Visual limitations shape audio-visual integration

Alexis Pérez-Bellido

Department of Basic Psychology, University of Barcelona, Barcelona, Catalonia, Spain
Department of Neuroscience, Baylor College of Medicine, Houston, TX, USA

✉

Marc O. Ernst

Cognitive Neuroscience Department and Cognitive Interaction Technology-Center of Excellence, Bielfeld University, Bielefeld, Germany

✉

Salvador Soto-Faraco

Department of Information and Communication Technologies, Pompeu Fabra University, Barcelona, Catalonia, Spain
Catalan Institution for Research and Advanced Studies (ICREA), Barcelona, Catalonia, Spain

✉

Joan López-Moliner

Department of Basic Psychology and Institute for Brain Cognition and Behavior (IR3C), University of Barcelona, Barcelona, Catalonia, Spain

✉

Recent studies have proposed that some cross-modal illusions might be expressed in what were previously thought of as sensory-specific brain areas. Therefore, one interesting question is whether auditory-driven visual illusory percepts respond to manipulations of low-level visual attributes (such as luminance or chromatic contrast) in the same way as their nonillusory analogs. Here, we addressed this question using the double flash illusion (DFI), whereby one brief flash can be perceived as two when combined with two beeps presented in rapid succession. Our results showed that the perception of two illusory flashes depended on luminance contrast, just as the temporal resolution for two real flashes did. Specifically we found that the higher the luminance contrast, the stronger the DFI. Such a pattern seems to contradict what would be predicted from a maximum likelihood estimation perspective, and can be explained by considering that low-level visual stimulus attributes similarly modulate the perception of sound-induced visual phenomena and "real" visual percepts. This finding provides psychophysical support for the involvement of sensory-specific brain areas in the expression of the DFI. On the other hand, the addition of chromatic contrast failed to produce a change in the strength of the DFI despite it improved visual sensitivity to real flashes. The null impact of chromaticity on the cross-modal illusion might suggest a weaker interaction of the parvocellular visual pathway with the auditory system for cross-modal illusions.

Introduction

Human perception and behavior are often driven by the integration of information from multiple sensory inputs. This tendency for integration is so strong that conflicting signals arriving from different sensory modalities often lead to perceptual illusions in the observer (Bertelson & Radeau, 1981; McGurk & MacDonald, 1976; Shams, Kamitani, & Shimojo, 2000). Recent neuroimaging studies have provided mounting evidence that some of these multisensory interactions express in brain areas classically considered unimodal (Driver & Noesselt, 2008; Kayser, Petkov, Augath, & Logothetis, 2005; Lakatos, Chen, O'Connell, Mills, & Schroeder, 2007), in addition to the well-known convergence of sensory information in heteromodal brain areas (subcortical: Meredith & Stein, 1983; Nagy, Eördegh, Paróczy, Márkus, & Benedek, 2006; Wallace, Meredith, & Stein, 1998; or cortical: Barbas et al., 2005; Beauchamp, Lee, Argall, & Martin, 2004; Cohen & Andersen, 2004; Helbig et al.,

2012; Werner & Noppeney, 2010). Indeed, the cross-modal activation of sensory cortices associated with the perception of cross-modal illusions (Mishra, Martinez, Sejnowski, & Hillyard, 2007; Watkins, Shams, Tanaka, Haynes, & Rees, 2006; for a review, Driver & Noesselt, 2008) suggests that illusory and nonillusory percepts experienced in a given sensory modality might be supported by similar neural underpinnings. In the present study we addressed whether sound-induced visual illusions respond to physical stimulus manipulations in a manner similar to real (i.e., nonillusory) visual percepts. For this purpose we utilized one well-known case of auditory driving, named the double flash illusion (DFI; Shams, 2002; Shams et al., 2000; see also Shipley, 1964). In this illusion, a single visual flash accompanied by two auditory beeps presented in rapid sequence is perceived as two flashes. That is, due to the auditory stimulus, two flashes are seen when physically only one is present. According to neuroimaging studies, the DFI involves the activity of subcortical convergence structures like the superior colliculus (SC), cortical association areas like the superior temporal sulcus (STS; Watkins et al., 2006), and, critically, sensory cortical areas like the primary visual cortex (V1; Mishra et al., 2007). Thus, one might expect that if primary visual cortex is causally involved in cross-modal induced visual illusions like the DFI, then the DFI should be sensitive to the stimulus manipulations that are known to modulate primary visual cortex. In the present study we focus on the temporal limitations of the perception of fast visual events, which strongly depend on processing in early sensory visual areas (Jiang, Zhou, & He, 2007; Mullen, Thompson, & Hess, 2010; Rager & Singer, 1998; Ress & Heeger, 2003; Simonson & Brožek, 1952). Specifically we assess whether the visual perceptual limitations induced by luminance contrast and chromaticity manipulations also apply to the perception of illusory events arising from cross-modal interactions. It is well known that luminance contrast affects visual temporal resolution, that is, the ability to perceptually resolve two flashes presented in a rapid sequence. For instance, increasing contrast leads to an improvement in the temporal resolution for flicker detection (De Lange, 1958; Tyler & Hamer, 1990). Similarly, additional chromatic information also leads to improvements in visual temporal resolution compared to same luminance contrast achromatic stimuli (Pokorny & Smith, 1997; Sun, Pokorny, & Smith, 2001; Swanson, Ueno, Smith, & Pokorny, 1987). Thus, we capitalize on these prior findings to evaluate if the ability to resolve two real flashes corresponds with the perception of two illusory flashes across manipulations in luminance and chromaticity.

In order to guide our analysis, we must consider two nonmutually exclusive ways in which luminance and chromatic contrast manipulations might affect visual temporal resolution and the strength of the auditory-induced visual illusion. On one hand, decreases in luminance/chromatic contrast might reduce the precision of sensory estimates. From the maximum likelihood estimation (MLE) framework (see Ernst & Banks, 2002; Ernst & Bülthoff, 2004), one can posit that changes in the relative precision of each sensory modality will affect the weight of each sensory input on the outcome of multisensory integration. Thus, we predicted that reductions in the precision of the visual estimate due to contrast manipulation should lead to a corresponding reduction in the relative weight of the visual estimate and, therefore, to an increase in the strength of the DFI. Based on this prediction, one would expect a negative correlation between visual temporal resolution and the strength of the DFI—that is, the lower the ability to discriminate one from two real flashes, the stronger the illusion. On the other hand, decreases in luminance/chromatic contrast might also affect the accuracy of sensory estimates. That is, low contrasts may degrade visual temporal resolution because observers might be perceptually biased toward seeing one instead of two flashes. For instance, according to Bloch's law of vision (Bloch, 1885), the capacity to resolve two consecutive visual stimuli in time depends on the contrast of the stimuli and their duration. When presenting two flashes at low levels of contrast, the luminance energy of the first flash might not be sufficient to reach the threshold required to induce an individuated visual percept before the second flash is presented. As a consequence, the observer integrates the second visual event over time and perceives only one flash. This results in an inaccurate (i.e., biased) visual percept. We hypothesize that the same mechanism that biases real visual perception can operate on illusory visual perception. Therefore, an observer discriminating two illusory flashes at low levels of contrast might be more biased to perceive one flash than an observer discriminating at higher levels of contrast. This should lead to fewer two-flash reports at low levels of contrast (i.e., a weaker DFI). Consequently, visual accuracy enhancements (i.e., reductions in bias) should positively correlate with the DFI.

Given the aforementioned, it is important to note that changes in precision and in accuracy to temporally resolve visual events lead to effects in opposite directions regarding the DFI strength. However, these two effects can co-occur, as visual precision and accuracy might both change together upon the manipulation of stimulus parameters like luminance or chromaticity. Therefore, for completeness, we must also consider a third, less likely possibility of a null correlation between the DFI and visual temporal resolution. This would only occur if luminance contrast or chromaticity do not have an effect on the DFI or if

there are concomitant reductions in precision and accuracy that influence the DFI with equivalent strength but in opposite directions. In this case, one would not be able to determine if the lack of correlation between visual performance for real flashes and DFI is the result of accuracy and precision modulations cancelling each other out or if it is the result of the absence of modulation by contrast (i.e., null effect) in the strength of the illusion. These different predictions are tested in two experiments. In Experiment 1 we modulate the luminance contrast and in Experiment 2 we add chromatic information.

