



LUND
UNIVERSITY

DEPARTMENT OF PSYCHOLOGY

*Can Attentional Processes Modulate Bimodal
Interference in V-EAR Synaesthesia?*

Robert Johansson

Bachelor Thesis Spring 2017

Supervisor: Geoffrey Patching

ABSTRACT

Even though the prevalence of synaesthetic-like ‘Visually-Evoked Auditory Responses’ (V-EAR) in the general population is probably immense by all standards, it is hardly ever mentioned in the literature on general synaesthetic phenomena. We decided to try to reproduce earlier findings that about 1/5 experience motion-conjured sound percepts. We did so with a good margin (27% of our participants claimed to ‘hear’ motion). We were also interested in whether these V-EAR experiencers would show of the enhanced visual sequence discrimination capabilities reported in two earlier studies. This finding was not successfully replicated. Neither could we manage to find any intuitively reasonable connections between V-EAR, attention processes and destructive interference between incongruently sequenced bimodal stimuli. Instead, we found evidence of *constructive* interference between incongruently sequenced bimodal stimuli, something that might be due to the ‘auditory driving’ effect. We discuss these findings within the theoretical frameworks as proposed by earlier research on ‘classic’ synaesthesia.

1. BACKGROUND

1.1 V-EAR Synaesthesia

In case you have never heard of V-EAR synaesthesia, you are most definitely not alone. V-EAR stands short for ‘Visually-Evoked Auditory Responses’, and is the ‘synaesthetic’ phenomena in which the sight of simple motion is not only encoded as visual information, but also ‘leaks’ to auditory regions. Consequently, it also provokes an auditory percept. V-EAR is probably a form of ‘synaesthesia’, which is the somewhat problematic umbrella term denoting a number of atypical perceptual binding processes. But we can not be absolutely sure that it really is, because currently no one agrees exactly on how to define synaesthesia in a proper and all-encompassing way (Rogowska, 2011). Usually it involves the coupling or leakage of information between two (or more) distinctly separated modalities of phenomenal conscious content. The most common forms are of sensory nature, but synaesthesia comes in a wide selection of varieties and flavors. These include tactile-to-feeling, feeling-to-color and the more conceptually based ‘ideosthesia’ (Rogowska, 2011). For the sake of general clarity and coherence of analogy between V-EAR and other forms of synaesthesia, the word will from here on only be used to denote the sensory-to-sensory varieties.

Occurrences of these strange ‘crossmodal correspondences’ have been a recognized phenomena since the beginning of the 19th century when it was first in the medical literature by German physiologist Georg Sachs. But by then it had already stimulated the imaginations of natural philosophers like Isaac Newton and Robert Boyle for hundreds of years, mostly through rumor and word-of-mouth (Jewanski, Day & Ward, 2009). It was first empirically studied by the ‘father of eugenics’ Francis Galton. He noticed how the trait seems to run in families, and was the first to discriminate between the two most common varieties of synaesthetic correspondences. (Galton, 1881). Today we call these two ‘auditory-visual’ and ‘grapheme-color’ synaesthesia. Although neuroatypical, synaesthesia is usually regarded as a trait rather than a form of pathology. Estimates of its occurrence rate in the general population usually range from between

0.05 % and 4 %. But it is noteworthy, however, that these approximations seems to have been consistently underestimated. In a more recent survey it was reported that 24% of all respondents claimed to experience one or more type(s) of synaesthesia (Rouw & Scholte, 2016). But response biases are known to flourish in surveys like this. And since no universal definition of the trait exists, we simply do not know. What is clear, however, is that V-EAR seems to have slipped under the radar of earlier surveys through its relative obscurity. Experimental findings now points to a populatory prevalence of V-EAR as high as that of all other forms of synaesthesia combined. With 22% of participants providing self-reports of V-EAR in a newly conducted randomized sample study (Fassnidge, Marcotti & Freeman, 2017), it might very well constitute the most commonly occurring form of synaesthetic-like phenomena presently identified. These numbers have however not been confirmed through replication yet. This article is the first attempt to do just that.

Moreover, no data on the male-to-female ratio of self-reported V-EAR has been published, and hence this article fills at least one holes in the (somewhat sparse) current body of knowledge surrounding the phenomena. This ratio is of intrested because (although it remains controversial) synaesthesia might have a slight bias of prevalence among women over men (Baron-Cohen et al. 1996). This has led to speculations of a sex-linked mechanism of genetic heritability, but the significance of this link remains uncertain (Simner et al. 2006). If the high incidence rate reported by (Fassnidge, Marcotti & Freeman) is succesfully replicated, one is tempted to look for patterns in its 'epidemiology'. What factors seem to play a role in its genesis? Can we gain insight into this enigma by regarding the relationship between neurotypical multimodal integration, V-EAR phenomenology and other forms of synaesthesia? Could V-EAR be better described simply as a spandrel-like¹ amplification of ordinary crossmodal brain wiring, the epiphenomena of some other trait independent of synaesthesia? And can V-EAR be subject to voluntary control? Because of the relative lack of reliable experimental data on the subject matter, it will be the goal of this article to take all of these questions into account in an integrative manner, while also in its own right contributing to the data accessible. We will do so through a design that primarily puts emphasis on the role that attentional processes and intermodal interference might play in V-EAR. But first we need some theoretical background.

¹ The somewhat arcane term 'spandrel' denotes a behavioural trait without apperant adaptive value (Gould & Lewontin, 1979).

1.2 Phenomenology and Classification

These auditory responses as experienced by V-EAR synaesthetes consist of simple, non-linguistic percept and are often described as very realistic sounds. They fade in and out of amplitude in a smooth and streamlined manner, just like waves. There seems to be a certain amount of temporal instability to these percepts (Hubbard, 2008) and there have been reports of synaesthetic plasticity, e.g. the tendency of sounds to get “mapped” onto specific visual motion stimuli when repeatedly presented together. This suggests that the phenomenal content of the V-EAR percepts have a component of temporal morphology. This is not to say that the occurrence of the V-EAR percepts themselves are unstable or temporary (e.g. as in narcotic and post-hypnotic synaesthesia). But it is worth noting that this is somewhat unique to V-EAR, since other forms of synaesthesia tend to be relatively stable over time (Rogowska, 2011).

