The number of stimulus-onset asynchronies affects the perception of the sound-induced flash illusion in young and older adults

Jason S. Chan*, Shannon K. Connolly, and Annalisa Setti

School of Applied Psychology, University College Cork, Cork, Ireland

----

*Corresponding Author
jason.chan@ucc.ie

Abstract

The sound-induced flash illusion is a multisensory illusion occurring when one flash is presented with two beeps and perceived as two flashes. Younger individuals are largely susceptible to the illusion when the stimulus onset asynchrony between the first and the second beep falls within the temporal window of integration, but the susceptibility falls dramatically outside of this short temporal range. Older individuals, and particularly older prone to falling and mild cognitive impairment patients, show an extended susceptibility to the illusion. This plausibly indicates that they have inefficient multisensory integration, particularly in the temporal domain. In the present study, we investigated the reliability of the illusion across younger and older people, guided by the hypothesis that the experimental context, i.e. exposure to a wider or smaller number of stimulus onset asynchronies, would modify the intra-personal susceptibility to the illusion at shorter asynchronies vs longer asynchronies, likely due to the gathering of model evidence based on Bayesian inference. We tested 22 young adults and 29 older adults and verified these hypotheses. Both groups showed higher susceptibility to the illusion when exposed to a smaller range of asynchronies, but only for longer ones, not within the 100 ms window. We discuss the thereotical implications in terms of online perceptual learning and practical implications in terms of standardisation of the experimental context when attempting to find normative values.

Introduction

The sound-induced flash illusion (SiFI) is a well-established multisensory phenomenon where a single flash, accompanied by two beeps, in short succession, leads to the perception of two flashes (Shams, Kamitani, & Shimojo, 2000, 2002). This simple illusion has been associated with the size of the temporal binding window (Stevenson et al., 2015; Stevenson & Wallace, 2013). In fact, the occurrence of the illusion is determined by the duration of the stimulus onset asynchrony (SOA) between the beep-flash pair and the second beep (which can precede or follow the other stimuli), providing a good example of the temporal rule of multisensory integration. When the SOA is under 100-150 ms healthy young adults perceive the illusion, while at larger SOAs, the illusion is no longer perceived. This is plausibly due to the fact that the young healthy brain integrate stimuli that are not received synchronously to allow for differences in travel time of sound and light from the output source and different transduction times (Vroomen & De Gelder, 2000, 2004; Vroomen & Keetels, 2010). However, beyond 100-150 ms the difference in travel time and transduction time are considered too great and the visual and auditory stimuli are considered distinct. Assuming that this pattern of susceptibility to the SiFI characterises an efficient multisensory integration of vision and hearing, departures from this pattern have been associated with deficits in integrating information across the senses. Schizophrenic patients (Vanes et al., 2016) and people with autism spectrum disorder (Foss-Feig et al., 2010) have been shown to be differently susceptible to the SiFI compared with healthy controls. Interestingly, healthy older adults also differ from healthy young adults as they are more susceptible to the SiFI compared to their younger counterpart (see also Bedard & Barnett-Cowan, 2016 for additional examples of impaired audiovisual timing in older adults; DeLoss, Pierce, & Andersen, 2013; Setti, Burke, Kenny, & Newell, 2011) indicating that they may present an over-extended temporal window of integration (Setti et al., 2014), as already shown by Diederich, Colonius & Schomburg (2008) with saccadic eye movement in response to visual stimuli with or without sound (Temporal Window of Integration, TWIN model; see also Diederich & Colonius, 2012; Diederich, Colonius, & Kandil, 2016).

Specifically, previous studies have demonstrated that the SiFI is affect by ageing, in general (McGovern, Roudaia, Stapleton, McGinnity, & Newell, 2014; Setti et al., 2011). These studies demonstrate older adults perceiving more illusions compared to younger adults. Additionally, specific functional (Merriman, Whyatt, Setti, Craig, & Newell, 2015; Stapleton, Setti, Doheny, Kenny, & Newell, 2014) and cognitive deficits (Mild Cognitive Impairment

(MCI; Chan et al., 2015) associated with older age have also been associated with increased susceptibility to the SiFI, plausibly indicating more marked temporal discrimination deficits across modalities in these specific populations compared with normal ageing. Temporal discrimination deficits may in fact have a comorbidity awareness of falling (Lupo & Barnett-Cowan, 2017). Considering that specific groups (MCI, older fallers, schizophrenics) display enhanced (and likely inefficient) susceptibility to the SiFI, while presenting very different primary deficits, it is possible to argue that susceptibility to the SiFI may be an indicator of a broad inefficiency in multisensory temporal processing and therefore tapping into temporal integration deficits that may go beyond the integration of simple auditory and visual stimuli. Particularly in relation to older adults, the higher susceptibility to the illusion could be associated with the relatively poor unisensory processing of the visual stimuli, which increases the need to rely on audition for reliable perception. However, this process of integration becomes maladaptive, hence inefficient, when the SOAs are very long. Therefore is emerging evidence that the SiFI could be a simple, reliable, non-verbal, and easy to administer test to assess the efficiency of multisensory temporal integration.