In parallel, our design will allow us to examine how changes in luminance contrast and chromaticity affect the flash fusion illusion (FI; Andersen, Tiippana, & Sams, 2004), whereby two visual flashes accompanied by one auditory beep are often perceived as one flash. Nevertheless, we will not use the FI to test our hypotheses. This is because we expect the same result on the FI independent of contrast manipulations affecting the precision or the accuracy: If increments in the level of contrast predominantly enhance visual precision, then, according to the MLE predictions, the level of contrast should be negatively correlated with the FI prevalence (i.e., more visual precision, less FI). Complementarily, if increments in the level of contrast predominantly affect the visual accuracy (e.g., biasing the participants to report one flash at low levels of contrast), the FI should also correlate negatively with visual luminance contrast (i.e., again better visual accuracy, less FI). These two indistinguishable patterns of interaction make the FI a less useful phenomenon to evaluate our main experimental hypothesis.

General methods

Participants

Eleven naive participants (eight female) took part in Experiment 1 (mean age $\pm$ SD = 25 $\pm$ 3) and an additional 12 naive participants (seven female) took part in Experiment 2 (mean age $\pm$ SD = 26 $\pm$ 4). None of them reported any visual or auditory deficits. The experimental protocols were approved by the local ethics committee (CEIC Parc de Mar).

Stimuli

Visual stimuli were displayed on a 21-inch Philips CRT monitor (120 Hz; Philips, Amsterdam, The Netherlands). Participants were asked to maintain their gaze throughout each trial on a 0.7° red fixation square (42.5 cd/m^2 luminance) presented in the center of the monitor on a gray background (39.7 cd/m^2). Just below the fixation square a 5.5° $\times$ 4.5° dark rectangle (0.55 cd/m^2) for the achromatic condition or a red (13.6 cd/m^2) rectangle for the chromatic condition served as the pedestal for stimulus presentation (see Figure 1a). The visual targets consisted of briefly flashed (33 ms, four frames) squares (2.29° side) presented in the center of the pedestal square, 2° below fixation. This ensured a balanced stimulation of retinal cones and rods (Osterberg, 1935), which are both critical for visual perception in photopic and scotopic conditions. In Experiment 1, the (achromatic) flashes varied exclusively in luminance contrast with respect to the black background rectangle (seven possible values: 0.29, 0.64, 0.80, 0.91, 0.96, 0.98, and 0.99 in Michelson contrast units). In Experiment 2, a green square served as visual target and was also presented at one of seven different contrast values (averaged across participants: −0.93, −0.60, −0.13, −0.05, 0.03, 0.35, and 0.60 in Michelson contrast units; note that the negative values correspond to a luminance below the luminance of the pedestal) with respect to the red background rectangle. The luminance contrast levels in Experiment 2 were adapted in order to gain more information about the illusion around the point of chromatic isoluminance. Isoluminance was assessed for each participant individually prior to the DFI experiment using a flicker fusion paradigm (described in the following material in the procedures section). The remaining luminance levels were then set individually for each person. This resulted in an average point of isoluminance of 12.4 $\pm$ 2 cd/m^2 across participants.

In both experiments, when the two flashes were presented, the stimulus onset asynchrony (SOA) was 66 ms (eight frames), producing a 16.6 Hz brief flicker. This frequency was determined in previous piloting to produce around 75% correct discrimination performance at medium contrast levels (for Experiment 1) and medium chromatic luminance levels (for Experiment 2).

Auditory stimuli in both experiments were 3 kHz tones (48 dB) of 10 ms duration delivered from two speakers located at both sides of the screen and at the same height as the visual stimulus. Thus, participants would perceive the sounds centrally. When two beeps were presented to induce the DFI, the auditory SOA was 60 ms. Audio-visual timing of the stimuli was carefully adjusted using an oscilloscope.

Conditions

The different visual and audio-visual combinations resulted in six experimental conditions, two purely visual and four audio-visual combinations: **V1A0:** one
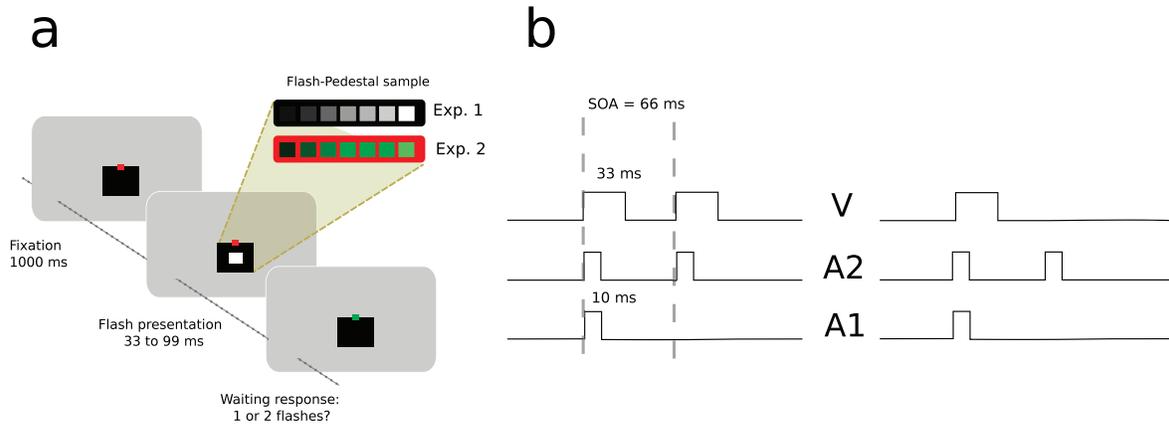
Figure 1. (a) Trial sequence for Experiments 1 and 2; *Inset:* Flash-Pedestal sample shows an approximate representation of the 14 visual stimuli used. (b) Schematic representation of the different possible audio-visual combinations in Experiments 1 and 2 (V = visual, A2 = 2 beeps, A1 = 1 beep conditions; Visual conditions include 2 flash trials (left) and 1 flash trials (right). Stimuli timings are specified for the 2 flash conditions but also apply to the 1 flash condition.

flash alone; **V1A1**: one flash and one beep; **V1A2**: one flash and two beep; **V2A0**: one flash; **V2A1**: two flash and one beep; **V2A2**: two flash and two beep. By introducing the one-beep condition we can test our hypothesis on the flash fusion illusion. We also expect to increment the uncertainty about the number of beeps at the same time that we balance our experimental design (i.e., one or two flash and one or two beep conditions).

Procedures

Participants sat in a darkened and sound-attenuated room with their head resting on a chinrest at 1 meter from the screen. Prior to the beginning of the experiment, they adapted to the room and screen luminance for 5 minutes. While maintaining their gaze on a fixation square, participants were instructed to report the number of perceived flashes (one or two) at the middle of the pedestal square regardless of any concurrent auditory information. They were asked to be as accurate as possible without time constraints. Each trial began with the appearance of the red fixation square and then, after 1000 ms, one or two flashes appeared. In the audio-visual conditions these flashes were presented accompanied by one, two, or no beeps. The first beep (or the only beep in the one beep condition) appeared temporally aligned with the first flash (see Figure 1a for more specific details on the timing).

In Experiment 2, prior to starting the main experimental task, participants were tested with the flicker-fusion procedure (Ives, 1912) to determine their individual point of perceived isoluminance between red and green. The visual display for estimating the point of isoluminance was identical to the one used in the

main experiment (red fixation square, green stimulus on a red pedestal, and a gray background). Participants directed their gaze to the fixation square while attending to a square flickering between green and red at 25 Hz at the center of the red pedestal rectangle. They were asked to adapt the luminance of the green target with the mouse buttons (in steps of 0.04) in order to minimize the sensation of flicker as much as possible. Once participants understood and had practiced the task, they performed five flicker-fusion trials from which we extracted the average isoluminance value for each participant.

Prior to starting any experiment, participants were warned that they would see stimuli of varying intensity, but the intensity was not informative about the number of flashes. After completing 30 training trials participants began with the experiment. Every observer took a short break of approximately 3 minutes between blocks. In both experiments, participants performed two experimental blocks completing 588 trials in total (18 minutes each block). The different possible audio-visual combinations and the level of contrast conditions appeared randomized across trials. Each experiment contained 14 trials per condition.

