

# PSYCHOACOUSTIC MANIPULATION OF THE SOUND INDUCED ILLUSORY FLASH

Sonia Wilkie  
s.wilkie@uws.edu.au

Catherine Stevens  
kj.stevens@uws.edu.au

Roger Dean  
roger.dean@uws.edu.au

MARCS Auditory Laboratories  
University of Western Sydney  
Locked Bag 1797  
Penrith South 1797  
Sydney, Australia

## ABSTRACT

Psychological research on cross-modal perception has focused on the manipulation of sensory information predominantly by visual information. There is a lacuna in using auditory stimuli to manipulate other sensory information. The Sound Induced Illusory Flash is one illusory paradigm that uses the auditory system to bias other sensory information. However, more research is needed into the different conditions under which the Sound Induced Illusory Flash manifests and is enhanced or reduced.

The experiment reported here investigates the effect of new auditory variables on the Sound Induced Illusory Flash. The variables to be discussed include the use of pitch intervals and harmonic relationships. The ultimate aim is to develop the illusory effect as a basis for new multi-media techniques and creative applications for the temporal manipulation and spatialisation of visual objects.

## 1. BACKGROUND

Research on cross-modal interactions in perception have focused predominantly on conditions where visual stimuli are used to manipulate auditory perception<sup>1 2</sup>. The results of these experiments suggest that vision is the dominant sense. However, the Sound Induced Illusory Flash<sup>3</sup> exploits the capacity of the auditory system to distort visual perception.

Conditions that give rise to this cross-modal illusion involve presentation of a visual stimulus consisting of a single white dot that is flashed once in the participant's peripheral visual field. This is accompanied by an auditory stimulus of multiple beeps of sound. The dual presentation of temporal stimuli appearing to emanate from a single source creates confusion regarding the number of physical flashes perceived and gives rise to the percept of the dot flashing equivalent to the number of

auditory beeps. This illusory percept appears to occur because of the superior resolution of the auditory system for rhythmic perception which, in this case, overrides visual information.

The Sound Induced Illusory Flash is a recent discovery. Whilst recent research<sup>4 5</sup> has focused on neural mechanisms that underpin the illusion, only the initial studies<sup>6 7 8</sup> explore basic structural variables that give rise to the illusion. It is these structural variables that are of interest for further exploration of the illusion experience. Such research outcomes provide new opportunities for creative application and multimedia transmission techniques where auditory stimuli might influence visual perception in novel and interesting ways.

## 2. THE SOUND INDUCED ILLUSORY FLASH AND AUDITORY STIMULUS VARIABLES – A REVIEW

### 2.1. Rhythm

The variables manipulated in previous research consist primarily of differing combinations of the number of auditory and visual stimuli presented<sup>9</sup>, with only minor adjustments to stimuli across experiments (including pitch frequency at 1kHz or 3.5kHz; the transmission of auditory stimuli through headphones or speakers; and background screen colour of grey or black). These minute alterations of the stimuli were not considered by Shams and colleagues as significant variables, hence

<sup>1</sup> McGurk and Macdonald (1976), 'Hearing lips and seeing voices' *Nature*. 264. 746 – 748.

<sup>2</sup> Alais and Burr (2004) 'The ventriloquist effect results from near-optimal bi-modal integration' *Current Biology*. 14. 257 – 262.

<sup>3</sup> Shams, Kamitani and Shimojo (2002) 'Visual illusion induced by sound' *Cognitive Brain Research*. 14: 147-152.

<sup>4</sup> Shams, Iwaki, Chawla and Bhattacharya. (2005) 'Early modulation of visual cortex by sound: an MEG study' *Neuroscience Letters*. 378: 76–81.

<sup>5</sup> Shams (2005) 'Sound induced flash illusion as an optimal percept' *Neuroreport*. vol 16: 17.

<sup>6</sup> Shams (2002) 'Integration in the brain - The subconscious alteration of visual perception by cross-modal integration' *Science and Consciousness Review*. 1: 1-4.

<sup>7</sup> Shimojo and Shams (2001) 'Sensory modalities are not separate modalities: plasticity and interactions' *Current Opinion in Neurobiology*. 11. 505-509.

