otherwise be independent. (Such a common proposal that it has been referred to as the "integration consensus" (Morsella et al., 2016)). A prime example of the proposed integrative function of consciousness is in the global (neuronal) workspace theory (GNWT) of consciousness (Baars, 2005; Dehaene & Changeux, 2011; Dehaene & Naccache, 2001; Mashour et al., 2020), which posits that non-conscious processing remains relatively siloed within specialized processing modules where it is inaccessible to other cognitive, mnemonic, motor, evaluative, and perceptual processes. Consciousness, on the other hand, reflects the contents that are broadcast beyond local processing and are widely accessible by other processing modules, putatively mediated by a long-range, recurrent, frontal-parietal network (Mashour et al., 2020). Such broadcast is thought to happen relatively late (around 250 ms post-stimulus) in a non-linear, "all-or-none" fashion, and to underlie a subject's ability to report their perceptual experiences (Dehaene et al., 2014; Del Cul et al., 2007; Salti et al., 2015; van Vugt et al., 2018).

One prediction that follows from the integration function proposed by GNWT theory is that unconscious processing in distinct sensory modules should not interact. Or, as Baars puts it *'consciousness is needed to integrate multiple sensory inputs'* (Baars, 2002 p. 47–48). Several lines of evidence support the idea that non-conscious sensory processing remains local and confined to the cortical areas selective for the type of sensory information. For example, masked visual stimuli that are reported as unseen by the observer elicit brain responses in early and mid-level visual areas but typically not in non-visual areas or higher-order association areas (Dehaene et al., 2001; Del Cul et al., 2007; Sergent et al., 2005). Non-conscious processing of auditory stimuli does not elicit spiking activity in somatosensory areas (Noel et al., 2019) and does not lead to global broadcast (Sadaghiani et al., 2009) particularly, when also task-irrelevant (Sergent et al., 2021). In contrast to this prediction, however, a growing number of studies point to conditions under which integration can occur non-consciously (Mudrik et al., 2014) even when both of the to-be-integrated stimuli are non-conscious and come from different sensory modalities (Scott et al., 2018).

A canonical example of multi-sensory integration is the Ventriloquist effect (VE), whereby the location of a visual stimulus biases the perceived location of a concurrent auditory stimulus, such as when perceiving the moving mouth of a ventriloquist dummy as the spatial source of vocalization (Alais & Burr, 2004; Vroomen & De Gelder, 2004). Recent evidence suggests that visual stimuli which are not consciously perceived may nevertheless bias the perceived location of auditory stimuli, causing a VE. Delong et al., (2018) used dynamic continuous flash suppression (dCSF) to render a brief flash invisible while having subjects judge the spatial location of a concurrent white noise burst. They found a reliable ventriloquist effect across two experiments, suggesting that the non-conscious flash still biased the perceived sound location (Delong et al., 2018). However, one potential drawback of the use of dCSF for masking in this context is that subjects may have actually perceived the brief flash stimulus consciously but simply confused it for a part of the mask, given the visually complex and dynamic appearance of the Mondrian patterns used, which varied over space and time in color, motion, and luminance, among other features. Additionally, prior studies have not yet investigated the timecourse of putative non-conscious neural processing in the VE.

Here, we conducted a multi-day behavioral and EEG experiment with 30 subjects using meta-contrast masking to render a brief flash invisible while subjects judged the location of a concurrent auditory stimulus. Using both objective criteria for awareness (chance performance in flash localization) and subjective criteria (reports of “no experience” on the perceptual awareness scale; PAS), we found a robust non-conscious VE that was present in virtually every subject. Decoding location information from concurrently recorded EEG signals revealed that non-conscious representations of the visual stimulus location were present only for an initial 300 ms, consistent with a lack of late, global broadcast and implying that the non-conscious integration with auditory processing occurs prior to 300 ms.