V-EAR was first identified in the literature by Saenz & Koch (2008) who noticed that people who claim to hear motion tended to outperform neurotypical controls in a ‘morse code’ visual sequence pair discrimination task. Target stimuli were presented in the form of paired morse sequences made up of either auditory sinewaves or white luminant disk oscillating in fast intervals over a black background. The control and V-EAR groups both performed at around 85% in the auditory condition. But in the visual condition, control performance fell close to chance levels with a mean of 56% pairs correct. The V-EAR synaesthetes showed higher consistency across conditions and managed to discriminate correctly between visual sequences 75% of the time. The authors concluded that unlike controls, the V-EAR group could use their unique auditory percepts in such a manner that they bypassed the relatively low temporal resolution of visual processing.

Rogowska (2011) has proposed a taxonomy of synaesthesia that can be applied to characterize V-EAR as follows: V-EAR seems to be *intermodal* rather than *intramodal* (e.g. involving ‘sensory leakage’ between two distinctly separated sensory modalities rather than between subunits of a single modality), *bimodal* rather than *unimodal* (e.g. conjuring up only one secondary percept), *partial* rather than *comprehensive* (because only motion and no other visual stimuli conjures the

secondary percept), *unidirectional* rather than bidirectional (the leakage is ‘one-way’ only) and finally *conceptual* rather than perceptual (simply meaning that the secondary percept is located ‘in ones head’ rather than in external space). The last dichotomy is similar to the distinction that is sometimes made between associative and projective synaesthesia. While associative secondary percept seems to be of endogenous origin, projective synaesthesia is phenomenologically ‘exogenous’.² These subjective features of V-EAR would according to Rogowskas classificatory system make the trait a synaesthesia of lesser magnitude. However, Rogowska also claims that the intensity of any individual subjects secondary percepts can be defined on a one-dimensional spectrum ranging from ‘subtle’ to ‘intense’.

Taken to its extreme, this view might lead to the conclusion all individuals are synaesthetes to a certain degree, and it has been emphasised by proponents of the ‘synaptic pruning hypothesis’ (see **1.3**) of synaesthesia that even neurotypicals know, for instance, which secondary percept goes with the sound of ‘chalkboard scratching’, eg. the very uncomfortable noise produced when fingernails are dragged against a classroom blackboard (Hubbard & Ramachandran, 2003). We will expand further on the relationship between V-EAR and neurotypical auditory-visual binding phenomena in the next section. Rogowskas spectrum hypothesis of synaesthesia does moreover point to the possibility of idiosyncracies in the phenomenology of V-EAR experience, something which has of yet not been studied. But even though the trait might very well be far from dichotomous in nature, we will for all practical purposes treat it as an all-or-nothing phenomena in our experimental design.

1.3 Crossmodal connectivity and the neural genesis of V-EAR

It has been pointed out by Hubbard (2008) that the V-EAR phenomena seems to involve brain regions that are not adjacent to each other. He goes on to imply that these appear to be V5 for motion processing and the primary and/or secondary auditory cortices. This stands in stark contrast to the synaptic pruning hypothesis of synaesthesia put forward by Ramachandran (2011) who claims that the synaesthetic condition is mediated through a form of local inhibition of neural darwinism which is ultimately dependent on the adjacency of the involved regions. This

² Associative synaesthetic percepts are often described to be taking place ‘in the minds eye’. See (Dixon, Smilek & Merikle, 2004) for a more elaborate overview of the associator/projector dichotomy.

hypothesis is based on observations of the mechanisms of grapheme-color synaesthesia, and Ramachandran goes on to claim that this condition is formed through the genetically hereditary 'cross-wiring' or 'cross-talk' between the color region V4 and the structurally adjacent fusiform gyrus. Many forms of synaesthesia do however seem to involve brain regions that are anatomically widely distributed, and these cases are not so easily explained by the local disruption of pruning. As of yet no reliable data on V-EAR connectivity has been reported, but Hubbard claims that V-EAR might very well be mediated through reasonably direct long-range connections between motion and auditory regions. Similar axonal wiring might mediate other forms of distributed synaesthesia.

Brang, Williams & Ramachandran (2011) have provided some evidence for a correlation between synaesthesia and a general increase in multisensory integration. They used the 'flash illusion paradigm', an illusion arising when two auditory bursts flank the presentation of a single white flashing disk presented on a monitor. Grapheme-color synaesthetes were more likelier than controls to report that not one, but *two* flashing disks had been presented. These synaesthetes were apparently ready to couple together sensory correspondences, independent of their own grapheme-color ones. The authors claim that this points to the possibility that synaesthesia (and consequently also V-EAR) might constitute what the authors refers to as "the tip of the iceberg" of an enhanced general connectivity phenomena, involving a global upregulation of transmodal axonal fibers. This supports the notion that synaesthetic connectivity might not be 'atypical' per se, but more accurately a *magnification* of neurotypical connectivity. Significant increases of white matter in the frontal, parietal and temporal regions has been observed among synaesthetes of the grapheme-color type, implying the non-locality of the condition and supporting the magnified global connectivity hypothesis (Rouw & Scholte, 2007). This is reflected in the tendency of comorbidity between different forms of synaesthesia, and it has for instance been shown that people who experience taste-shape synaesthesia are significantly more likely than neurotypicals to also experience the grapheme-color variety (Simner et al., 2006).