Data analysis

In order to assess the perceptual impact of luminance contrast and chromatic change on the DFI, we analyzed the present data using the signal detection theory (SDT) framework (Macmillan & Creelman, 2004). Previous studies using the DFI paradigm have successfully applied SDT analysis (Rosenthal, Shimojo, & Shams, 2009; Watkins et al., 2006; Wozny, Beierholm, & Shams, 2008) to demonstrate the sensory
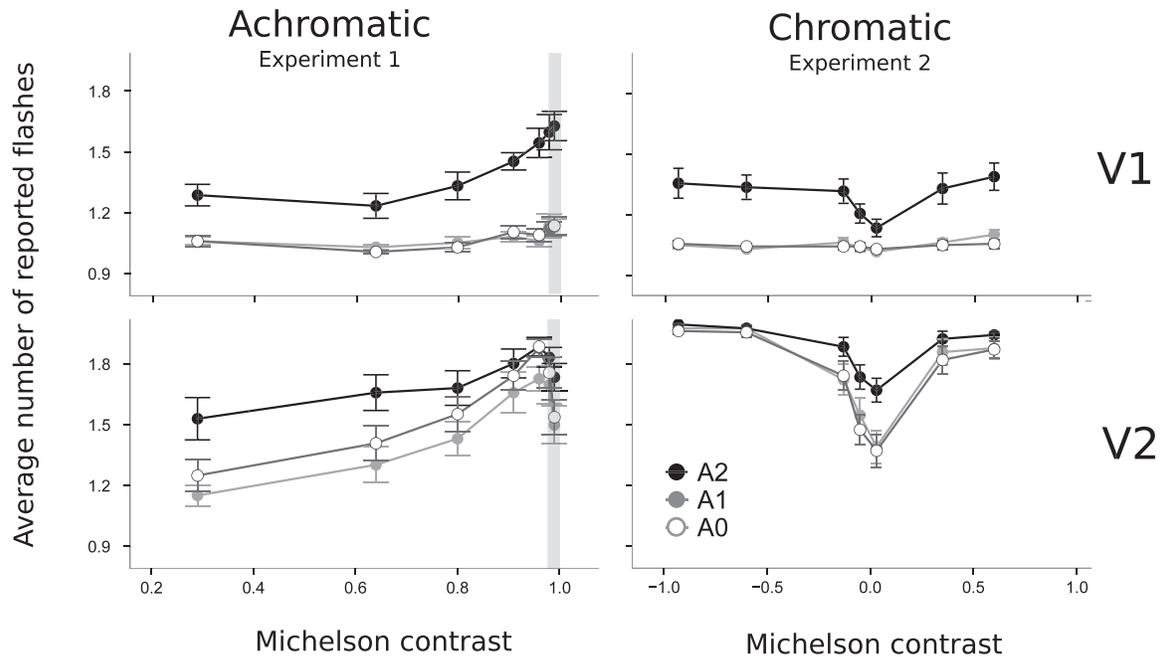
Figure 2. Proportion data: Average number of reported flashes at each luminance contrast level (Michelson units) in achromatic (Experiment 1) and chromatic (Experiment 2) conditions. The top and bottom panels show the data from the one (V1) and two (V2) flash conditions, respectively. Each plot depicts the number of flashes reported as a function of 0, 1, or 2 auditory beeps (gray scale coded; see legend). The shaded area in the achromatic experiment (left column) covers those luminance contrast values where visual performance is impaired by a flash-blindness phenomenon (see text for details). Error bars represent *SEM*.

nature of the sound-induced illusions. First of all, we specify that the sensitivity parameter (d') is determined by both the observers' visual accuracy and precision, and it describes the capacity to discriminate signal + noise (two flash) from noise (one flash). The criterion parameter (c) represents the observer bias. It can be determined by his/her decisional threshold to classify the stimuli as signal or noise. Furthermore, as the criterion parameter is also sensitive to perceptual biases, it is a relevant measure to quantify the DFI (Witt, Taylor, Sugovic, & Wixted, 2015).

In order to calculate the d' and c parameters we applied the following procedure: First, for each level of contrast we extracted the proportion of hits ("two flashes" correctly reported in the two flash conditions V2A0, V2A1, V2A2) and false alarms (incorrectly reported "two flashes" in the one flash conditions V1A0, V1A1, V1A2). For completeness, please note that reporting correctly "one flash" in a one flash condition is a correct rejection and reporting incorrectly only "one flash" in a two flash condition is a miss. Then we computed the d' and the c parameters to reflect performance in resolving two flashes for each observer and contrast level, under each different auditory condition separately (A0 = no sound, A1 = one beep or A2 = two beeps). That is, visual only performance (A0) is calculated using hits from the V2A0 condition and false alarms from the V1A0

condition. DFI is calculated using hits from the V2A2 condition and false alarms from the V1A2 condition. Strong DFI should result in the inability to discriminate between these two conditions (V2A2 vs. V1A2), and therefore, in low d'. Finally, the visual performance under single beep conditions (hits from V2A1 and false alarms from V1A1) will also be calculated. This condition should reflect the strength of fusion illusions (taking two flashes as one). In all cases, the computation of the parameters follows the SDT framework [$d' = z(hits) - z(false\ alarms)$; $c = -0.5*(z(hits) + z(false\ alarms))$]. Hit and false alarm rates of 0 and 1 were approximated by $1/N$ and $1 - (1/N)$ respectively to allow Z score conversion, where N is the number of trials tested for that particular condition and participant (Rosenthal et al., 2009). For congruence with previous studies and for simplicity, we assumed that the distributions in the one flash and two flash conditions have equal variances (although this is a simplification).

To test our hypotheses, we fit linear mixed-effects models (LMM) to the d' and c data. These models address the limitations of more traditional approaches to random effects and provide valuable information about the linear relationship between the different factors. This approach allowed us to easily interpret the effect of contrast in both experiments even though they included slightly different levels of luminance. Prior to conducting the LMM we translated (i.e., adding the

smallest constant to avoid negative numbers) and log-log transformed our data (Figure 3) to obtain linear trends and improve the quality of the fit. To make sure this transformation did not violate the assumption for the LMM, we checked for the model's assumptions via the criticism plots (e.g., model residuals, q-q plot and standardized residuals) (Baayen, 2008). The log-log transformation did not affect the shape of the residuals, so we did not have to fit generalized LMM models besides the one assuming the Gaussian residuals. We conducted all the analyses using the *lmer* function implemented in the *lme4* R package (v.1.0-6) (Bates, Mächler, Bolker, & Walker, 2012).

In the model, the participant factor was fitted as a random effect (varying the intercept and the slope) to control for the potential between-subjects variability within each condition. The number of beeps (factor; A0, A1 and A2), luminance contrast (continuous variable with five levels in Experiment 1 [0.29, 0.64, 0.80, 0.91, 0.96] and seven levels in Experiment 2, [−0.93, −0.60, −0.13, −0.05, 0.03, 0.35, 0.60], values expressed in Michelson contrast units), and chromaticity (factor; achromatic flashes in Experiment 1 and chromatic flashes in Experiment 2) were fixed effects. We used the visual condition (A0) of Experiment 1 as a reference for all the statistical comparisons.

It is important to note that in Experiment 2 we used positive and negative luminance contrast polarities (e.g., positive and negative polarities correspond to flash target luminance above and below the pedestal luminance, respectively). Thus, before regressing visual performance as a function of contrast levels, we ensured that observers' sensitivities were similar for luminance contrast values of equivalent magnitude but opposite in polarity. To do this, we introduced the variable polarity (four negative and three positive polarities) in a LMM. In congruence with McCormick & Mamassian (2008), we did not find any significant effect or interaction involving the polarity variable. Therefore, we decided to report the regression analyses using luminance contrast transformed to absolute units (Figure 4, right panel).

Hypotheses

If the DFI is a perceptual effect, then a single flash in the V1A2 stimuli should be experienced as two flashes, and therefore difficult to tell apart from two real flashes. Hence, because we have associated V1 (1 flash) and V2 (2 flash) to noise and signal + noise distributions respectively, we would thus expect a lower d' in the A2 condition compared to the A0 condition.

Directly addressing the motivating question in the present study, the visual stimulus manipulations in luminance contrast and chromaticity may affect the

DFI in two different ways: On the one hand, if the visual manipulations predominantly affect visual *precision*, the DFI prevalence should decrease with increments in contrast (in agreement with the MLE model). In this case, the d' in the A2 condition will increase relative to the A0 d' with increments in luminance contrast or chromaticity (i.e., a significantly steeper slope in the A2 compared to the A0 condition level). On the other hand, if a visual manipulation predominantly impacts visual *accuracy*, the A2 d' will be reduced relative to A0 d' with increments in contrast or chromaticity (i.e., a significantly shallower slope in the A2 condition). Finally, there is also the possibility that a visual manipulation may not have an impact on the strength of the DFI. In this case the linear model parameter estimates representing the slope of the A2 d' and A0 d' levels as a function of luminance contrast or chromaticity will not be significantly different. That is, both conditions (the condition reflecting DFI, and the condition reflecting visual performance) will behave the same under the visual manipulation.