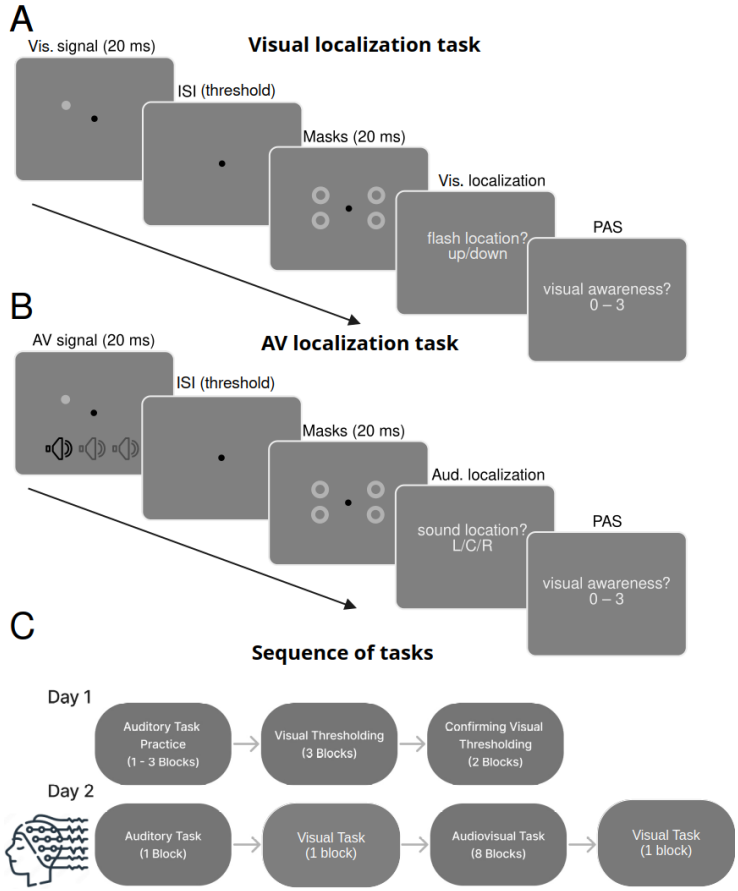

Figure 1. *Tasks and procedure.* ***A)*** *In the unimodal visual localization task, participants judged the vertical location (up or down) of a brief flash target and then rated subjective visibility on the 4-point Perceptual Awareness Scale (PAS). Targets were backwards-masked with metacontrast annuli that appeared at all possible target locations in each visual quadrant. The target luminance and inter-stimulus interval (ISI) between mask and target were determined in a separate thresholding session so as to produce near chance localization accuracy on trials judged as invisilbe (PAS = 0).* ***B)*** *In the main audio-visual (AV) localization task, participants were shown threshold flash stimuli concurrently with supra-threshold auditory stimuli presented at three simulated locations across the horizontal plane (using HRTFs; see methods)> Participants first judged the auditory stimulus location (left, center, or right) and then flash visibility using the PAS. Visual and auditory locations were randomized so that one did not inform the location of the other.* ***C)*** *Sequence of tasks across two days of testing. On Day 1 (behavior only) participants practiced auditory localization (in the absence of visual stimuli), then performed several blocks of visual thresholding in order to find the ISI and flash luminance that achieved at or near chance performance on trials rated as subjectively invisible. On Day 2 EEG was recorded while subjects performed a pre-task visual-only block to check for chance performance on invisible trials, the main audiovisual task, and then a post-task visual-only block.*

Methods

Participants

32 members of the UC Santa Cruz community participated in this two-day study (age range 18-32, 18 Female, 12 Male, 2 Non-Binary). All subjects were compensated with a $20 Amazon gift card and class credit when appropriate. One subject was excluded from analyses due to excessive noise in the EEG signal and another due to blinking in response to most auditory stimuli, leaving 30 subjects in the final analyses (unless otherwise noted). The study was approved by the UC Santa Cruz Institutional Review Board.

Stimuli

The experiment was implemented in MATLAB (The MathWorks Inc., 2022) and all stimuli were generated using the Psychophysics Toolbox Version 3 (Kleiner et al., 2007; Pelli, 1997). Stimuli were presented on a uniform gray background (approx. 50 cd/m^2) on a gamma-corrected VIEWPixx EEG monitor (53.4 x 30 cm, 1920 x 1080 resolution, 120 Hz refresh rate). Participants were seated in a dimly lit room approximately 74 cm away from the screen with their head stabilized on a chin rest. Auditory stimuli were delivered using Etymotic Research 3C insert earphones at an (individually adjusted) comfortable volume.

To deliver auditory stimuli originating from varying spatial locations, the MIT KEMAR head related transfer function (HRTF) was convolved with the auditory stimulus (Gardner & Martin, 1995). Specifically, the timing and amplitude profile of a 20ms burst of white noise (generated randomly on each trial) was manipulated for each ear to simulate sound from 5 deg. to the left, center, or 5 deg. right, along the azimuth with zero elevation.

The target visual stimulus was a circular luminance increment subtending 1 deg. of visual angle, whose luminance contrast with respect to the gray background was individually determined in an initial thresholding procedure (see below). The masking stimulus was an annulus with an outer diameter of 2 deg., an inner diameter of 1 deg., and a luminance of ~70 cd/m^2. A central crosshair plus bullseye (0.6 deg.) was used as a fixation point

General Procedure

We modeled our task design closely on that of Delong et al., (2018) with the primary difference being that we used meta-contrast masking rather than dCFS to render visual stimuli invisible. Participants completed this experiment over a two-day period (see Figure 1). On the first day, lasting approximately one hour, participants practiced the auditory localization task (described next) until they achieved above-chance performance. Then participants completed the visual thresholding task, designed to ensure at or close to chance performance on the visual localization task. Both of these tasks were in preparation for the main task on Day 2, during which we aimed for participants to be proficient at auditory localization but unaware of the visual stimulus on a majority of trials. On Day 2 (which occurred within one week of Day 1), participants were fitted with EEG, completed one more practice block of the auditory localization task, a visual only block, eight blocks of the main AV task, and then a final visual only block. Each of these tasks is described in greater detail below.

Auditory localization task

In the auditory localization task, the participant was presented with a 20ms burst of white noise originating from 5 deg. left, center, or 5 deg. right. The task was to judge the sound's location using one of three options, 'left', 'center', or 'right', using the keys 7, 8, and 9, respectively, with the right hand. An inter-trial-interval (ITI) randomly sampled between was 0.9 and 1.4 sec followed the response. Given the difficulty of discriminating sound locations separated by only 5 deg., we repeated blocks of 72 trials with visual feedback displaying the correct answer on every trial until each participant performed higher than 45% accuracy across an entire block (where chance is 33%). This typically took between 1-3 blocks, all performed on Day 1. A single block was also

run at the beginning of Day 2 to remind participants of the auditory judgment before proceeding to the rest of the experiment.

Visual thresholding task

The visual thresholding task was designed to find the ideal masking conditions for rendering the flash invisible on a majority of trials. To establish the optimal inter-stimulus interval (ISI) between the target flash and the mask for each participant, 3 blocks (576 trials total) of a purely visual masking task were conducted on Day 1. The visual task involved fixating while a brief (20ms) flash was presented at one of four quadrant locations, 1.2 deg. above or below and 5 deg. left or right from fixation. After a brief ISI, a mask was presented for 20ms in all four quadrants, encompassing the space around all potential flash locations. The ISI varied between 0-40ms (corresponding to SOAs of 20-60ms), in steps of 8ms. While doing the task, participants are asked 1) to rate their perceptual awareness of the target stimulus on the 4-point perceptual awareness scale (PAS; Sandberg et al., 2010) and 2) to discriminate the vertical location of the stimulus (i.e., upper or lower visual field). The numbers 0-3 of the PAS correspond to ‘not seen’, ‘brief glimpse’, ‘almost clear’, and ‘clear image’, respectively. Participants were repeatedly instructed to use a PAS response > 0 if they had any feeling whatsoever of having seen the stimulus. The PAS response aims to capture the participants' subjective invisibility on a trial-by-trial basis, i.e., does the participant believe they saw the flash? The location discrimination question aims to provide an objective measure of stimulus invisibility based on the reasoning that when location discrimination accuracy is indistinguishable from 50% (i.e., chance), the participant is likely unaware of the visual stimulus (Merikle et al., 2001).