Integration of visual and auditory modules is also apparent in the non-synaesthetic. It has been shown by Lovelace, Stein & Wallace (2003) how vision and audition can work synergetically and improve the signal detection capabilities of neurotypical subjects significantly. In their

paradigm participants were presented with signals of either auditory, visual audiovisual nature. But there were also trials where no stimuli at all was presented. The nature of the task was simple – was the signal present or not? Their synaesthetes made significantly more correct responses in the multimodal condition, but were also slightly more biased to respond with false alarms in the absence of an actual stimulus. In some respects, Lovelace, Stein & Wallace's use of multimodal cues in the audiovisual trials can be regarded as a digital analogue of V-EAR. The synergetic interplay between modalities they found was an example 'constructive interference', while the effect demonstrated in the flash illusion paradigm was interference of 'destructive' nature. Try imagining two waves out of phase with each other. Their crest and troughs will tend to outcancel each other, and this is what's called 'destructive interference'. Now imagine them *in phase*, so that the peaks stack on top of each other. This is 'constructive interference'. The different kinds of stimuli used in for instance Saenz & Koch (2008) paradigm can just like the percepts experienced by their V-EAR participants be described as waves with peaks that rise and fall with time. (See also **Fig.2**) The specific interference phenomena where rhythmic bursts of auditory stimulus 'hijacks' rhythmically incongruent visual stimuli and remolds their temporal structure is an effect known as 'auditory driving'. Auditory driving is not an atypical phenomena, and is also present among non-synaesthetes (Guttman, Gislroy & Blake, 2005). The pairing of temporally congruent visual and auditory stimuli has also been shown to cause constructive interference among neurotypicals. This interference results in enhanced BOLD fMRI-gain in both the primary visual and the primary auditory cortex signal, relative to when the same stimuli are presented independent of each other. This effect is believed to be mediated by multimodal regions in the superior temporal sulcus (Noesselt et al., 2007). This same sulcus partially comprises the areas that Hubbard suggested as candidates for the mediation of V-EAR.

It has also been speculated that all sensory processing is inherently multimodal, and that perception gives precedence to the cleanest signal at hand, e.g. what Welch (1999) termed the 'modality-appropriateness' hypothesis. Recordings from the sensory cortices of higher mammals have supported a multimodal nature of early, low-hierarchy sensory processing for some time (Spinelli, Starr & Barrett, 1968), and it has been observed that even in the absence of auditory stimuli, the simple act of presenting a macaque with an image of a visual landscape will increase the metabolic rate of the primary auditory cortex (Kayser et al. 2007). But not all people report

phenomena akin to that of synaesthesia, and this suggests a clear dividing line between typical and atypical multimodal integration. As already noted by Galton (1881) there seems to be a strong component of heritability among those who actually do experience synaesthesia. As we earlier noted, there has been some evidence pointing to a bias of occurrence among women over men, and therefore speculations of an x-chromosome mediated mode of inheritance have flourished. A meta-review of the literature supporting this hypothesis did however ascribe these findings to a number of methodological flaws. Most important among these were experimenter bias and the tendency of males to be somewhat more reluctant than females to report atypical phenomenal content (Simner & Carmichael, 2015). As of right now at least four distinct chromosomal pathways of inheritance have been identified (Asher, 2009).

However, the role of learning and epigenetic interactions is also speculated to play a very significant role in synaesthetic ontogenesis. For instance, Kadosch, Henik & Walsch (2009) looked at the correlation between the frequency of occurrence of the days of the week (as obtained through feeding them into a standard internet search engine) and the parameters of the color percepts experienced by native hebrew speakers with grapheme-color synaesthesia. Days of high frequency was significantly correlated with warmer color temperatures and higher luminance, but lesser saturation. This serves as an argument of learned synaesthesia, at least in terms of shaping its phenomenal content through experience. Consequently, findings like this poses a dilemma to the more orthodox genetic hypotheses of the neural origins of synaesthesia, and they are a divider amongst the proposed theoretical frameworks. Clinical training paradigms have been designed to create conditioned secondary percepts in neurotypical controls through repeated exposure of colored graphemes. The first one of only managed to induce a color priming driven stroop effect (Rothen & Meier, 2009). Another managed to recreate synaesthetic phenomenology proper. This experiment consisted of ~22 hours of associative training tasks divided over 9 weeks, and involved thirteen specific grapheme-color correspondences. In a post-training follow up, half of all participants provided self-reports of associative secondary percepts, and one claimed that his/her conditioned percept was actually located in exterior space (Bor et al., 2014). There is of course a probability that these participants were just genetically disposed to develop synaesthesia, and the succes of the conditioning trials might partially reflect this

disposition. In the future, large-scale analysis of genetic disposition and associative readiness could be used synergetically to determine the nature of conditioned associative responses.

Hubbard (2008) put emphasis on the rapid development of novel visual-auditory correspondences that was reported by V-EAR experiencers in Saenz and Koch's (2008) sequence discrimination paradigm – correspondences that were cemented within a single session. Plastic visual-auditory correspondences have also been reported among the blind, where it has been observed that auditory spatial processing can 'hijack' V5 through life-long plastic processes in the absence of visual input, making it respond primarily to spatially encoded auditory motion (Saenz et al. 2008).

1.4 Top-down influences

Until only quite recently it was a commonly held belief that synaesthetic percepts arise in an automatic fashion and as such they can not be subject to conscious executive control by the perceiver. Much of this was due to the fact that synaesthetic percepts seem to gain cognitive saliency quite early in the sensory processing hierarchy. Grapheme-color synaesthetes that experience color percepts in conjunction with numeral figures have, for instance, been shown to outperform controls in a visual search task where hidden patterns of achromatic target numerals was arranged in structural patterns organized in either rows or columns. The global arrangements would automatically 'pop out' to the synaesthetes by virtue of their chromatic encoding of the target numerals (Ramachandran & Hubbard, 2001). The same effect could be seen in crowding tasks where achromatic numerals were embedded as geometrical shapes in a soup of numeral flankers (Hubbard, Arman, Ramachandra & Boynton, 2005). These findings supports the notion of a mechanistic, bottom-up basis of synaesthetic percepts.