The interpretation of the criterion parameter is less straightforward in the case of perceptual biases (Witt et al., 2015). This is because criterion is shifted by any source of bias, whether decisional or perceptual in origin. The DFI is an example of sound-induced perceptual bias. Accordingly, previous literature has found c to be lower in the A2 compared to the A0 conditions (Rosenthal et al., 2009; Watkins et al., 2006; Wozny et al., 2008). This represents a sound-induced increment in the proportion of reported flashes. We expect a similar pattern in our A2 condition. Furthermore, if there is a bias to report more flashes at high compared to low contrast levels, we expect a reduction in the c parameter with increments in contrast. This c reduction should be more pronounced in the A2 relative to the A0 condition if the source of bias induced by the DFI increases with contrast. However, it is important to note that in a DFI paradigm, the SDT analysis cannot be used to determine whether shifts in criterion are of a perceptual or decisional nature.

Results

Figure 2 shows the average number of reported flashes as a function of contrast for the different conditions (levels of gray) and the two experiments (different panels). We can observe an overall increase in the number of flashes reported as a function of contrast, but it is also noticeable that at the two highest luminance contrast levels (Experiment 1, V2 conditions), the average number of reported flashes declines sharply (shadowed areas in Figure 2). The most parsimonious explanation for this sharp decline in two
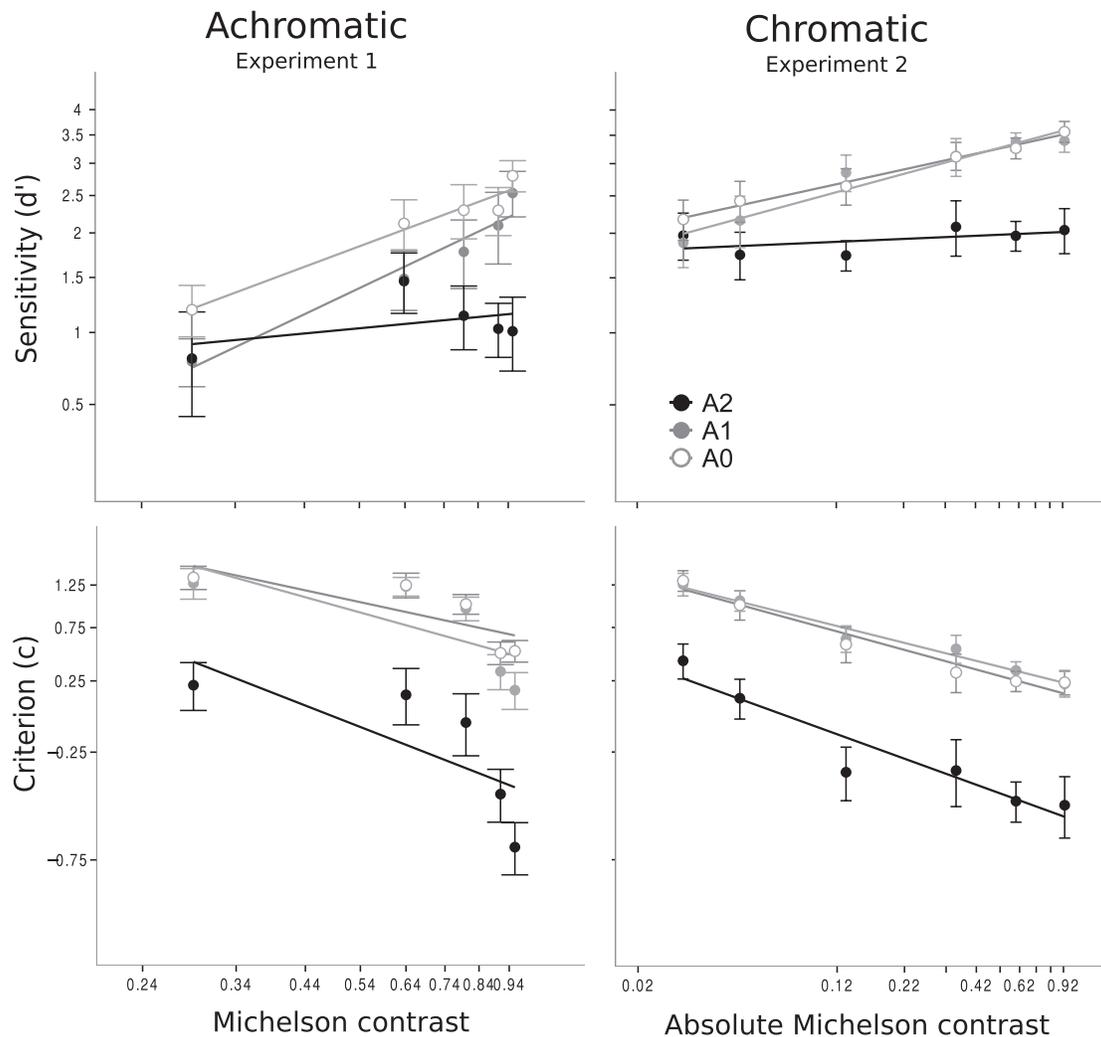
Figure 3. The average *d'* and *c* values (top and bottom panels, respectively) across participants for the achromatic (Experiment 1) and chromatic (Experiment 2) experiments are represented as a function of luminance contrast (log-log transformed scores). The visual (A0) and the one (A1) and two beep (A2) conditions are indexed in different colors. Error bars represent *SEM*. The continuous lines represent the average linear regression fits for each experiment and auditory condition. Luminance contrast units in the chromatic experiment were transformed to absolute contrast units. The −0.6 and 0.6 Michelson contrast levels in the chromatic experiment were averaged into a single data point after this transformation.

flash perception at high contrast is the flash-blindness phenomenon, caused by oversaturation of the retinal pigment (Brown, 1965; Miller, 1965) following the presentation of the first flash at high luminance. Therefore, as this decrement in performance is likely to have a retinal origin that is not present in the lower contrast values, the involved levels of contrast were removed for subsequent analyses based on the SDT framework.

Sensitivity (*d'*) analyses

It is important to note that reductions in the A1 *d'* or A2 *d'* relative to the A0 *d'* condition (baseline) can be interpreted as a signature of the perceptual nature of the FI and the DFI (i.e., the larger the *d'* reduction the stronger the illusion). After fitting the LMM, we conducted an ANOVA with luminance contrast, number of beeps, and chromaticity (i.e., Experiment) as fixed effects and subjects as random effects. Degrees of freedom were calculated by the Satterthwaite's approximation (Satterthwaite, 1946) to obtain corresponding *p*-values. The analysis yielded a significant effect of luminance contrast, $F(1, 51.46) = 31.06$, $p < 0.001$; number of beeps, $F(2, 364) = 36.35$ $p < 0.001$; chromaticity, $F(1, 364) = 14.73$, $p < 0.001$; an interaction between number of beeps × luminance contrast, $F(2, 364) = 4.39$, $p < 0.02$; and an interaction between chromaticity × luminance, $F(1, 364) = 14.73$, $p < 0.001$. The significant interaction between number of

Fixed Effect	Estimate	Std. Error	df	t value	Pr(> | t |)
Intercept (A0)	1.203	0.096	31.800	12.505	0.000***
Luminance	0.456	0.117	365.500	3.891	0.000***
A1	−0.184	0.098	364.000	−1.882	0.061
A2	−0.623	0.098	364.000	−6.372	0.000***
Chromaticity	0.297	0.097	364.000	3.069	0.002**
Luminance: A1	0.027	0.163	364.000	0.164	0.870
Luminance: A2	−0.319	0.163	364.000	−1.952	0.051
Luminance: Chromaticity	−0.312	0.121	364.000	−2.587	0.010*
A1: Chromaticity	0.155	0.137	364.000	1.130	0.259
A2: Chromaticity	0.145	0.137	364.000	1.061	0.290
Luminance: A1: Chromaticity	−0.062	0.170	364.000	−0.361	0.718
Luminance: A2: Chromaticity	0.196	0.170	364.000	1.149	0.251

Table 1. Summary of parameter estimates obtained from the linear mixed model fitted to the *d'* (Significant codes: '***' 0.001 '**' 0.01 '*' 0.05).

beeps × luminance contrast is congruent with a modulation of the double flash audio-visual illusion with luminance contrast. However, the null interaction between number of beeps × chromaticity, $F(2, 364) = 0.80$, p = 0.44, suggests that the addition of chromatic information did not have an impact on the double flash illusion.