At the end of all three thresholding blocks on Day 1, plots were produced displaying accuracy at discriminating the flash location at each ISI. The timing at which the participant was closest to chance is selected. If needed, the luminance of the flash was also reduced to reduce visibility in some subjects. Two subsequent shorter blocks (96 trials) of the same task are completed to confirm the decision and modify parameters slightly if needed.

Visual Localization Task

The masking parameters from the Day 1 thresholding procedure were fixed for all subsequent visual tasks in Day 2. Specifically, on Day 2, participants performed two blocks (48 trials each) of a visual only task identical to the visual thresholding task but with a constant ISI and flash luminance. The purpose was to test that performance was still near chance on trials rated as invisible.

Main Audio-Visual Task

The main task involved the presentation of a temporally congruent audio-visual (AV) stimulus with varying spatial locations. On Day 2, while collecting EEG data, participants complete eight blocks of the combined AV task (720 trials total), wherein the flash target was presented in temporal congruence with the sound burst, followed by the visual masks. Participants are asked (1) The location of the sound (left, center, or right) and (2) their perception of the visual stimulus on the PAS scale (0-3). The visual location judgment was omitted from this main task based on pilot data where subjects indicated occasionally confusing response options and difficulty remembering all responses when asked three questions per trial.

EEG Recording and Pre-Processing

EEG data were recorded at 1000 Hz using a 64-channel Ag/AgCl gel-based active electrode system (actiCHamp Plus, Brain Products) with electrode FCz as the online reference. The data were pre-processed using custom MATLAB scripts (version R2023a) and the EEGLAB toolbox (Delorme & Makeig, 2004). First, a band-pass filter was applied at from 0.1 – 50 Hz (zero-phase, Hamming-windowed sinc FIR filter), the data were

downsampled to 500 Hz, and then re-referenced to the median of all electrodes. Epoched data were then visually inspected and trials with excessive noise, muscle artifacts, or ocular artifacts overlapping with the stimulus presentation were rejected. Independent components analysis (ICA) was performed using the INFOMAX algorithm to remove 1 or 2 components per subject corresponding to blinks or eye movements. An average of 55 trials were removed (range: 10-167). Channels with excessive noise were removed and spherically interpolated (average of 1.4 channels; range: 0-6). Lastly, the data were re-referenced to the average across channels and a 200 ms pre-stimulus baseline was subtracted from each timepoint, channel, and trial.

Behavioral Analysis

All behavioral analyses were conducted on data from day two, which comprised the auditory-only task, the main AV task, as well as the two visual-only blocks conducted immediately before and after the main AV task (see Figure 1). We categorized trials as invisible when PAS=0 and visible when PAS>0. We computed a ventriloquist effect (VE) index by considering the mean sound localization report in deg. (i.e., -5, 0, 5) as a function of the actual sound location (in deg.) and the flash location (left or right or absent). We computed the VE (in deg.) as the difference in mean sound localization report between trials with a left or right flash. Accordingly, if the flash location had no influence on sound localization, the VE would be 0 and if the flash attracts (repels) the sound localization report, the VE is positive (negative). Thus, the classic VE would correspond to a positive value using this metric. We tested whether an unconscious VE occurred by computing the VE index separately for invisible (PAS=0) and visible (PAS>0) flash trials and comparing each index to 0 using repeated-measures t-tests. Bayes factors were also used to evaluate the evidence for the null (BF_{null}) or alternative (BF_{alt}) hypothesis using the Bayes Factor MATLAB package from Bart Krekelberg (https://github.com/klabhub/bayesFactor), which uses a JZS prior.

EEG Decoding Analysis

We used multivariate pattern analysis of EEG recordings to decode stimulus information from brain responses in order to study the timecourse of neural representations of stimulus location. We first sought to decode visual location information in order to track how neural processing of the flash changed with the subject's awareness. Based on the predictions of GNWT theory, we hypothesized that the flash location on invisible trials may be decodable early on (reflecting putatively unconscious processing) but should not be "broadcast" to the fronto-parietal workspace supporting sustained representation and thus should no longer be decodable after the time broadcast is assumed to occur (around 300 ms). To this end, we trained a linear discriminant analysis (LDA) classifier to predict the horizontal location (left or right) of the flash stimulus during the main AV task. Single-trial EEG data from all sensors were first downsampled to 250 Hz to speed up computation, z-scored across time, and undersampled to ensure an equal number of left and right flash trials so as not to bias classification. Classification was performed using 5-fold cross validation repeated a total of 5 times per participant and averaged over repetitions to ensure results were not driven by any random stratification during the creation of folds. We then averaged classification accuracy across all trials to get a general sense of decoding accuracy as well as by visibility rating (invisible, PAS = 0; visible PAS > 0). Because some subjects had too few trials either rated as visible or invisible for stratification into 5 folds, we restricted all decoding analysis to subjects with at least 50 trials in both visible and invisible conditions, leaving 23 subjects for decoding. Decoding was implemented in MATLAB using the MVPA-light toolbox (Treder, 2020).