But there has been some anecdotal reports contradicting thie bottom-up hypothesis. While only 15% of respondents in a large scale qualitative survey of grapheme-color synaesthesia claimed to be in complete control of their secondary percepts, half of them acknowledged how simply attending to a stimuli could increase the magnitude of the lexical-color response. Fatigue, stress and the ingestion of psychoactive substances were also reported as possible modulatory factors (Rich, Bradshaw & Mattingley, 2005). Later the same authors (2006) constructed a synaesthetic

analogue of the classical stroop task, where numeral targets were chromatically congruent or incongruent with the participants idiosyncratic grapheme-color correspondences. A significant decrease of priming effects was observed when attention was divided between the stroop test and a secondary visual task. This suggests that increased cognitive load upon visuospatial working memory can downregulate the expression of grapheme-color correspondences. Modulatory effects of higher order top-down processes were also studied by Ramachandran (2001), who presented grapheme-color synaesthetes with material of multistable synaesthetic nature. Multiples of the numeral three were arranged in such a manner that their global shape resembled the numeral five. When presented with this stimuli the participants reported the possibility to conjure up and 'choose' between local and global color percepts in a manner similar to that whereby the non-synaesthete 'chooses' how to perceive a Necker-cube.

Because we were intrigued by the possibility that similar attentional processes might play a role in modulating V-EAR synaesthesia, we designed an extended version of the Saenz & Koch (2008) paradigm. Based on the findings from earlier studies, we expected the V-EAR group to experience significant constructive interference between the visual target stimuli and their secondary percepts. This would manifest as a main effect *of* or an interaction effect *between* groups and/or target modality. Our first null hypothesis was that no such interference would take place, and that no significant main or interaction effects related to groups and target modality would be observed. We made the addition of visual distractor sequences, and hypothesized that the V-EAR groups performance would indicate amplified destructive interference between distractor-elicited secondary percepts and target sequences of both modalities. Our second null hypothesis was that no such interactions between group and distractor modality would be found.

2. METHODS

2.1 Participants

Twenty-eight individuals between the ages of 18 and 59 (Mean= 28.9, SD= 7.69) partook in in the experiment. These participants were recruited by word-of-mouth. One of them could however not finish the entire procedure, and another asked for his result to be withdrawn from the final data analysis. The final sample consisted of (N=26) participants (14 female and 12

male). One of these participants ('Participant L') was however contacted beforehand in order to participate in a free-flowing interview about V-EAR. All participants had the choice to go through the experimental procedure either in their own home or at the authors housing.

2.2 Apparatus and Stimuli

A Fujitsu Lifebook S751 laptop with an LED backlit 14" screen running MATLAB (R2016b, The MathWorks, Inc.) controlled the experiments. Stimulus presentation and timing were controlled using the Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997).

Participants were seated at about arms length (~57 cm) from the computer monitor. A set of V7 HA510-2NP stereo headphones with volume control set at maximum level was used for auditory stimulus, with laptop audio output set to 50%. Two bursts of morse code would be presented in either the visual or the auditory modality. A crosshair at the center of the monitor provided the coordinates for the visual target stimuli. This visual target stimuli consisted of a white radial disk that would flash in and out of obscurity against a black background in periodic cycles of either 75 or 300 milliseconds. The auditory modality target stimuli were sine wave tones of 360 Hz frequency. Long and short bursts of these sine waves were just like in the visual modality presented for either 300 or 75 milliseconds respectively.

New to this design was the addition of a distractor stimuli in the form of an initial luminant disk, similar in nature to the target stimuli. This distractor would appear in the left or the right periphery of the screen. These flanker stimuli would on half of all trials be of static nature and appear/disappear simultaneously with the target stimulus ("distractor static condition"), or appear as target-incongruent morse sequences ("distractor flash condition"). Participants responded by way of the horizontal arrow keys labelled with green and red colored sticker to signify 'same' or 'different' responses respectively.

2.3 Design

Our design was modelled after the sequence discrimination paradigm developed by Saenz & Koch (2008). All in all the experiment consisted of 160 sequence pairs, out of which half were congruent and half were incongruent (see **Fig.1**). These trials were evenly distributed over a four conditions. These were Auditory Static, Auditory Flash, Visual Static and Visual Flash. These

names denote the modality of the target stimuli and the mobility of the visual distractor stimuli.

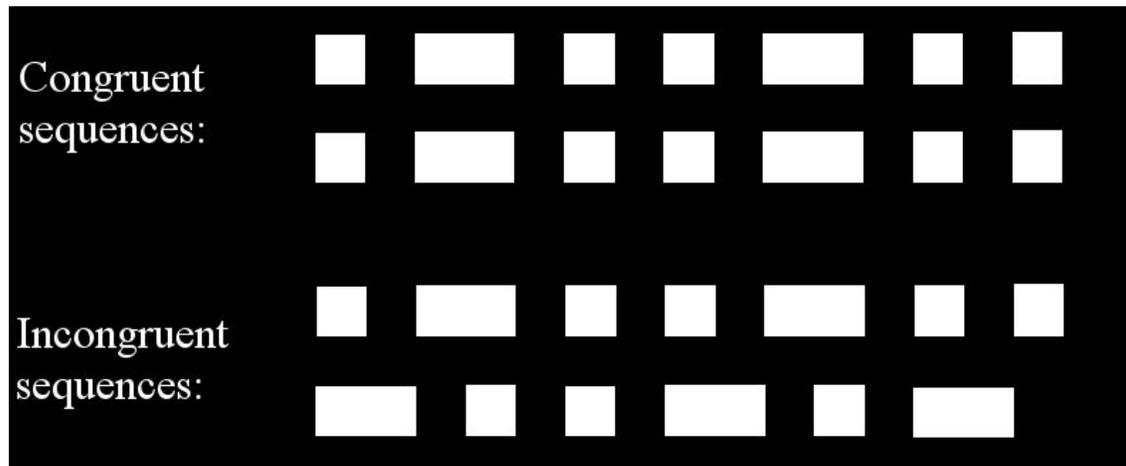


Fig. 2 Approximate illustration of morse code sequence pairs and their temporal relationship. Squares signify 75 ms bursts, rectangles signify 300 ms bursts. All distractor sequences were incongruent with target sequences.