In order to gain further insight into the structure of the effects, we report the estimated parameters of the LMM (with their corresponding significances) in Table 1. As we predicted (De Lange, 1958; Tyler & Hamer, 1990), visual sensitivity (measured by *d'*) increases with luminance contrast in the visual condition (A0; Luminance parameter in Table 1) with a slope of 0.456, which is significantly larger than zero (t = 3.891, $p < 0.0001$). Moreover, *d'* in the A2 conditions is −0.623 units significantly smaller (t = −6.372, $p < 0.0001$) on average than in the A0 condition, in congruence with the occurrence of the DFI. Interestingly the slope of *d'* as a function of luminance contrast for the 2 flashes condition (A2) is significantly smaller (t = −1.952, p = 0.051) than the slope for the visual condition (A0) by an amount of 0.319 (luminance contrast: A2 parameter). This reduction in the slope relative to the visual condition is the manifestation of a luminance contrast based modulation of the DFI, and is congruent with the interaction between number of beeps × luminance reported in the ANOVA. The significantly shallower slope can be seen in Figure 3 (upper left panel). Specifically, this result demonstrates that DFI prevalence increments with luminance contrast (i.e., the drop in *d'* for the A2 condition relative to the A0 condition was larger at high compared to low luminance levels). The null interaction between chromaticity × luminance contrast × A2 represents that the same pattern of modulation holds for achromatic (Experiment 1) and chromatic (Experiment 2) conditions. On the other hand, there were no significant differences between the

A1 *d'* and the A0 *d'*. This suggests that the FI was not strong enough in our experiments to drive a significant effect or interaction.

The statistically significant positive intercept estimate for the chromaticity effect (t = 3.069, $p < 0.003$) indicates that the overall sensitivity was larger (by an increment of 0.297) in the chromatic than in the achromatic experiment. This result fits well with previous literature showing that temporal resolution improves by adding chromatic contrast information to available luminance contrast information (as reported by Pokorny & Smith, 1997; Sun et al., 2001; Swanson, Ueno, Smith, & Pokorny, 1987). However, in congruence with the ANOVA, we did not find any interaction between the number of beeps × chromaticity. This result points toward a null contribution of chromatic contrast (chromaticity) to the DFI. Finally, we also found a significant interaction between luminance × chromaticity (t = −2.587, p = 0.02). This interaction can be attributed to a different visual sensitivity gain across levels of luminance for the chromatic and the achromatic conditions.

Criterion (*c*) analyses

We conducted a similar ANOVA on the criterion parameter as a dependent variable with luminance contrast, number of beeps, and chromaticity as fixed effects and subjects as random effects. The analysis yielded a significant effect of luminance contrast, $F(1, 90.91) = 74.20$, $p < 0.001$; number of beeps, $F(2, 374) = 108.51$, $p < 0.001$; chromaticity, $F(1, 374) = 16.67$, $p < 0.001$; and an interaction between chromaticity × luminance, $F(1, 364) = 13.48$, $p < 0.001$. The null interactions between number of beeps × luminance, $F(1, 374) = 2.31$, p = 0.1, and number of beeps × chromaticity, $F(2, 374) = 1.27$, p = 0.28, suggests that

Fixed Effect	Estimate	Std. Error	df	t value	Pr($>$ $\mid$ t $\mid$)
Intercept (A0)	0.644	0.065	78.200	9.941	0.000***
Luminance	−0.353	0.095	376.200	−3.698	0.000***
A1	0.111	0.080	374.000	1.386	0.167
A2	−0.685	0.080	374.000	−8.543	0.000***
Chromaticity	−0.118	0.079	374.000	−1.492	0.137
Luminance: A1	0.086	0.134	374.000	0.639	0.524
Luminance: A2	−0.139	0.134	374.000	−1.037	0.300
Luminance: Chromaticity	0.221	0.099	374.000	2.234	0.026*
A1: Chromaticity	−0.170	0.112	374.000	−1.516	0.130
A2: Chromaticity	−0.036	0.112	374.000	−0.319	0.750
Luminance: A1: Chromaticity	−0.097	0.140	374.000	−0.694	0.488
Luminance: A2: Chromaticity	0.063	0.140	374.000	0.447	0.655

Table 2. Summary of parameter estimates obtained from the linear mixed model fitted to the *c* (Significant codes: '***' 0.001 '**' 0.01 '*' 0.05).

the chromatic and luminance manipulations did not interact with the sound-induced biases.

In order to understand the directionality of the effects reported in the ANOVA, here we will interpret the LMM parameters (Table 2). The criterion (*c*) in the visual (A0) condition decreases significantly with contrast with a slope of −0.353 (t = −3.698, $p < 0.0001$; see Figure 3). That is, the observers show a bias to report two flashes more often at higher levels of luminance. In addition, there is an overall reduction of criterion (by an amount of 0.685) for the A2 conditions with respect to the visual one (t = −8.543, $p < 0.0001$) denoting a significant increment in the bias in reporting two flashes when two beeps are presented (Rosenthal et al., 2009; Watkins et al., 2006; Wozny et al., 2008). In congruence with a null FI effect, criterion was not significantly different in the A0 and the A1 conditions. Finally, the overall criterion was modulated to a lesser extent (smaller slope) in Experiment 2 (chromatic) with respect to Experiment 1 (achromatic). This decrement of the slope (0.221) was significant (t = 2.234, $p = 0.026$).

Visual temporal resolution shapes the double flash illusion

In the previous analyses we demonstrated that visual luminance correlates with the strength of the DFI. To complement these results and directly evaluate whether the observers' visual temporal resolution (given by sensitivity or bias measures) correlates with the DFI strength, we regressed the A2 condition as a function of the A0 condition in a LMM, using the same levels of luminance and chromaticity as in the previous analyses. Hence, the same regression model was adjusted for sensitivity (*d'*) and criterion (*c*) measures (Figure 4; left and right respectively). In this linear regression analysis, a slope smaller than 1 reflects increases in the

DFI strength with enhancements in visual temporal resolution. A slope equal to 1 describes a null modulation of the DFI strength with visual temporal resolution changes, and a slope significantly larger than 1 represents a reduction in the DFI strength with increments in the temporal resolution.

Sensitivity (d') analyses

The linear regression analyses revealed that A2 *d'* increased with A0 *d'* (t = 2.682, $p < 0.01$) for the achromatic and chromatic experiments (Experiment 1 and 2, respectively), with a slope significantly smaller than 1 (averaged between experiments; confidence intervals: 0.06 and 0.45). This result signifies that the higher the observer's visual sensitivity, the larger the DFI (that is, the A2 *d'* reduction, relative to A0). Therefore in compliance with our previous results, we can claim that luminance based changes in visual sensitivity correlate positively with the size of the DFI. Interestingly, the chromatic intercept was significantly higher than the achromatic intercept (t = 2.638, $p < 0.01$). This establishes that visual sensitivity was improved when chromatic information was available, nevertheless, this sensitivity increment was not paralleled by a complimentary increment in the DFI prevalence (i.e., a proportional A2 *d'* reduction in the chromatic conditions). This is congruent with the lack of chromaticity effect on the DFI, reported in our previous results.

Criterion (c) analyses

The A2 *c* was positively correlated with the A0 *c* (t = 7.358, $p < 0.0001$). Moreover, the slopes were significantly smaller than 1 in both the achromatic and the chromatic experiments (averaged confidence intervals: 0.48 and 0.83). This pattern suggests that for similar levels of luminance contrast, the bias in the A2
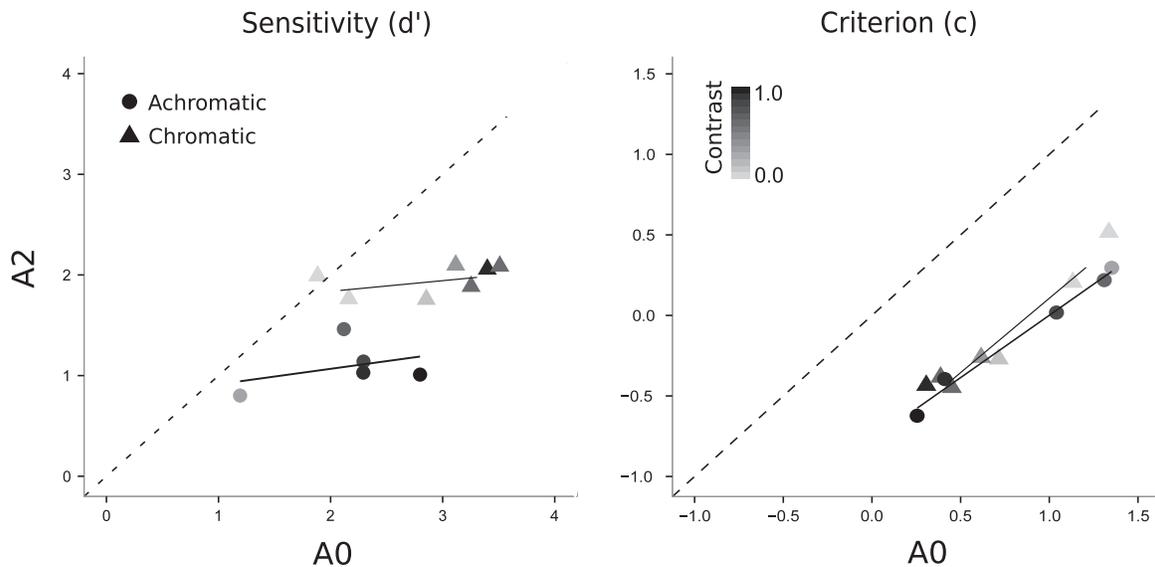
Figure 4. The A2 conditions were linearly regressed as a function of their correspondent A0 values for the *d'* and *c* parameters (left and right panels respectively) for the achromatic (circles; Experiment 1) and chromatic (triangles; Experiment 2) experiments. Contrast level is represented by the white-black gradient (Luminance contrast units in the chromatic experiment are represented in their absolute values). The continuous lines are the averaged linear regression fits for each experiment. The dashed diagonal lines are the unity lines (i.e., same sensitivity/criterion in the A0 and A2 conditions).