To statistically evaluate decoder performance, we tested whether classification proportion correct at each time point differed from chance levels (0.5) and whether classification differed between visible and invisible trials using two-tailed repeated-measures t-tests. Cluster correction using the cluster size method (Maris & Oostenveld, 2007) was used to correct for multiple comparisons across time. Specifically, the mapping between

data and condition was randomized 5,000 times and all timepoints were tested for each permutation. The size of the largest cluster of contiguous significant timepoints was stored for each of the 5,000 permutations and only cluster sizes in the real data that exceeded the 95 percentile of this distribution of cluster sizes expected under the null hypothesis were considered statistically significant.

We also attempted to decode the sound location (left/center/right) using the same approach described above, however the decoder never reached significant above-chance classification at any timepoint when training and testing on all AV trials together so we did not further explore sound location decoding.

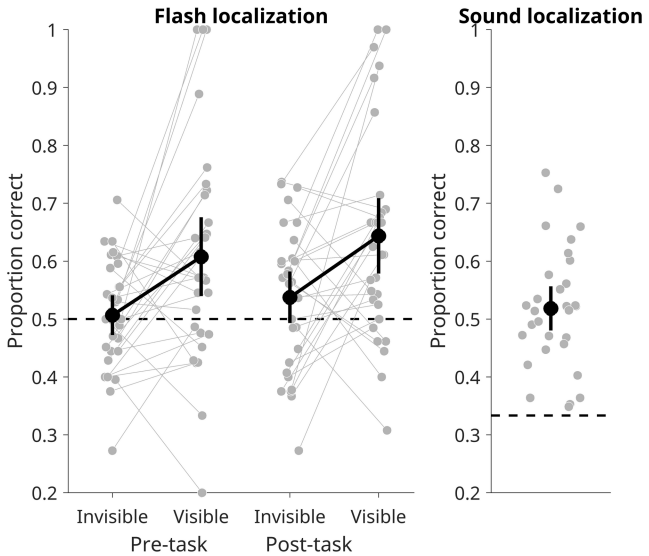

Figure 2. *Unimodal visual and auditory behavior on Day 2. Left panel shows pre- and post-AV task unimodal visual localization accuracy on trials judged as subjectively invisible (PAS = 0) and visible (PAS > 0). Error bars are 95% confidence intervals which overlap chance performance (0.5) on invisible trials, indicating that subjects could not, on average, localize the visual stimuli better than chance levels when they were subjectively unseen. Visible trials, on the other hand, were localized significantly better than chance. Right panel shows unimodal sound localization accuracy tested at the beginning of the Day 2 session. All subjects could localize the sounds better than chance levels (0.33). Error bars are 95% confidence intervals.*

Results

Unimodal Visual Localization and Awareness

We first determined whether flash localization accuracy was at chance levels on trials rated as invisible (i.e., PAS = 0; see Figure 2). In the pre-task visual only block, we found that mean (*SEM)* flash localization accuracy on invisible trials was 50.6% (1.75%) and did not differ from chance ($t(29) = 0.377$, $p = 0.708$) with a BF indicating substantial evidence for the null ($BF_{null} = 4.816$). Likewise, in the post-task visual only block, mean accuracy on invisible trials was 53.7% (2.27%), which did not differ from chance performance ($t(29) = 1.650$, $p = 0.109$), with a BF providing only inconclusive evidence in favor of the null ($BF_{null} = 1.534$). Comparing the two, invisible flash localization accuracy measured pre- and post-task did not differ ($t(29) = -0.961$, $p = 0.344$, $BF_{null} = 3.371$), suggesting that the thresholding procedure produced stable performance over the course of the experiment. Flash localization accuracy on trials judged as visible, however, was significantly above chance in both the pre-task block (M = 60.7%, SEM = 3.47%, $t(29) = 3.100$, $p = 0.004$, $BF_{alt} = 9.308$) and in the post-task block (M = 64.3%, SEM = 3.32%, $t(29) = 4.318$, $p = 0.004$, $BF_{alt} = 163.041$), with no difference between the

two ($t(29) = -0.634$, $p = 0.531$, $BF_{null} = 4.272$) .Together, this indicates that performance was stable across the experiment, was better than chance when subjectively aware, and was not better than chance when subjectively unaware.

Unimodal Auditory Localization

We next checked that auditory localization was above chance performance (Figure 2), which ensures that participants were using the timing and amplitude cues in the stimuli to localize the sounds. Indeed, average auditory localization accuracy was 51.8% (1.91%), well above the chance level of 33% ($t(29) = 9.664$, $p < 0.001$, $BF_{alt} > 999$).

Conscious and Non-Conscious Ventriloquist Effect

As seen in Figure 3, when considering all trials together, the presence of the flash on the left or right side of fixation caused sound location judgments to be shifted to the left and right, respectively, relative to flash absent trials. The difference in mean auditory localization reports between left and right flash trials provided an index of the VE, which we further broke down by trials rated as visible (PAS > 0) and invisible (PAS = 0). On visible trials, we observed a clear VE with visual stimuli biasing auditory localization towards them by an average of 2.31 deg. (0.29), which was significantly different than zero ($t(29) = 7.776$, $p < 0.001$, $BF_{alt} > 999$). Regarding our main hypothesis, we observed a VE of 1.90 deg. (0.24) even on invisible trials, which represents decisive evidence against the null hypothesis ($t(29) = 7.799$, $p < 0.001$, $BF_{alt} > 999$). Although the conscious VE was numerically larger than the non-conscious effect, the two were not significantly different ($t(29) = -1.747$, $p = 0.091$), however, the strength of evidence for the null was inconclusive ($BF_{null} = 1.335$). Together these results suggest that subjective awareness of a visual stimulus is not needed for multi-sensory integration and, under conditions of meta-contrast masking, the unconscious VE can be nearly as large as the conscious one.