The entire procedure was divided into two equal blocks. One block contained the flashing distractor stimuli and the other contained a static circle accompanying the target. Between these two sessions there was a chance for participants to leave the monitor for a short break. The order of the flash block and the static block was ABBA-counterbalanced between participants.

The two independent within-subject variables of the sequence discrimination task were the sensory nature of the target sequences (auditory or visual) and the mobility of the distractor stimuli (either static or flashing). The dependent variable was the rate of correct discriminatory responses to the rate of incorrect responses. After the sequence discrimination task a scripted interview consisting of six questions was conducted. The participants were placed in the V-EAR or control group through a 'Yes' or 'No' response to question number two (see 2.4). We were interested whether the secondary percepts experienced by the V-EAR group would result in any significant differences in performance between groups. (see Fig.2) All stimuli were presented in a randomized order for counterbalancing, except for the three first and the three last sequence pairs which were presented in identical order to all participants

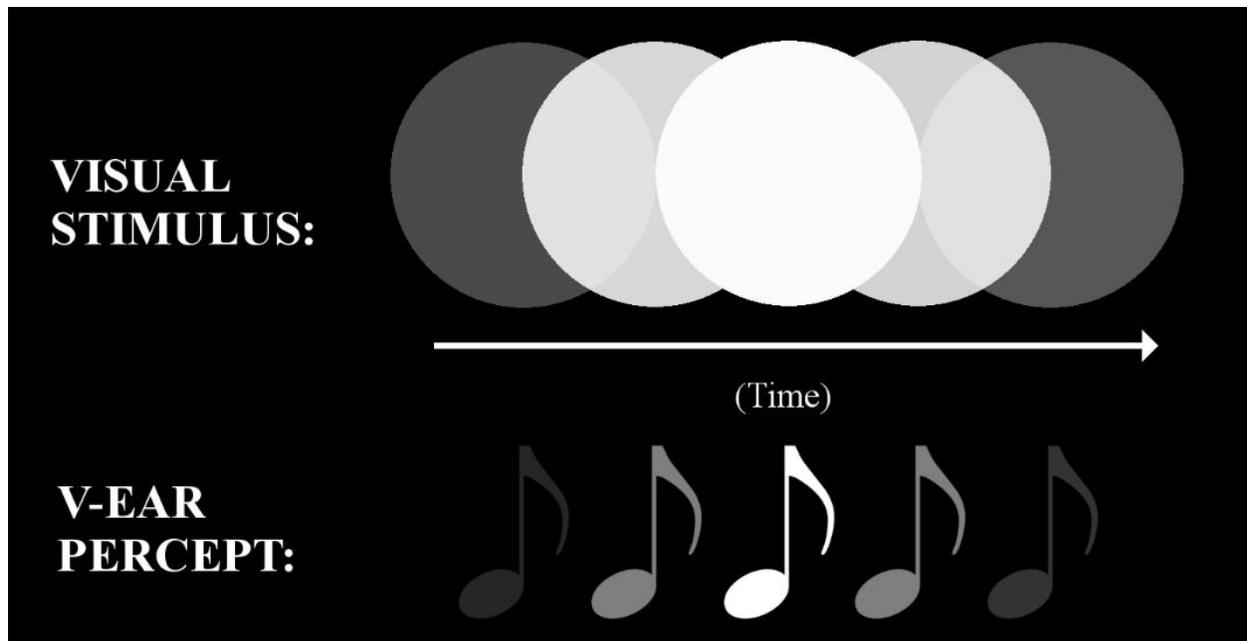


Fig.2 An illustration of one periodic burst of the flashing disk that we used as both visual target and visual distractor stimuli, and one burst of the secondary percept uniquely reported by V-EAR experiencers. Both the physical stimuli and cognitive percept rises and falls with time.

2.4 Procedure

Before commencing with the experiment, all of the participants were informed about the basic logic of the experimental task. Each participant had the chance to make special requests regarding the parameters of the auditory stimuli presented in case they experience a high or low threshold to visual or auditory stimuli in everyday life. All participants were however fine with being presented with auditory stimuli at agreed upon volume, independent of individual auditory threshold. One participant requested the brightness of the screen to be lowered by three units because of migraine attacks associated with highly ruminant stimuli. Confounding sources of visual and/or auditory stimuli (such as room lighting and the humming of electrical machinery) was not balanced for, and interfering sources of sound could have made it through the static noise. About half of all participants provided reports of computer lag during the procedure, but this was never present in more than three sequence pairs. After completing both blocks of the sequence discrimination task, all participants partook in a scripted interview consisting of the same six questions that Fassnidge, Marcotti and Freeman (2017) made use of in their study.

These were (1.) what strategy they used to encode the flashing circles, (2.) if they encoded the flashing circles as auditory beeps, (3.) if they could actually hear the circles flashing, (4.) whether they experienced V-EAR phenomena in everyday life, (5.) if they experienced other synaesthetic phenomena and finally (6.) if they suspected that they might themselves be synaesthetes. The answer to question (3.) (if they could actually *hear* the circles) was used as the criteria for placing any participant in the V-EAR (*yes*) or control group (*no*) respectively. In case there was any problem with the clarity of this question the participant would be shown the Saenz & Koch (2008b) V-EAR self-diagnosis demo on a 5'' Samsung Galaxy S4 monitor. After the interview there was an optional debriefing where the logic and variables of the procedure were revealed to the participants.