condition is greater than the bias in the A0 condition. This is congruent with an enlargement in the DFI prevalence with increments in the A0 bias. The achromatic and chromatic intercepts were 0.64 units smaller on average than 0 (t = −4.894, *p* < 0.0001), representing a general bias to report two flashes in the A2 condition. However, in consonance with our previous results, the intercepts were not significantly different from each other. This demonstrates that, as opposed to luminance contrast, chromaticity did not introduce any type of bias (perceptual or decisional) in our experiments.

Discussion

Using direct measures of performance and SDT analysis in two experiments, we found that the effect of luminance on illusory perception is the same as it is on veridical perception: The higher the luminance contrast (and therefore, visual sensitivity), the stronger the double flash illusion. These results can be framed within our hypothesis by assuming that perceptual biases for real flashes (captured by the *c* reductions with luminance contrast increments) have led to corresponding perceptual biases for (sound-induced) illusory flashes. In fact, this pattern of modulation can be explained by considering that low-level visual stimulus attributes modulate the perception of "real" and sound-induced visual phenomena in a similar way. The exceptions to this pattern are the constraints that arise

from peripheral stages of processing at the retinal level (i.e., flash-blindness phenomenon), which affected the perception of the real, but not the illusory flashes.

These results are consistent with previous studies showing that sound-induced visual illusory percepts share similar perceptual characteristics as real ones. For instance, Berger, Martelli, and Pelli (2003) demonstrated that visual sensitivity to tilt was enhanced by increasing the number of perceived events, independently of whether this was achieved by incrementing the number of real stimuli or by adding beeps (DFI) that consequently induced redundant illusory visual stimuli. Another study using the DFI paradigm (McCormick & Mamassian, 2008) showed that sound-induced illusory flashes interact at a perceptual level with real flashes presented in spatio-temporal concurrence. In a more recent study, Takeshima & Gyoba (2015) showed that the perception of the DFI depended on the spatial frequency (SF) of the visual flashes: the higher the SF the weaker the DFI. It is important to note that visual temporal resolution for low spatial frequencies is better than for high spatial frequencies (Kulikowski, 1971; Lennie, 1980). Therefore, Takeshima and Gyoba's results fit well with the account presented here, that visual limitations (in temporal resolution) shape the perceptibility of the DFI, and potentially extend the applicability of this explanation to a new set of visual manipulations (SF instead of luminance contrast). These studies collectively provide evidence that audio-visual integration is bound by similar sensory limitations that are characteristic of early stages of visual processing.

Our analyses also revealed a trend towards a fusion illusion. Unfortunately, the power of that illusion was not strong enough to drive statistically significant effects or interactions with contrast. This is not completely surprising considering that fusion effects have often been reported as being weaker than "fission," i.e., double flash, effects (Andersen, Tiippana, & Sams, 2004; Innes-Brown & Crewther, 2009; Shams et al., 2000). Future experiments could be designed to test our main hypothesis on the FI (e.g., the more likely two visual stimuli are to fuse, the stronger the sound-induced fusion illusion should be).

Introducing the chromatic contrast manipulation in Experiment 2 allowed us to measure the DFI at isoluminance (or near-isoluminance) levels (where the achromatic flashes of the first experiment would be barely visible). This manipulation induced a general improvement in visual sensitivity in the visual-only condition for comparable luminance values regardless of chromatic information (Benimoff, Schneider, & Hood, 1982; Sun et al., 2001). However, this sensitivity enhancement was not associated with a concurrent increase or reduction in the DFI strength. As mentioned in the Introduction, a symmetric enhancement in visual precision and accuracy induced by adding chromatic contrast could produce equal magnitude but opposite effects on the DFI and thus account for this null result. That is, people might experience less audio-visual interaction (due to an increment in visual precision), while concurrently becoming better at resolving two illusory flashes in time (due to a comparable increment in visual accuracy). However, the nonsignificant chromatic-induced changes in criterion suggest that accuracy does not increase or reduce with the chromatic manipulation, ruling out this hypothesis. In accordance with recent literature (Jaekl, Pérez-Bellido, & Soto-Faraco, 2014), we propose a simpler, more parsimonious explanation: Given previous reports on the potentially distinct role of the magno- and parvo-cellular visual pathways for multisensory integration (Jaekl & Soto-Faraco, 2010; Pérez-Bellido, Soto-Faraco, & López-Moliner, 2013), it is likely that chromatic information, which is primarily processed by the parvocellular pathway (Derrington, Krauskopf, & Lennie, 1984; Lee, Pokorny, Smith, Martin, & Valberg, 1990; Merigan, 1989), might not contribute to the DFI (or it does so to a smaller extent). The DFI seems to be taking place at early processing stages in the visual cortex (Bolognini, Convento, Fusaro, & Vallar, 2013; Watkins et al., 2006). Therefore, the functional and neuroanatomical organization of the early cortical stages of the visual system could have an impact on audio-visual integration. For instance, in the granular layers of V1 (i.e., $4C\alpha$ and $4C\beta$), the information conveyed by the magno- and parvo-cellular subcortical visual pathways remains

segregated (Nassi & Callaway, 2009). The differences in integration of achromatic/chromatic fast visual and auditory information in the DFI might reflect this division. Further research applying neuroimaging techniques could illuminate the neuroanatomical basis of this visual asymmetry in multisensory integration. At present, this hypothesis has received some empirical support and seems to account well for previous multisensory results (Jaekl & Soto-Faraco, 2010; Pérez-Bellido et al., 2013).

Our results showing a parallel perceptual pattern for real and illusory phenomena may seem to contradict what would be predicted from an MLE perspective, which has become a popular framework with which to model multisensory integration, including tasks involving numerosity judgments (e.g., Andersen et al., 2005; Bresciani, Dammeier, & Ernst, 2006; Bresciani et al., 2005). The MLE framework—in which the weight of each sensory signal during multisensory integration depends on its precision—predicts that reductions in visual temporal resolution (given by reductions in the visual precision) should produce an enhancement of the visual illusion (i.e., stronger dominance of sound). However, we found that at low luminance contrast values, where temporal resolution is poor and hence greater auditory influence should be expected, the DFI prevalence was in fact weaker. Yet, the pattern of results produced by luminance contrast manipulations is consistent with MLE when considering the impact of visual accuracy changes on the strength of the DFI. That is, if luminance contrast manipulations predominantly affect visual accuracy (more than precision), they will translate to similar perceptual biases in the outcome of the DFI. Indeed, as the c analyses indicated, luminance contrast manipulations produced a bias to report more flashes with increments in the level of contrast. Thus, the present results are interesting because, to our knowledge, the effect of accuracy reductions on cross-modal integration tasks had not yet been considered. This could be articulated by considering a Bayesian account (see Shams et al., 2005; Ernst, 2006; Wozny et al., 2008) that incorporates a prior describing the "sparseness in numerosity" to embody the present perceptual bias as a function of luminance. That is, in congruence with the c pattern across different levels of contrast, one could speculate about the existence of a visual prior to assume that "one flash is more likely than two flashes." This prior would contribute more to estimates on "number of flashes" for low compared to high contrast levels and might have induced the observed DFI modulation. These kinds of visual priors have been suggested to account for other phenomena in vision. For example, the visual system might assume a prior that leads subjects to perceive moving stimuli experienced at low

luminance contrast levels as being slower or smoother (Stocker & Simoncelli, 2006). In our case, one could also speculate that a manipulation in chromaticity might not interact with a "sparseness in numerosity prior" at all (in congruence with the previously reported null effect of chromaticity in the c parameter). This could provide an explanation for the null contribution of the chromatic information to the DFI. An experimental design oriented towards data modeling and correspondent analysis would help to better elucidate the mechanisms driving these DFI patterns and uncover such hypothetical priors.