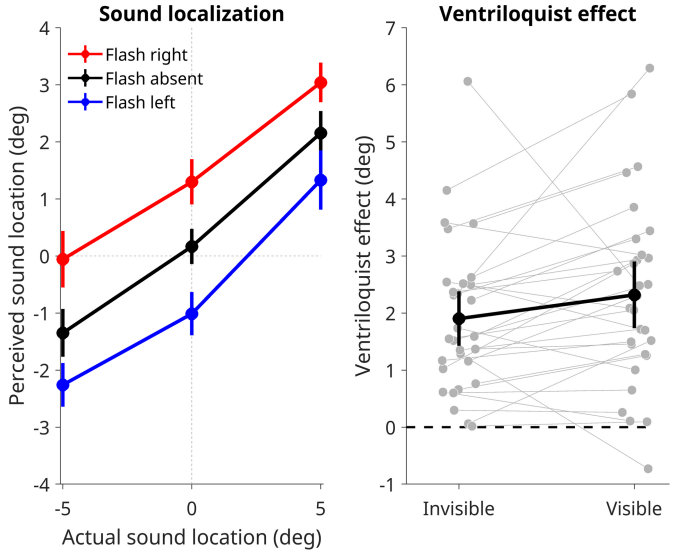

Figure 3. *Ventriloquist effect (VE) for visible and invisible trials. Left panel shows the mean reported sound location (left = -5, center = 0, right = +5) as a function of the actual (simulated) location across all flash absent (black), flash right (red), and flash left (blue) trials. Reports generally followed the true sound location but the rightward and leftward shift of sound localization on right and left flash trials respectively indicates a clear VE, which was quantified as the difference in mean sound localization report on right and left flash trials. Error bars are 95% confidence intervals. Right panel shows the VE index separately for trials where the flash was rated invisible (PAS = 0) and visible (PAS > 0). Error bars are 95% confidence intervals and do not overlap zero in either visible or invisible conditions, indicating a significant VE for subjectively unseen stimuli.*

Behavioral Control analyses

While at the group level participants were not better than chance at localizing the visual stimulus on invisible trials, it is always possible that certain individual subjects were significantly above chance and are driving the non-conscious VE. We addressed this in two ways. First, if this were the case one may expect a positive correlation between an individuals localization accuracy on invisible trials (here, collapsing across pre- and post-task blocks) and the size of their VE on invisible trials. However, a regression analysis revealed no such effect (slope = 0.018, SEM = 0.038, p = .644; see Figure 4) which a Bayesian correlation indicates is substantial evidence for the null (BF_{null} = 6.36). Moreover, as seen in Figure 4, the 95% confidence interval on the regression line at chance performance on the x-axis does not include a VE effect of zero, indicating that even for objective flash unconsciousness (p(correct) = 0.5), the predicted level of performance in our data is a non-zero VE of around 1.88 deg.

As an additional control, we sought to screen subjects who may have been above or below chance levels of performance when they rated the flash as invisible. To do this we ran a Bayesian binomial test on each participant's accuracy on invisible trials to see if there was evidence for or against the null hypothesis of the data coming from a binomial distribution with success probability of 0.5 (i.e., 50% accuracy). Only participants with a BF substantially favoring the null hypothesis of chance performance (BF_{null} > 3.33) were retained and the main VE analysis was repeated on this subset. We found 21 of our 30 participants had evidence in favor of the null hypothesis of chance performance and that using this subset produced a clear but slightly reduced invisible VE of 1.52 deg. (0.19) that remained statistically significant ($t(20) = 7.872$, $p < 0.001$) and remained decisive evidence for a non-conscious VE effect (BF_{alt} > 999).

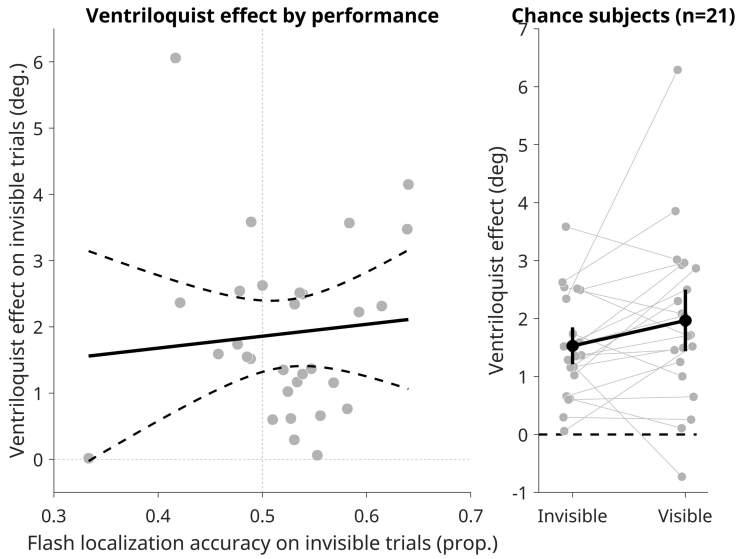

Figure 4. *Non-conscious VE does not depend on task performance. The left panel shows each participant's VE effect on invisible trials as a function of their flash localization accuracy on invisible trials. A lack of clear correlation and a significant intercept term at p(correct) = 0.5 on the x-axis suggests that the non-conscious VE was not driven by individuals with higher task performance. Dashed lines represent 95% confidence intervals on the regression line. The right panel shows the VE for subjectively invisible and visible trials in a subset (n = 21) of participants whose flash localization accuracy on invisible trials indicated substantial support (BF_{null} > 3.33) for the null hypothesis of chance performance. Together these analyses suggest that the non-conscious ventriloquist effect holds even for objectively chance levels of performance. Error bar are 95% confidence intervals.*

EEG decoding of flash location

We decoded the left/right location of the flash from single-trial EEG signals using multivariate LDA in 23 subjects. We found significant flash location decoding when considering all trials together (Figure 5, left) for a period extending 80 to 520 ms post-stimulus (cluster $p < 0.001$), suggesting location information was present for an extended period of time despite masking. When separately examining decoder performance on visible and invisible trials, we found that invisible flash location was significantly decodable from 63 to 224 ms (cluster $p = 0.008$), indicating preserved early, putatively feed-forward processing during unconscious masked trials. On visible trials, however, decoding was sustained longer with one cluster spanning a range from 112 to 220 ms (cluster $p = 0.007$) and another from 232 to 320 ms (cluster $p = 0.007$). Lastly, visible trials were associated with significantly higher decoding accuracy from 304 to 318 ms (cluster $p = 0.034$).