2.6 Ethics

All participants partook in the procedure under informed consent and the procedure followed the ethical guidelines as established by the Swedish research center Vetenskapsrådet.

3. Results

3.1 Data Analysis

In the first instance the proportion of correct responses (hits and correct rejections) and the proportion of incorrect responses (misses and false alarms) was calculated for computation of $d' = Z(\textit{correct}) - Z(\textit{incorrect})$ (Green & Sweets, 1966). Following the procedures of Fassinidge, Marcotti and Freeman (2017), all subsequent analysis were conducted using d' as the dependent measure. The multivariate results obtained using mixed analysis of variance are reported to avoid problems associated with sphericity. All analysis we conducted using IBM SPSS versions 24 and the mixed ANOVA was modeled with Type II sums of squares (as opposed to the default Type III sums of squares) to counter problems associated with unbalanced designs. A split plot ANOVA was conducted to compare the main effects of the within-participants factors module (vision and audition) and distractor mobility (static and flash) and the interaction effect of module and mobility on d' scores. We used an alpha level of 0.25 for all statistical tests.

3.2 Findings

The results of our multifactorial analysis revealed a main effect of auditory module $F(1, N=24) = 52.90, p < 0.025$, indicating the mean d' score was significantly greater for the auditory flash condition ($M=1.41, SD=0.4$) than for auditory static ($M= 1.28, SD=0.6$). No other main effect or interaction were found to be statistically significant (all p 's > 0.025). **Fig.3** shows d' and standard deviations of V-EAR and control groups across conditions.

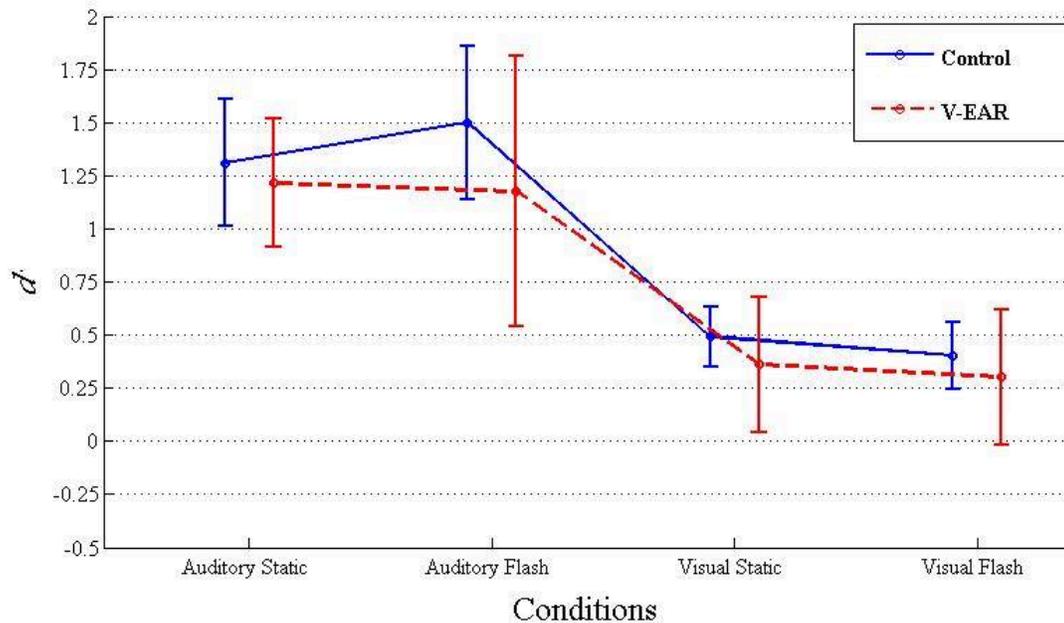


Fig. 3 Shows mean d' scores of both groups plotted across all four conditions.

During the scripted post-task interview, (N=7) or ~27% out of the total number of participants (N=26) answered 'yes' when questioned whether the flashing circles elicited auditory responses. All of the 'yes' responders were female, and the total amount of female participants was (N=14). The most commonly reported strategies used for visual target encoding across both groups was counting the flashes, using the fingers to tap along with their rhythms, and variations of 'just feeling it' or 'going with the flow'. Five participants in the V-EAR group and four controls reported synaesthetic correspondences in everyday life independent of hearing motion.

4. Discussion

The results reported by Saenz & Koch (2008) and Fassnidge, Marcotti & Freeman (2017) were not successfully replicated. The V-EAR group did not show of the predicted performance trajectory between conditions (relative to their control group counterparts) and both of our null hypotheses were true. The magnitude of constructive and destructive interference can not be properly evaluated from the data at hand, and the real relationship between selective attention and V-EAR remains uncertain. This uncertainty is at least partially a reflection of the asymmetry of standard deviations that was observed between groups. This is a problem inherent to the nature of unbalanced designs. But it might also reflect difficulties in the self-evaluation process of the scripted interview. In the future, more precise experimental paradigms independent of those that we tried to replicate here could be designed to evaluate the relationship between V-EAR and selective attention in more detail.

The relatively absent effect of peripheral distractor interference in our data could be due to the cortical magnification factor, an effect that arises from the fact that the photoreceptor density of the eye decreases when the distance and relative angle from the fovea is increased. This results in lesser cortical dedication to the peripheral visual field in the occipital lobes retinotopographic mapping (Anstis, 1998). This hardwired compression of the peripheral visual field could theoretically result in a absence of interfering cognitive load upon working memory capacity. However eye tracking devices would have had to be used to determine whether the participants actually kept their gazes fixed upon the target crosshair to ensure that the distractor stimuli was presented to the peripheral visual field. It is equally likely that the distractor stimuli might have elicited an orienting response through the rapid change of its parameters, thereby imposing an involuntary attracting force. But this is not all. The *improvement* in performance that was observed in the Auditory Distractor condition does actually require an independent explanation, beyond that of the mere lack of interference. Could it be that the auditory target stimuli did actually ‘hijack’ the distractor stimuli, thereby remolding its temporal structure to that of its own, as in the case of auditory driving? This would thus lead to a form of ‘driven’ destructive interference, promoting double encoding of the auditory target stimuli. This driving could have played a part in mediating the results observed.