<sup>8</sup> Shimojo, Scheier, Nijhawan, Shams, Kamitani and Watanabe (2001) 'Beyond perceptual modality: auditory effects on visual perception' *Acoustic Science and Technology*. 22: 2. 61-67.

<sup>9</sup> Shams, Kamitani, and Shimojo (2002) 'Visual illusion induced by sound' *Cognitive Brain Research*. 14: 147-152.

there was little discussion of their effects or interactions.

Brief discussion of the visual stimulus concludes that the illusory percept is stronger when the dots are placed in the periphery rather than fovea<sup>10</sup>, but this concept has not expanded into research on exact spatial location in 360° peripheral vision. The number of dots presented has been manipulated but their correlation with spatialisation has not been investigated.

A broad examination of research into auditory perception reveals that rhythm of stimuli and its timescale, may be important variables. Research has investigated the duration of the stimulus gap between the auditory and visual stimuli before illusory fragmentation occurs<sup>11</sup>.

Whilst this research explores the elementary formations of rhythm, research has not investigated the gap size that would cause perceptual fusion between the auditory and visual stimuli, nor has it investigated the combination of various durations to form rhythmic motifs, or the potential for auditory rhythm motifs to create perception of visual rhythm. Shams noted that "*Pilot behavioural work confirmed that whether beeps and flashes were presented simultaneously or with slight temporal offset made little difference to behavioural reports of illusory perception*"<sup>12</sup>.

To have confidence in this important conclusion, there is a need for closer and systematic examination of the spatial disparity and duration of the gap for both large and short durations and the effects on fusion.

## 2.2. Manipulation of Fine-Grained Time Scale

Time scale is an important variable that has been employed and manipulated in many illusory paradigms. The capture of sensory information is accrued by the provision of timescale; therefore, misperception sometimes occurs from insufficient time to acquire sensory information. Employment of the micro time scales as a variable includes Microsounds<sup>13</sup> with stimuli generated at 600ms or less; the Octave/ Scale/ Chromatic Illusion<sup>14 15</sup> involving auditory stimuli of 250ms; The Illusory Continuity of Tones<sup>16</sup> consisting of noise at 50ms or less; and the Auditory Driving of Visual Flicker<sup>17</sup>

comprising auditory stimuli at 150ms. The micro time scale and limited duration of the auditory stimulus employed in the Sound Induced Illusory Flash is an important variable as the illusion fragments when the beeps expand in duration from 70ms onwards to 100ms<sup>18</sup>.

The structural variable of rhythm has been explored in some depth in current research into the Sound Induced Illusory Flash. However, expansion of this variable would be problematic as the micro time scale of the auditory stimulus is too short for it to be perceived as rhythm. Further, participants would require musical knowledge of simple rhythms to be able to report the visual rhythm they perceived.

One way to further explore the illusion, particularly with respect to creative applications, is to examine variables that have been manipulated in the generation of other uni-modal and cross-modal illusions that might be applied to the Sound Induced Illusory Flash.

## 2.3. Frequency and Pitch Interval

Frequency as a stimulus variable for cross-modal manipulation was introduced by Marks's exploration on the Mediation of Brightness, Pitch, and Loudness<sup>19</sup>. This psychophysical research does not exhibit cross-modal manipulation, but cross-modal association between auditory and visual stimuli (greyscale hue and pitch). However, the employment of pitch as a variable has been effective in uni-modal auditory illusions.

The Octave<sup>20</sup>, Scale<sup>21</sup> and the Chromatic illusions<sup>22</sup> pit the perceptual grouping principles of similarity of frequency and spatialisation against one another resulting in an illusory percept based on pitch proximity. Both of these variables – pitch proximity and spatial location – may translate to cross-modal illusions.

## 2.4. Harmonic Relationship

The variable of intervallic close harmonic relationship can be used to motion direction of motion of a tone or pitch. This variable is most notably exploited in investigations of the Tritone Paradox<sup>23</sup> that employ multiple layered frequencies at the octave; and Shepard Tones<sup>24 25</sup> that consist of multiple layered frequencies at the octave or the augmented 5th.

There is a need for investigation of new variables – pitch interval and spatialisation – and their effects on

<sup>10</sup> Shimojo, Scheier, Nijhawan, Shams, Kamitani and Watanabe (2001) 'Beyond perceptual modality: auditory effects on visual perception' *Acoustic Science and Technology*. 22: 2. 61-67.