Together these findings indicate that invisible trials were initially processed comparably to visible ones, but were not processed to the point of global "broadcast", thought to occur around 250 – 300 ms (Del Cul et al., 2007; Salti et al., 2015; Sergent et al., 2021). This suggests that any multisensory integration occurring on invisible trials likely took place before 250 ms (i.e., before conscious broadcast according to GNWT).

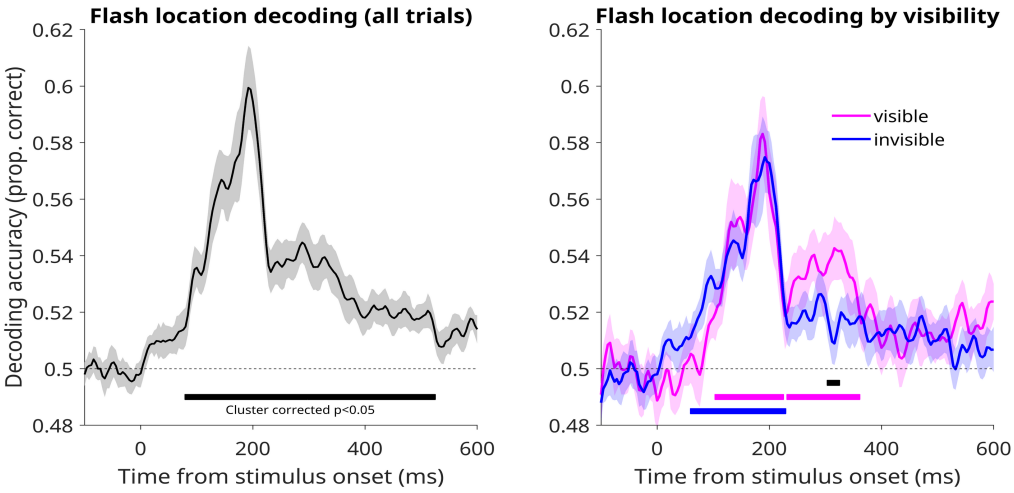

Figure 5*. EEG decoding of visual stimulus location (n = 23). Left panel shows EEG decoding accuracy over time for a classifier trained to predict the location (left/right) of the flash stimulus using all trials and all EEG channels. Solid black squares at the bottom of the plot indicate timepoints of significant above-chance decoding after cluster correction for multiple comparisons. Right panel shows decoding accuracy separated by subjectively visible and invisible trials with magenta squares below representing significant (cluster corrected) timepoints for visible trials, blue squares for invisible trials, and black squares for the difference between visible and invisible. Initial decoding accuracy was comparable for visible and invisible trials but only visible trial decoding persisted beyond ~230 ms and significantly differed from invisible trials around 300 ms. Results are consistent with a preserved early, feed-forward sweep of processing during invisible trials but a lack of sustained "broadcast" of non-conscious information. Shaded regions denote the SEM.*

Discussion

We investigated the role of conscious awareness in a well-studied multisensory integration phenomenon known as the Ventriloquist effect (VE). We found that visual stimuli rendered non-conscious via metacontrast masking nevertheless biased auditory perception towards them. Visual stimuli could not be localized better than chance levels yet our results show that they were still integrated into auditory perception, revealing a clear non-conscious VE. Moreover, we found that EEG decoding of invisible flash location showed an initial representation comparable to that of seen stimuli, but was not sustained beyond ~230 ms, indicating that the non-

conscious integration likely occurred prior to 230 ms, before global broadcast is thought to occur acording GNWT.

Our results add to a growing body of findings suggesting that integration of non-conscious inputs is more pronounced than once believed, particularly in the domain of multisensory integration. Indeed, recent findings indicate that associative learning can occur between two non-conscious stimuli in different sensory modalities, as measured by RT priming effects (Scott et al., 2018). Another study found semantic priming effects on RT for masked audio-visual primes (Faivre et al., 2014). Our results critically extent these findings, however, by going beyond RT effects to show that an unconscious stimulus in one modality actually alters the conscious perception of a stimulus in another modality. In other words, our findings suggest that non-conscious multisensory integration actually shapes the contents of conscious experience and not only produces priming effects in RT tasks. In this vein, our findings are concordant with two recent studies showing similar effects. First, the (Delong et al., 2018) study that directly motivated ours found a similar VE to invisible flashes using dCFS as a means of suppressing flash awareness. Additionally, a recent study found that motion after effects caused by adaptation to a continuous motion signal could be boosted by a congruent auditory motion stimulus and that this effect was observable even when the adapting visual motion stimulus was rendered non-conscious using CFS (Park et al., 2024). Thus, while both studies come to the same general conclusions as ours, it should be noted that CFS has received criticism as a tool for studying unconscious processing due to the possibility (and some evidence) that observers may briefly be aware of low-level features of the suppressed stimuli but simply confuse them as part of the mask, which typically varies across trials along many dimensions of color, shape, size, luminance, etc (Gelbard-Sagiv et al., 2016). Thus, our results add important evidence in favor of non-conscious multisensory integration using a different masking approach less susceptible to those concerns. Additionally, the decoding analysis of concurrent EEG signals helps to pinpoint the general time frame of non-conscious multisensory integration in the VE, which our data indicate occurs within ~230 ms. Although we could not measure a neural signature of integration directly due to poor decoding performance of the sound location (see Methods), this would be an important step for future research.