The case of our V-EAR groups performance is even more perplexing. Based on earlier findings, one might intuitively have thought that the participants who claimed to hear the flashing circles might have performed (a.) more consistently across conditions through selective inhibition of ‘unwanted’ secondary percepts, or (b.) significantly worse in the auditory distractor condition because of destructive interference between stimuli and secondary percept putting a higher cognitive load upon phonological working memory. A simple (but perhaps not very satisfying) explanation of our data would be that it reflects random noise inherent to the experiment because of an interaction between our small V-EAR sample size and the magnitude of computer lag that our participants reported. But the magnitude of these noise factors remain hard to evaluate. Even small sources of noise can however be amplified when working in synergy with the uncertainty of unbalanced designs.

Moreover, a great amount of factors beyond the mere presence or absence of self-reported V-EAR is likely to have played a role in mediating the results that we observed. Many different strategies of stimulus encoding were reported among participants, and the control groups outperformance of the V-EAR group could reflect a more effective chunking together of individual stimuli bursts. Here a higher level of musical intelligence and familiarity with rhythmic sequences could have provided advantages in terms of phonological working memory capacity. Musicians know quite well how the act of restructuring a rhythmic sequence into more manageable units of doublets and triplets of whole and quarter ‘notes’ tends to facilitate performance in these matters. Indeed, one of the most commonly reported strategies across both groups was counting and/or using fingers to tap along with the rhythms of the target sequences.

Another explanation of our data could be that the variability in performance among our synaesthetes might reflect idiosyncrasies of V-EAR phenomenology and connectivity. Rogowska (2011) noted that the phenomenal amplitude of secondary percepts can vary between individuals, and consequently there could also exist individual differences in general synaesthetic connectivities. Indeed, the associator-projector dichotomy observed in many other forms of synaesthesia seem to demand such differences, if we are not to be left with an explanatory gap. Perhaps the hypothesis of direct, long-range connections between primary sensory regions proposed by Hubbard (2008) does not apply to all of our V-EAR experiencers. Participants with ‘detour’ connectivity or sensory leakage of higher origin in the sensory processing hierarchy may

not have benefited equally well from the increase in auditory temporal resolution relative to that of the visual modality. In those cases where participants were uncertain as to whether the visual stimuli had promoted V-EAR, the self-diagnosis demo that we utilized might still have conjured such ‘detour’ responses. This demo consisted of visual motion burst cycles of about 1 second periodicity – cycles way slower than our target oscillations. Consequently there was then a longer lasting critical window opened for V-EAR percepts to gain cognitive salience. The plastic learning processes and genetic differences involved in synaesthetic ontogenesis could play a role in modulating such idiosyncracies of V-EAR connectivity and phenomenology. With their relatively modest sample sizes, earlier studies could easily have bypassed the variable of individual differences by chance alone.

We also found that half of all female participants responded ‘Yes’ when asked whether they heard the flashing circles. This leaves us with a slight response bias of self-reported V-EAR. Earlier studies have not reported this (Saenz & Koch, 2008, Fassnidge, Marcotti & Freeman, 2017). Some might attribute this to the general tendency of synaesthesia to be overrepresented among women (Baron-Cohen, 1996). But it is also worth keeping the female response bias in mind, as emphasized by Simner & Carmichael (2015). For instance, only those participants who were uncertain as to whether they had actually heard the flashes were shown the Saenz & Koch (2008b) self-diagnosis demo. The female response bias could in this case have interacted with implicit experimenter biases. We further also noted that V-EAR experiencers were more than three times likelier than controls to report synaesthetic percepts in their everyday life, independent of V-EAR phenomena. As noted by Simner (2006), comorbidity between varieties is one of the hallmarks of general synaesthetic connectivity. Our measure of the prevalence of synaesthesia beyond that of V-EAR was however not optimal. Indeed, many participants responded that their main strategy was just trying to ‘feel’ the rhythms. Should this be regarded as a form of visual-tactile synaesthesia or just as a figure of speech? As emphasized in the introduction, no one really knows. Taken together, however, our significant female bias and the tendency of comorbidity between V-EAR and other forms of synaesthesia does seem to support the notion that V-EAR should *by definition* be regarded as a form of synaesthesia. As such it might constitute but one of the many ‘tips’ of the multimodal ‘iceberg’.

Like so many other findings from synaesthesia research, our results seem to demand carefully executed replications conducted with larger sample sizes. Only then will we understand why it is that so many people claim to hear the constant dancing of the photons taking place on the receptive fields of their eyes. Little did the natural philosophers of the scientific revolution know, when theorizing about the nature of synaesthesia, that this enigma would continue to perplex science for hundreds of years to come. And it has done so by merit of its conceptual nature, because synaesthesia points to something beyond itself - right into the heart of the nature of perceptual binding processes. Until this more primary perceptual puzzle is solved, even the task of simply defining what synaesthesia really is will most likely remain elusive.