<sup>11</sup> Shams (2002) 'Integration in the brain - the subconscious alteration of visual perception by cross-modal integration' *Science and Consciousness Review*. 1: 1-4.

<sup>12</sup> Shams (2005) 'Sound induced flash illusion as an optimal percept' *Neuroreport*. vol 16: 17.

<sup>13</sup> Roads (2001) *Microsound*. The MIT Press: Cambridge, Massachusetts.

<sup>14</sup> Deutsch (1981) 'The octave illusion and auditory perceptual integration' *Hearing Research and Theory*. 1. 99-142.

<sup>15</sup> Deutsch (1975) 'Two-channel listening to musical scales' *Journal of the Acoustical Society of America*. 57: 1156-1160.

<sup>16</sup> Bregman (1999) *Auditory scene analysis*. The MIT Press: Cambridge Massachusetts.

<sup>17</sup> Shipley (1964) 'Auditory flutter driving of visual flicker' *Science*. vol 145: 1328-133.

<sup>18</sup> Shams (2005) 'Sound induced flash illusion as an optimal percept' *Neuroreport*. vol 16: 17.

<sup>19</sup> Marks (1974) 'On the associations of light and sound: the mediation of brightness, pitch, and loudness' *American Journal of Psychology*. Vol 87: 173-188.

<sup>20</sup> Deutsch (1974) 'An auditory illusion' *Nature*. 251: 307-309.

<sup>21</sup> Deutsch (1975) 'Two-channel listening to musical scales' *Journal of the Acoustical Society of America*. 57: 1156-1160.

<sup>22</sup> Deutsch (1988) 'The semitone paradox' *Music Perception*. 6: 2. 115-132.

<sup>23</sup> Deutsch (1986) 'A musical paradox' *Music Perception*. 3: 275-280.

<sup>24</sup> Shepard (1964) 'Circularity in judgements of relative pitch' *The journal of the Acoustical Society of America*. 36: 2346-2353.

<sup>25</sup> Risset (1972) *Musical Acoustics*. IRCAM: Paris.

auditory-visual perception. A lacuna is evident in studies of the Sound Induced Illusory Flash that only manipulate the variables of micro time scales with allusions to rhythm, to distort visual temporal perception. Pitch interval and spatialisation were used for purposes of illusory emphasis, the generation of perceived motion and spatialisation, therefore there application as variables to articulate visual rhythm, create visual motion and spatialisation should translate to the Sound Induced Illusory Flash.

### 2.5. Aim, Design and Hypotheses

The aim of the experiment was to enhance illusory perception. The experimental design consisted of three independent variables: beep pitch separation (unison, octave), presentation (monaural, binaural), and beeps (2, 3, 4, 5). The dependent variable was the number of flashes perceived. Based on the foregoing it was hypothesised that a contrast in auditory stimulus will articulate each beat emphasising the illusory effect. That the variable of pitch separation at the octave is anticipated to create a greater illusory percept than unison, and that the variable of binaural presentation will create greater illusory percept than monaural presentation.

The application of pitch interval as an auditory stimulus contrast will generate two auditory fixation points corresponding to the high and low pitches (reflective of the Octave illusion), articulating the apparent dot to flicker accordingly and rhythmically, whilst emphasising the illusory effect. An interval of an octave is employed, as this interval is the closest harmonically to the unison, and therefore the most conservative option for manipulation of the variable

Monaural versus binaural presentation is used as an auditory contrast to draw spatial attention to individual beeps that may further enhance the illusory effect.

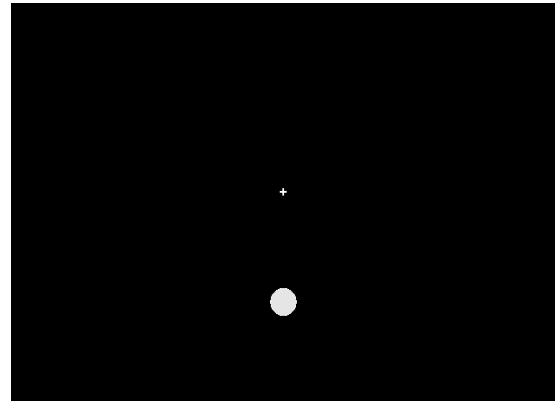
## 3. METHOD

### 3.1. Participants

A sample of 40 participants naïve to the illusion were recruited. They were Psychology 1A students from the University of Western Sydney and received course credit for their participation. Participants were aged between 17 and 54 years ( $M = 21.18$  years,  $SD = 6.68$ ), with more female participants than male participants (36 female, 4 male). People reporting a hearing impairment, visual impairment (corrected to normal vision allowed), severe migraines or epilepsy were excluded from testing.