Conclusions

The main hypothesis addressed in this study was that if perception of illusory flashes relies on the same underlying visual mechanisms that determine visual temporal resolution for real flashes, the perceptibility of flashes should vary with physical parameters (such as luminance or chromaticity) in a manner similar to real, nonillusory percepts. In this sense, we showed that variations in luminance contrast led to corresponding modulations in performance in real and illusory flashes; if the observers were not able to disentangle two brief flashes at a particular luminance contrast, they were not able to disentangle two illusory flashes induced by similar visual stimulation.

A second interesting finding that arose from the present work is that the addition of chromatic contrast information improves visual temporal resolution, but it does not impact the strength of the DFI in any way. This result is compatible with a functional division between the magno- and parvocellular pathways for audio-visual integration (Jaekl et al., 2014). Moreover, the differential sensitivity of the DFI to luminance and chromatic changes is consistent with multisensory interaction at the early sensory level (e.g., V1), in which the processing of these visual features (luminance and chromatic contrast) might remain partially segregated.

So far, most of the evidence favoring an "early expression" hypothesis for cross-modal integration has been derived from neuroimaging studies showing that activity in sensory specific brain areas correlated with illusory percepts (Bhattacharya, Shams, & Shimojo, 2002; Keil, Müller, Hartmann, & Weisz, 2014; Lange, Oostenveld, & Fries, 2011; Mishra et al., 2007; Watkins et al., 2006, see also Cecere, Rees, & Romei, 2015 for causal evidence and Murray et al., 2015 for a related review). The present study complements these findings by showing comparable performance patterns in a psychophysical task under unisensory and multisensory conditions after multiple visual feature manipulations.

Corresponding author: Alexis Pérez-Bellido.
Email: alexisperezbellido@gmail.com.
Address: Department of Basic Psychology, University of Barcelona, Barcelona, Spain; Department of Information and Communication Technologies, Pompeu Fabra University, Barcelona, Catalonia, Spain; and Department of Neuroscience, Baylor College of Medicine, Houston, Texas.

References

Andersen, T. S., Tiippana, K., & Sams, M. (2004). Factors influencing audiovisual fission and fusion illusions. *Brain Research. Cognitive Brain Research, 21*(3), 301–308, doi:10.1016/j.cogbrainres.2004.06.004.

Andersen, T. S., Tiippana, K., & Sams, M. (2005). Maximum likelihood integration of rapid flashes and beeps. *Neuroscience Letters, 380*(1-2), 155–160, doi:10.1016/j.neulet.2005.01.030.

Baayen, R. H. (2008). *Analyzing linguistic data. A practical introduction to statistics using R*. Cambridge, UK: Cambridge University Press.

Barbas, H., Medalla, M., Alade, O., Suski, J., Zikopoulos, B., & Lera, P. (2005). Relationship of prefrontal connections to inhibitory systems in superior temporal areas in the rhesus monkey. *Cerebral Cortex, 15*(9), 1356–1370, doi:10.1093/cercor/bhi018.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2012). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 51*. Retrieved from http://arxiv.org/abs/1406.5823

Beauchamp, M. S., Lee, K. E., Argall, B. D., & Martin, A. (2004). Integration of auditory and visual information about objects in superior temporal sulcus. *Neuron, 41*(5), 809–823.

Benimoff, N. I., Schneider, S., & Hood, D. C. (1982). Interactions between rod and cone channels above threshold: A test of various models. *Vision Research, 22*(9), 1133–1140, doi:10.1016/0042-6989(82)90078-5.

Berger, T. D., Martelli, M., & Pelli, D. G. (2003). Flicker flutter: Is an illusory event as good as the real thing? *Journal of Vision, 3*(6):1, 406–412, doi:10.1167/3.6.1 [PubMed] [Article].

Bertelson, P., & Radeau, M. (1981). Cross-modal bias and perceptual fusion with auditory-visual spatial discordance. *Perception & Psychophysics, 29*(6), 578–584.

Bhattacharya, J., Shams, L., & Shimojo, S. (2002). Sound-induced illusory flash perception: Role of gamma band responses. *Neuroreport, 13*(14), 1727–1730.

Bloch, M. A.-M. (1885). Experiments in vision [Expériences sur la vision]. *Essai d'Optique sur la gradation de la lumie're Comptes Rendus de Séances de La Société de Biologie, Paris, 37,* 493–495.

Bolognini, N., Convento, S., Fusaro, M., & Vallar, G. (2013). The sound-induced phosphene illusion. *Experimental Brain Research, 231,* 469–478, doi:10.1007/s00221-013-3711-1.

Bresciani, J.-P., Dammeier, F., & Ernst, M. O. (2006). Vision and touch are automatically integrated for the perception of sequences of events. *Journal of Vision, 6*(5):2, 554–564, doi:10.1167/6.5.2. [PubMed] [Article]

Bresciani, J.-P., Ernst, M. O., Drewing, K., Bouyer, G., Maury, V., & Kheddar, A. (2005). Feeling what you hear: Auditory signals can modulate tactile tap perception. *Experimental Brain Research, 162*(2), 172–180, doi:10.1007/s00221-004-2128-2.

Brown, J. L. (1965). Flash blindness. *American Journal of Ophthalmology, 60*(3), 505–520.

Cecere, R., Rees, G., & Romei, V. (2015). Report individual differences in alpha frequency drive crossmodal illusory perception. *Current Biology, 25*(2), 231–235, doi:10.1016/j.cub.2014.11.034.

Cohen, Y. E., & Andersen, R. A. (2004). Multimodal spatial representations in the primate parietal lobe. In C. Spence & J. Driver (Eds.). *Crossmodal space and crossmodal attention* (pp. 227–319). Oxford, UK: Oxford University Press.

Derrington, A. M. Krauskopf, J., & Lennie, P. (1984). Chromatic mechanisms in lateral geniculate nucleus of macaque. *The Journal of Physiology, 357,* 241–265.

De Lange Dzn, H. (1958). Research into the dynamic nature of the human fovea-cortex systems with intermittent and modulated light. I. Attenuation characteristics with white and coloured light. *Journal of the Optical Society of America, 48*(11), 777–784. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/13588450

Driver, J., & Noesselt, T. (2008). Multisensory interplay reveals crossmodal influences on "sensory-specific" brain regions, neural responses, and judgments. *Neuron, 57*(1), 11–23, doi:10.1016/j.neuron.2007.12.013.

Ernst, M. O. (2006). A Bayesian view on multimodal cue integration. In G. Knoblich, I. M. Thornton, M. Grosjean, & M. Shiffrar (Eds.), *Human body perception from the inside out* (pp. 105–131). Oxford, UK: Oxford University Press.

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature, 415*(6870), 429–433, doi:10.1038/415429a.

Ernst, M. O., & Bülthoff, H. H. (2004). Merging the senses into a robust percept. *Trends in Cognitive Sciences, 8*(4), 162–169, doi:10.1016/j.tics.2004.02.002.

Helbig, H. B., Ernst, M. O., Riccardi, E., Pietrini, P., Thielscher, A., Mayer, K. M., … Noppeney, U. (2012). The neural mechanisms of reliability weighted integration of shape information from vision and touch. *NeuroImage, 60*(2), 1063–1072. http://doi.org/10.1016/j.neuroimage.2011.09.072

Innes-Brown, H., & Crewther, D. (2009). The impact of spatial incongruence on an auditory-visual illusion. *PloS One, 4*(7), e6450, doi:10.1371/journal.pone.0006450.

Ives, H. E. (1912). Studies of the photometry of different colours: I. Spectral luminosity curves obtained by equality of brightness photometry and the flicker photometer under similar conditions. *Philosophical Magazine, 6*(24), 149–188.

Jaekl, P., Pérez-Bellido, A., & Soto-Faraco, S. (2014). On the "visual" in "audio-visual integration": A hypothesis concerning visual pathways. *Experimental Brain Research, 232,* 1631–1638, doi:10.1007/s00221-014-3927-8.

Jaekl, P., & Soto-Faraco, S. (2010). Audiovisual contrast enhancement is articulated primarily via

Pérez-Bellido, Ernst, Soto-Faraco, & López-Moliner

the M-pathway. *Brain Research, 1366,* 85–92, doi: 10.1016/j.brainres.2010.10.012.

Jiang, Y., Zhou, K., & He, S. (2007). Human visual cortex responds to invisible chromatic flicker. *Nature Neuroscience, 10*(5), 657–662, doi:10.1038/nn1879.

Kayser, C., Petkov, C. I., Augath, M., & Logothetis, N. K. (2005). Integration of touch and sound in auditory cortex. *Neuron, 48*(2), 373–384, doi:10.1016/j.neuron.2005.09.018.