References

- Anstis, S. (1998). Picturing peripheral acuity. *Perception*, 27(7), 817-825.
- Asher, J. E., Lamb, J. A., Brocklebank, D., Cazier, J., Maestrini, E., Addis, L., Monaco (2009). A whole-genome scan and fine-mapping linkage study of auditory-visual synesthesia reveals evidence of linkage to chromosomes 2q24, 5q33, 6p12, and 12p12. *American Journal Of Human Genetics*, 84(2), 279-285.
- Baron-Cohen, S., Burt, L., Smith-Laittan, F., Harrison, J., & Bolton, P. (1996). Synaesthesia: prevalence and familiarity. *Perception*, 25(9), 1073-1079
- Beat, e., & Nicolas, e. (2015). Developing synaesthesia: A primer. *Frontiers In Human Neuroscience*, Vol 9 (2015)
- Brainard, D. H. (1997) The Psychophysics Toolbox, *Spatial Vision* 10:433-436.
- Brang, D., Williams, L., & Ramachandran, V. (2012). Grapheme-color synesthetes show enhanced crossmodal processing between auditory and visual modalities. *Cortex: A Journal Devoted To The Study Of The Nervous System & Behavior*, 48(5), 630-637.
- Bor D, Rothen N, Schwartzman D, Clayton S, Seth A. Adults can be trained to acquire synesthetic experiences. *Scientific Reports*, November 18, 2014; 4:7089.
- Dixon, M. J., Smilek, D., & Merikle, P. M. (2004). Not all synaesthetes are created equal: projector versus associator synaesthetes. *Cognitive, Affective & Behavioral Neuroscience*, 4(3), 335-343.
- Fassnidge, C., Cecconi Marcotti, C., & Freeman, E. (2017). A deafening flash! Visual interference of auditory signal detection. *Consciousness And Cognition*, 49:15-24.

Galton, F. (1880). "Visualized Numerals". *Nature*. 21 (543): 494–5.

Green, D.M., & Sweets, J.A. (1966) *Signal Detection Theory and Psychophysics*, New York: Wiley

Hubbard, E. M., Arman, A. C., Ramachandran, V. S., & Boynton, G. M. (2005). Individual Differences among Grapheme-Color Synesthetes: Brain-Behavior Correlations. *Neuron*, 45975-985.

Hubbard, E. M. (2008). Dispatch: Synaesthesia: The Sounds of Moving Patterns. *Current Biology*, 18R657-R659. doi:10.1016/j.cub.2008.06.035

Hubbard, E.M., Ramachandran, V.S, (2003). The Phenomenology of Synaesthesia. *Journal Of Consciousness Studies*, 10(8), 49.

Hubbard, E. M., & Ramachandran, V. (2005). Neurocognitive Mechanisms of Synesthesia. *Neuron*, 48(3), 509-520.

Jewanski, J., Day, S. A., & Ward, J. (2009). A Colorful Albino: The First Documented Case of Synaesthesia, by Georg Tobias Ludwig Sachs in 1812. *Journal Of The History Of The Neurosciences*, 18(3), 293-303.

Kadosh, R. C., Henik, A., & Walsh, V. (2009). Synaesthesia: learned or lost?. *Developmental Science*, 12(3), 484-491.

Mattingley, J. B., Payne, J. M., & Rich, A. N. (2006). Attention: Attentional Load Attenuates Synaesthetic Priming Effects in Grapheme-Colour Synaesthesia. *Cortex*, 42213-221.

Meier, B., & Rothen, N. (2009). Note: Training grapheme-colour associations produces a synaesthetic Stroop effect, but not a conditioned synaesthetic response. *Neuropsychologia*, 471208-1211.

Noesselt, T., Rieger, J. W., Schoenfeld, M. A., Kanowski, M., Hinrichs, H., Heinze, H., & Driver, J. (2007). Audiovisual temporal correspondence modulates human multisensory superior temporal sulcus plus primary sensory cortices. *The Journal Of Neuroscience: The Official Journal Of The Society For Neuroscience*, 27(42), 11431-11441.

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

Ramachandran, V. S. (2011). *The Tell-Tale brain : A Neuroscientist's Quest For What Makes Us Human*. New York : W.W. Norton

Ramachandran, V.S. (2001). Synaesthesia -- A window into perception, thought and language. *Journal Of Consciousness Studies*, 8(12), 3.

Ramachandran, V. S., & Hubbard, E. M. (2001). Psychophysical Investigations into the Neural Basis of Synaesthesia. *Proceedings: Biological Sciences*, (1470). 979.

Rich, A. N., Bradshaw, J. L., & Mattingley, J. B. (2005). A Systematic, Large-Scale Study of Synaesthesia: Implications for the Role of Early Experience in Lexical-Colour Associations. *Cognition*, 98(1), 53-84.

Rogowska, A. (2011). Categorization of synaesthesia. *Review Of General Psychology*, 15(3), 213-227.

Rouw, R., & Scholte, H. S. (2007). Increased structural connectivity in grapheme-color synesthesia. *Nature Neuroscience*, 10(6), 792-797.

Rouw, R., & Scholte, H. S. (2016). Personality and cognitive profiles of a general synesthetic trait. *Neuropsychologia*, 88(Special Issue: Synaesthesia and Multisensory Processes), 35-48.

Saenz, M., & Koch, C. (2008). The sound of change: visually-induced auditory synesthesia. *Current Biology*, 18(15), R650-R651.

Saenz, M., & Koch, C. (2008b). sciencentral, 11 September 2008, *Hearing motion – synaesthesia diagnosis video*, [Video file], Retrieved from <https://www.youtube.com/watch?v=o39TiACe4mw>

Saenz, M., Lewis, L. B., Huth, A. G., Fine, I., & Koch, C. (2008). Visual Motion Area MT+/V5 Responds to Auditory Motion in Human Sight-Recovery Subjects. *The Journal Of Neuroscience: The Official Journal Of The Society For Neuroscience*, 28(20), 5141-5148.

Simner, J., & Carmichael, D. A. (2015). Is synaesthesia a dominantly female trait?. *Cognitive Neuroscience*, 6(2/3), 68-76.

Simner, J., Mulvenna, C., Sagiv, N., Tsakanikos, E., Witherby, S. A., Fraser, C., & ... Ward, J. (2006). Synaesthesia: the prevalence of atypical cross-modal experiences. *Perception*, 35(8), 1024-1033.

Welch, R. B. (1999). Meaning, attention, and the “unity assumption” in the intersensory bias of spatial and temporal perceptions. *Advances In Psychology*, 129, 371-387.