### 3.2. Stimuli

The visual stimulus consisted of a centrally located fixation point, and a single white dot positioned below the centre of the screen that was located in the participants' peripheral vision. The visual angle of dot was  $2^\circ$  below the fixation point.



**Figure 1.** Screen Capture of the visual stimulus. The fixation point is the centrally located cross. The dot appeared below the fixation point for 17 ms.

In 2-beep illusory trials the dot was presented 23ms after the auditory stimulus onset and for the duration of 17ms. The visual stimulus remained the same in the trials, with the independent variables concerning only the aural stimuli.

The auditory stimulus consisted of a sine tone generated every 50ms and lasted for a total of 7ms (attack 2ms, sustain 3ms, decay 2ms).

The variables manipulated were:

- The number of auditory stimulus beeps, presented at two, three, four and five generations.
- Intervallic pitch of beeps with the Unison set at 261.5Hz versus an Octave separation of 261.5 and 523Hz.
- Monaural versus binaural presentation via headphones.

The interval variable was presented within subjects and the presentation variable between subjects.

Yielding 16 conditions in total, each condition was presented six times.

### 3.3. Equipment

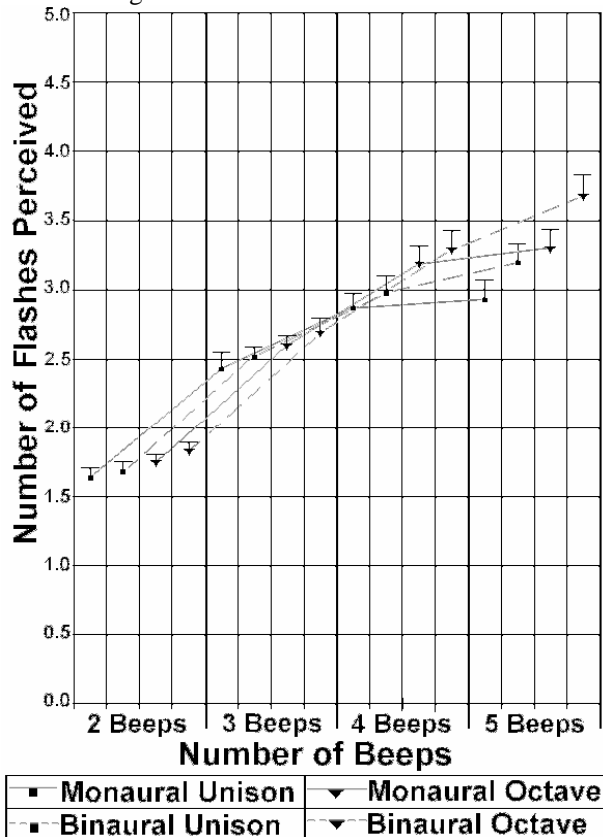
Participants were located at computer workstations with their head positioned on a chin rest 40cm from the computer monitor and eyes level with the fixation point. A Mac Pro G5 with a Diamond digital CRT monitor was used with sound transmitted through AKG K601 headphones. MAX / MSP was used to construct an application that generated the auditory and visual stimulus; presented the trials in a randomised and collected order (using the urn object); generated the questionnaire; and collected the participants responses in a text file.

### 3.4. Procedure

Participants were instructed to place their head on the chin rest and focus on the fixation point. They were asked to use their peripheral vision to count the number of times a dot was presented. The task required them to state on a multiple choice questionnaire the number times the dot was presented, ranging from one event to five events, within an 8 second time limit.

### 3.5. Results

The data collected in the experiment reports the number of flashes perceived. The experiment recovered the illusory effect with results suggesting that pitch interval and binaural transmission enhanced the illusory effect. The mean perceived number of flashes, as a function of beep number, spatial presentation, and pitch separation, are shown in Figure 2.