Keil, J., Müller, N., Hartmann, T., & Weisz, N. (2014). Prestimulus beta power and phase synchrony influence the sound-induced flash illusion. *Cerebral Cortex, 24*(5), 1278–1288, doi:10.1093/cercor/bhs409.

Kulikowski, J. J. (1971). Some stimulus parameters affecting spatial and temporal resolution of human vision. *Vision Research, 11*(1), 83–93.

Lakatos, P., Chen, C.-M., O'Connell, M. N., Mills, A., & Schroeder, C. E. (2007). Neuronal oscillations and multisensory interaction in primary auditory cortex. *Neuron, 53*(2), 279–292, doi:10.1016/j.neuron.2006.12.011.

Lange, J., Oostenveld, R., & Fries, P. (2011). Perception of the touch-induced visual double-flash illusion correlates with changes of rhythmic neuronal activity in human visual and somatosensory areas. *NeuroImage, 54*(2), 1395–405, doi:10.1016/j.neuroimage.2010.09.031.

Lee, B. B., Pokorny, J., Smith, V. C., Martin, P. R., & Valberg, A. (1990). Luminance and chromatic modulation sensitivity of macaque ganglion cells and human observers. *Journal of the Optical Society of America. A: Optics and Image Science, 7*(12), 2223–2236.

Lennie, P. (1980). Parallel visual pathways: A review. *Vision Research, 20*(7), 561–594, doi:10.1016/0042-6989(80)90115-7.

Macmillan, N., & Creelman, C. (2004). Detection theory: A user's guide. Retrieved from http://books.google.es/books?hl=es&lr=&id=P094AgAAQBAJ&oi=fnd&pg=PP1&dq=macmillan+creelman&ots=bbxk1zQag2&sig=Xa1Q5jt8kBjL1BECwcW4wSbgib0

McCormick, D., & Mamassian, P. (2008). What does the illusory-flash look like? *Vision Research, 48*(1), 63–69, doi:10.1016/j.visres.2007.10.010.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature, 264*(5588), 746–748, doi: 10.1038/264746a0.

Meredith, M., & Stein, B. E. (1983). Interactions among converging sensory inputs in the superior colliculus. *Science, 221*(4608), 389–391, doi:10.1126/science.6867718.

Merigan, W. H. (1989). Chromatic pathway and achromatic vision of macaques: Role of the P pathway. *The Journal of Neuroscience, 9,* 776–783.

Miller, N. (1965). *Visual Recovery from High Intensity Flashes. USAF School of Aerospace Medicine, AFSC, Brooks Air Force Base.* Retrieved from http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=AD0627325

Mishra, J., Martinez, A., Sejnowski, T. J., & Hillyard, S. A. (2007). Early cross-modal interactions in auditory and visual cortex underlie a sound-induced visual illusion. *The Journal of Neuroscience, 27*(15), 4120–4131, doi:10.1523/JNEUROSCI.4912-06.2007.

Mullen, K., Thompson, B., & Hess, R. (2010). Responses of the human visual cortex and LGN to achromatic and chromatic temporal modulations: An fMRI study. *Journal of Vision, 10*(13):13, 1–19, doi:10.1167/10.13.13. [PubMed] [Article]

Murray, M. M., Thelen, A., Thut, G., Romei, V., Martuzzi, R., & Matusz, P. J. (2015). The multisensory function of the human primary visual cortex. *Neuropsychologia,* E-pub ahead of print. http://dx.doi.org/10.1016/j.neuropsychologia.2015.08.011i.

Murray, M. M., Thelen, A., Thut, G., Romei, V., Martuzzi, R., & Matusz, P. J. (2015). The multisensory function of primary visual cortex in humans. *Neuropsychologia,* [epub ahead of print.]

Nagy, A., Eördegh, G., Paróczy, Z., Márkus, Z., & Benedek, G. (2006). Multisensory integration in the basal ganglia. *The European Journal of Neuroscience, 24*(3), 917–924, doi:10.1111/j.1460-9568.2006.04942.x.

Nassi, J. J., & Callaway, E. M. (2009). Parallel processing strategies of the primate visual system. *Nature Reviews Neuroscience, 10*(5), 360–372, doi:10.1038/nrn2619.

Osterberg, G. (1935). Topography of the layer of rods and cones in the human retina. Copenhagen [Levin & Munksgaard]. Retrieved from http://worldcat.org/oclc/6108551.

Pérez-Bellido, A., Soto-Faraco, S., & López-Moliner, J. (2013). Sound-driven enhancement of vision: Disentangling detection-level from decision-level contributions. *Journal of Neurophysiology, 109*(4), 1065–1077, doi:10.1152/jn.00226.2012.

Pokorny, J., & Smith, V. C. (1997). Psychophysical

signatures associated with magnocellular and parvocellular pathway contrast gain. *Journal of the Optical Society of America A: Optics, Image Science, and Vision, 14*(9), 2477–2486.

Rager, G., & Singer, W. (1998). The response of cat visual cortex to flicker stimuli of variable frequency. *The European Journal of Neuroscience, 10*(5), 1856–1877.

Ress, D., & Heeger, D. J. (2003). Neuronal correlates of perception in early visual cortex. *Nature Neuroscience, 6*(4), 414–420, doi:10.1038/nn1024.

Rosenthal, O., Shimojo, S., & Shams, L. (2009). Sound-induced flash illusion is resistant to feedback training. *Brain Topography, 21*(3-4), 185–192, doi:10.1007/s10548-009-0090-9.

Satterthwaite, F. E. (1946). An approximate distribution of estimates of variance components. *Biometrics, 2*(6), 110–114, http://doi.org/10.2307/3002019

Shams, L. (2002). Visual illusion induced by sound. *Cognitive Brain Research, 14*(1), 147–152, doi:10.1016/S0926-6410(02)00069-1.

Shams, L., Kamitani, Y., & Shimojo, S. (2000). Illusions: What you see is what you hear. *Nature, 408,* 788.

Shams, L., Ma, W. J., & Beierholm, U. (2005). Sound-induced flash illusion as an optimal percept. *Neuroreport, 16*(17), 1923–1927. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/16272880.

Shipley, T. (1964). Auditory flutter-driving of visual flicker. *Science, 145*(3638), 1328–1330. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/14173429.

Simonson, E., & Brožek, J. (1952). Flicker fusion frequency: Background and applications. *Physiological Reviews, 32,* 349–378.

Stocker, A. A., & Simoncelli, E. P. (2006). Noise characteristics and prior expectations in human visual speed perception. *Nature Neuroscience, 9*(4), 578–585. http://doi.org/10.1038/nn1669.

Sun, H., Pokorny, J., & Smith, V. C. (2001). Rod-cone interactions assessed in inferred magnocellular and parvocellular postreceptoral pathways. *Journal of Vision, 1*(1):5, 42–54, doi:10.1167/1.1.5. [PubMed] [Article]

Swanson, W. H., Ueno, T., Smith, V. C., & Pokorny, J. (1987). Temporal modulation sensitivity and pulse-detection thresholds for chromatic and luminance perturbations. *Journal of the Optical Society of America A: Optics and Image Science, 4*(10), 1992–2005.

Takeshima, Y., & Gyoba, J. (2015). Spatial frequency modulates the degree of illusory second flash perception. *Multisensory Research, 28* (1–2), 1–10, http://doi.org/10.1163/22134808-00002468.

Tyler, C. W., & Hamer, R. D. (1990). Analysis of visual modulation sensitivity. IV. Validity of the Ferry-Porter law. *Journal of the Optical Society of America A, 7*(4), 743. http://doi.org/10.1364/JOSAA.7.000743.

Wallace, M. T., Meredith, M. a, & Stein, B. E. (1998). Multisensory integration in the superior colliculus of the alert cat. *Journal of Neurophysiology, 80*(2), 1006–1010.

Watkins, S., Shams, L., Tanaka, S., Haynes, J.-D., & Rees, G. (2006). Sound alters activity in human V1 in association with illusory visual perception. *NeuroImage, 31*(3), 1247–1256, doi:10.1016/j.neuroimage.2006.01.016.

Werner, S., & Noppeney, U. (2010). Superadditive responses in superior temporal sulcus predict audiovisual benefits in object categorization. *Cerebral Cortex, 20*(8), 1829–1842, doi:10.1093/cercor/bhp248.

Witt, J. K., Taylor, J. E. T., Sugovic, M., & Wixted, J. T. (2015). Signal detection measures cannot distinguish perceptual biases from response biases. *Perception, 44*(3), 289–300, doi:10.1068/p7908.

Wozny, D., Beierholm, U., & Shams, L. (2008). Human trimodal perception follows optimal statistical inference. *Journal of Vision, 8*(3):24, 1–11, doi:10.1167/8.3.24. [PubMed] [Article]