It is also worth noting that the magnitude of the non-conscious VE effect we observed was rather large, comparable even to that of the visible flash trials. This departs from the finding of Delong et al. (2018) who observed an invisible VE of ~20% of the visible one, compared to about 82% in the current study. One possible explanation for the large effect observed here is that metacontrast masking may preserve more unconscious processing compared to CFS, which may suppress information processing at an earlier stage. Indeed, previous work has found a strong advantage for meta-contrast masking over CFS in affording unconscious perceptual processing, suggesting that effects observed during CFS may result from partial awareness, rather than unconscious processing (Peremen & Lamy, 2014).

Although our study was inspired by predictions made by the GNWT of consciousness, our results provide constrains on any approach to consciousness that emphasizes integrative functions of conscious perception. With respect to GNWT, however, we can see several ways the theory might adapt in light of these findings. It could be that some integration is possible unconsciously but not all forms of integration. Specifically, it could be the case that relatively low-level (i.e., spatial or temporal congruency-based) forms of integration, such as in the VE, can occur either subcortically or via direct auditory-visual cortico-cortico projections (Ghazanfar & Schroeder, 2006). However, perhaps some forms of integration, such as those based on semantic congruency could require higher-level processing only possible if stimuli are consciously perceived and activate global workspace neurons. This is indeed suggested by recent findings showing that AV integration based on semantic cues does not seem to occur when stimuli are presented non-consciously (Delong & Noppeney, 2021). Similar arguments

have been made on the basis of the McGurk effect seeming to depend on conscious perception of the visual component of the stimulus (Palmer & Ramsey, 2012). However, this finding has also recently been challenged (Ching et al., 2019). Thus, it might be the case that certain forms of multi-sensory integration require conscious perception, it seems clear that not all forms do. Our findings begin to constrain the nature and timing of multisensory interactions in the absence of awareness, putting pressure on integration-based theories of consciousness to better specify the role of consciousness in multisensory processing.

References

Alais, D., & Burr, D. (2004). The Ventriloquist Effect Results from Near-Optimal Bimodal Integration. *Current Biology*, *14*(3), 257–262. https://doi.org/10.1016/j.cub.2004.01.029

Baars, B. J. (2002). The conscious access hypothesis: Origins and recent evidence. *Trends in Cognitive Sciences*, *6*(1), 47–52. https://doi.org/10.1016/S1364-6613(00)01819-2

Baars, B. J. (2005). Global workspace theory of consciousness: Toward a cognitive neuroscience of human experience. *Progress in Brain Research*, *150*, 45–53. https://doi.org/10.1016/S0079-6123(05)50004-9

Block, N. (1995). On a confusion about a function of consciousness. *Behavioral and Brain Sciences*, *18*(2), 227–247. https://doi.org/10.1017/S0140525X00038188

Ching, A. S. M., Kim, J., & Davis, C. (2019). Auditory–visual integration during nonconscious perception. *Cortex*, *117*, 1–15. https://doi.org/10.1016/j.cortex.2019.02.014

Cohen, M. A., & Dennett, D. C. (2011). Consciousness cannot be separated from function. *Trends in Cognitive Sciences*, *15*(8), 358–364.

Dehaene, S., & Changeux, J.-P. (2011). Experimental and Theoretical Approaches to Conscious Processing. *Neuron*, *70*(2), 200–227. https://doi.org/10.1016/j.neuron.2011.03.018

Dehaene, S., Charles, L., King, J.-R., & Marti, S. (2014). Toward a computational theory of conscious processing. *Current Opinion in Neurobiology*, *25*, 76–84.

Dehaene, S., & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: Basic evidence and a workspace framework. *Cognition*, *79*(1–2), 1–37. https://doi.org/10.1016/S0010-0277(00)00123-2

Dehaene, S., Naccache, L., Cohen, L., Bihan, D. L., Mangin, J. F., Poline, J. B., & Rivière, D. (2001). Cerebral mechanisms of word masking and unconscious repetition priming. *Nature Neuroscience*, *4*(7), 752–758. https://doi.org/10.1038/89551

Del Cul, A., Baillet, S., & Dehaene, S. (2007). Brain Dynamics Underlying the Nonlinear Threshold for Access to Consciousness. *PLOS Biol*, *5*(10), e260. https://doi.org/10.1371/journal.pbio.0050260

Delong, P., Aller, M., Giani, A. S., Rohe, T., Conrad, V., Watanabe, M., & Noppeney, U. (2018). Invisible Flashes Alter Perceived Sound Location. *Scientific Reports*, *8*(1), 1. https://doi.org/10.1038/s41598-018-30773-3

Delong, P., & Noppeney, U. (2021). Semantic and spatial congruency mould audiovisual integration depending on perceptual awareness. *Scientific Reports*, *11*(1), 10832. https://doi.org/10.1038/s41598-021-90183-w

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Faivre, N., Mudrik, L., Schwartz, N., & Koch, C. (2014). Multisensory integration in complete unawareness: Evidence from audiovisual congruency priming. *Psychological Science*, *25*(11), 2006–2016. https://doi.org/10.1177/0956797614547916

Gardner, W. G., & Martin, K. D. (1995). HRTF measurements of a KEMAR. *The Journal of the Acoustical Society of America*, *97*(6), 3907–3908. https://doi.org/10.1121/1.412407

Gelbard-Sagiv, H., Faivre, N., Mudrik, L., & Koch, C. (2016). Low-level awareness accompanies “unconscious” high-level processing during continuous flash suppression. *Journal of Vision*, *16*(1), 3–3. https://doi.org/10.1167/16.1.3

Ghazanfar, A. A., & Schroeder, C. E. (2006). Is neocortex essentially multisensory? *Trends in Cognitive Sciences*, *10*(6), 278–285. https://doi.org/10.1016/j.tics.2006.04.008

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What’s new in psychtoolbox-3. *Perception*, *36*(14), 1–16.