**Figure 2.** The mean number of flashes perceived. Error bars refer to standard error of mean

For the 2-beep condition, there was a significant main effect of interval  $F(1,38)=40.52$ ,  $p<.05$ . Unison pitched beeps elicited fewer perceived flashes ( $M = 1.66$ ,  $SD = 0.27$ ) than pitched beeps that were separated by an octave ( $M = 1.78$ ,  $SD = 0.24$ ). There was no main effect of presentation and no interval  $\times$  presentation interaction. A similar pattern of results was obtained in 3 beep, 4 beep, and 5 beep conditions, ie main effect of interval (unison less than octave), no main effect of presentation and no interval  $\times$  presentation interaction.

Participants recorded greater mean accuracy for 2-beep presentations with accuracy decreasing as the number of beeps increased. Over the two to five beep conditions, the binaural octave condition consistently elicited the highest mean, followed by the monaural octave and binaural unison conditions, with the monaural unison condition eliciting the lowest mean.

These results indicated that modifying the auditory stimulus with the octave interval and binaural transmission enhanced the illusory effect. Separation of beeps by an octave in pitch influences perceived visual rhythm.

### 4. REFERENCES

- [1] Alais, D. and D. Burr. 2004. 'The ventriloquist effect results from near-optimal bi-modal integration' *Current Biology*. 14. 257 - 262.
- [2] Bregman, A. 1999. *Auditory Scene Analysis*. The MIT Press: Cambridge Massachusetts.
- [3] Deutsch, D. 1974. 'An auditory illusion' *Nature*. 251: 307-309.
- [4] Deutsch, D. 1975. 'Two-channel listening to musical scales' *Journal of the Acoustical Society of America*. 57: 1156-1160.
- [5] Deutsch, D. 1981. 'The octave illusion and auditory perceptual integration' *Hearing Research and Theory*. 1. 99-142.
- [6] Deutsch, D. 1983. 'The octave illusion in relation to handedness and familial handedness background' *Neuropsychologia*. 21: 289-293.
- [7] Deutsch, D. 1988. 'The semitone paradox' *Music Perception*. 6: 2. 115-132.
- [8] Deutsch, D. 1986. 'A musical paradox' *Music Perception*. 3: 275-280.
- [9] Marks, L. 1974. 'On the associations of light and sound: the mediation of brightness, pitch, and loudness' *American journal of psychology*. 87.
- [10] McGurk, H. and J. Macdonald. 1976. 'Hearing lips and seeing voices' *Nature*. 264. 746 - 748.
- [11] Risset, J-C. 1972. *Musical Acoustics*. IRCAM: Paris.
- [12] Roads, C. 2001. *Microsound*. The MIT Press: Cambridge, Massachusetts. 86-106, 330-351.
- [13] Shams, L. 2002. 'Integration in the brain - the subconscious alteration of visual perception by cross-modal integration' *Science and Consciousness Review*. 1: 1-4.
- [14] Shams, L. 2005. 'Sound induced flash illusion as an optimal percept' *Neuroreport*. vol 16: 17.
- [15] Shams, L., S. Iwaki, A. Chawla, and J. Bhattacharya. 2005. 'Early modulation of visual cortex by sound: an MEG study' *Neuroscience Letters*. 378: 76-81.
- [16] Shams, L., Y. Kamitani, and S. Shimojo. 2002. 'Visual illusion induced by sound' *Cognitive Brain Research*. 14: 147-152.
- [17] Shepard, R. 1964. 'Circularity in judgements of relative pitch' *The journal of the Acoustical Society of America*. 36: 2346-2353.
- [18] Shimojo, S., C. Scheier, R. Nijhawan, L. Shams, Y. Kamitani, and K. Watanabe. 2001. 'Beyond perceptual modality: auditory effects on visual perception' *Acoustic Science and Technology*. 22: 2 61-67.
- [19] Shimojo, S. and L. Shams. 2001. 'Sensory modalities are not separate modalities: plasticity and interactions' *Current opinion in neurobiology*. 11. 505-509.
- [20] Shipley, T. 1964. 'Auditory flutter driving of visual flicker' *Science*. vol 145: 1328-1330.