Lau, H., & Rosenthal, D. (2011). Empirical support for higher-order theories of conscious awareness. *Trends in Cognitive Sciences*, *15*(8), 365–373. https://doi.org/10.1016/j.tics.2011.05.009

Ludwig, D. (2023). The functions of consciousness in visual processing. *Neuroscience of Consciousness*, *2023*(1), niac018. https://doi.org/10.1093/nc/niac018

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, *164*(1), 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024

Mashour, G. A., Roelfsema, P., Changeux, J.-P., & Dehaene, S. (2020). Conscious Processing and the Global Neuronal Workspace Hypothesis. *Neuron*, *105*(5), 776–798. https://doi.org/10.1016/j.neuron.2020.01.026

Merikle, P. M., Smilek, D., & Eastwood, J. D. (2001). Perception without awareness: Perspectives from cognitive psychology. *Cognition*, *79*(1), 115–134.

Morsella, E., Godwin, C. A., Jantz, T. K., Krieger, S. C., & Gazzaley, A. (2016). Homing in on consciousness in the nervous system: An action-based synthesis. *Behavioral and Brain Sciences*, *39*, e168. https://doi.org/10.1017/S0140525X15000643

Mudrik, L., Faivre, N., & Koch, C. (2014). Information integration without awareness. *Trends in Cognitive Sciences*, *18*(9), 488–496. https://doi.org/10.1016/j.tics.2014.04.009

Noel, J.-P., Ishizawa, Y., Patel, S. R., Eskandar, E. N., & Wallace, M. T. (2019). Leveraging Nonhuman Primate Multisensory Neurons and Circuits in Assessing Consciousness Theory. *Journal of Neuroscience*, *39*(38), 7485–7500. https://doi.org/10.1523/JNEUROSCI.0934-19.2019

Palmer, T. D., & Ramsey, A. K. (2012). The function of consciousness in multisensory integration. *Cognition*, *125*(3), 353–364. https://doi.org/10.1016/j.cognition.2012.08.003

Park, M., Blake, R., & Kim, C.-Y. (2024). Audiovisual interactions outside of visual awareness during motion adaptation. *Neuroscience of Consciousness*, *2024*(1). https://doi.org/10.1093/nc/niad027

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*(4), 437–442.

Peremen, Z., & Lamy, D. (2014). Comparing unconscious processing during continuous flash suppression and meta-contrast masking just under the limen of consciousness. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00969

Sadaghiani, S., Hesselmann, G., & Kleinschmidt, A. (2009). Distributed and Antagonistic Contributions of Ongoing Activity Fluctuations to Auditory Stimulus Detection. *Journal of Neuroscience*, *29*(42), 13410–13417. https://doi.org/10.1523/JNEUROSCI.2592-09.2009

Salti, M., Monto, S., Charles, L., King, J.-R., Parkkonen, L., & Dehaene, S. (2015). Distinct cortical codes and temporal dynamics for conscious and unconscious percepts. *ELife*, *4*, e05652. https://doi.org/10.7554/eLife.05652

Samaha, J. (2015). How best to study the function of consciousness? *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00604

Sandberg, K., Timmermans, B., Overgaard, M., & Cleeremans, A. (2010). Measuring consciousness: Is one measure better than the other? *Consciousness and Cognition*, *19*(4), 1069–1078.

Scott, R. B., Samaha, J., Chrisley, R., & Dienes, Z. (2018). Prevailing theories of consciousness are challenged by novel cross-modal associations acquired between subliminal stimuli. *Cognition*, *175*, 169–185. https://doi.org/10.1016/j.cognition.2018.02.008

Sergent, C., Baillet, S., & Dehaene, S. (2005). Timing of the brain events underlying access to consciousness during the attentional blink. *Nature Neuroscience*, *8*(10), 1391–1400. https://doi.org/10.1038/nn1549

Sergent, C., Corazzol, M., Labouret, G., Stockart, F., Wexler, M., King, J.-R., Meyniel, F., & Pressnitzer, D. (2021). Bifurcation in brain dynamics reveals a signature of conscious processing independent of report. *Nature Communications*, *12*(1), 1. https://doi.org/10.1038/s41467-021-21393-z

Seth, A. K., & Bayne, T. (2022). Theories of consciousness. *Nature Reviews Neuroscience*, *23*(7), 7. https://doi.org/10.1038/s41583-022-00587-4

Treder, M. S. (2020). MVPA-Light: A Classification and Regression Toolbox for Multi-Dimensional Data. *Frontiers in Neuroscience*, *14*. https://doi.org/10.3389/fnins.2020.00289

van Vugt, B., Dagnino, B., Vartak, D., Safaai, H., Panzeri, S., Dehaene, S., & Roelfsema, P. R. (2018). The threshold for conscious report: Signal loss and response bias in visual and frontal cortex. *Science*, eaar7186. https://doi.org/10.1126/science.aar7186

Vroomen, J., & De Gelder, B. (2004). Perceptual Effects of Cross-Modal Stimulation: Ventriloquism and the Freezing Phenomenon. In *The handbook of multisensory processes* (pp. 141–150). Boston Review. https://doi.org/10.7551/mitpress/3422.003.0012