Interestingly, the age difference in the susceptibility to the illusion is specific to the fission (one flash perceived as two flashes when presented with two beeps), as it was not found for the fusion (two flashes perceived as one when presented with one beep) (McGovern, Roudaia, Stapleton, McGinnity, & Newell, 2014). Importantly, as McGovern et al. (2014) suggested, that these differences highlight the need for carefully considering the specific characteristics of each task.

The SiFI has been found to be very robust; for example, participants aware of the illusion and receiving feedback were still susceptible with an SOA of 53 ms (Rosenthal, Shimojo, & Shams, 2009). Nonetheless, when the focus is on the *pattern* of susceptibility across different SOAs, some methodological and theoretical considerations are in order, in view of potentially utilising the SiFI as an assessment tool for efficient multisensory perception in ageing. The reliability within participants, especially for older individuals, remains to be established. Specifically, differences in susceptibility related to the internal experimental context, namely the variety of SOAs tested, should be taken into account when determining what pattern of susceptibility across different SOAs should be expected in different populations.

From the theoretical point of view, the rationale behind most illusions, including the SiFI, is that they violate what we believe occurs outside the laboratory, i.e. that congruent and redundant information is provided by the senses, provided that the stimuli are temporally or spatially coincident (de Gelder & Bertelson, 2003; Groh & Werner-reiss, 2002; Kayser & Shams, 2015). It is suggested that the SiFI is due to a calculation of Bayesian priors, whereby expectations based on prior probabilities, i.e. that two beeps are associated with two flashes (Cuppini, Magosso, Bolognini, Vallar, & Ursino; see Shams & Beierholm, 2010 for a review; Shams, Ma, & Beierholm, 2005). Considering that individuals acquire information during the experiment, i.e. they presumably learn what a flash or two flashes presented with or without beeps, potentially exposing individuals to a wider or smaller set of asynchronies may provide them with different amounts of information, i.e. to gather model evidence, and consequently modifying their susceptibility to the illusion.

The literature on temporal recalibration (Fujisaki, Shimojo, Kashino, & Nishida, 2004; Vatakis & Spence, 2008; Vroomen & De Gelder, 2004) shows that information can be acquired rapidly, supporting the idea that the experimental context may modify the pattern of susceptibility to the illusion. To show how rapidly information can be acquired and utilised, Van der Burg and colleagues (2013) used a simultaneity judgment task consisting of a simple tone and flash with young adults and demonstrated that the point of subjective simultaneity of the stimulus onset was contingent on the modality order of the trial preceding it; such that an audition-led trial would make it more likely that the following audition-led trial would be judged as simultaneous. Furthermore, recalibration was found over a broader range of vision-led SOAs compared to auditory-led ones. This asymmetry is reflective of the natural dominance of vision, which is prudent given that for any audio-visual events occurring ten metres or more away from an individual the visual stimuli will reach the sensory system before auditory stimuli. This indicates that rapid sensory recalibration is most likely to be attributable to a real perceptual effect and not a change in response bias or strategy. Harvey, Van der Burg and Alais (2014) replicated these findings and determined that rapid audio-visual temporal recalibration occurs only for multisensory but not unimodal stimuli.

In sum, it is known that humans quickly adapt to the features of the multisensory stimuli presented, such as in temporal recalibration studies, and this information is likely used to build expectations on what the world looks/sounds like. If we consider that the SiFI is emerging as a useful tool assess temporal integration efficiency, it is of relevance to understand what is the effect of the experimental context, if any, on the susceptibility to the illusion. In light of the above described evidence, in the present study, we assessed whether older and younger participants present a different pattern of susceptibility to the SiFI when exposed to a long (5 SOAs) and short (3 SOAs) version of the task. We hypothesised that no difference should be found at shorter SOAs, 70 ms and 110 ms, because they fall unambiguously within the temporal window of integration; in addition, we hypothesised that, at longer SOAs, in the long version of the task participants would present lower susceptibility to the illusion than in the short version of the task. This hypothesis is driven by the idea that being exposed to 5 asynchronies allows participants to make more accurate inferences, and therefore predictions based on priors, on when a given percept is constituted by one single flash or two flashes. While we do not directly test a Bayesian model here, based on previous literature we propose that the exposure to 5 asynchronies will allow more accurate inference, while not increasing the response bias.

Methods

Participants

Twenty-two young adults (females = 5) between the ages of 18 and 34 years (mean = 25 years) and twenty-nine older adults (females = 18) between the ages of 60 and 81 years (mean = 68 years) took part in this study. Older participants were recruited from active retirement groups and classes for older adults across Cork and Wexford city and county (Ireland). All participants in the final sample self-reported normal or corrected to normal hearing and vision, and were free from neurological illness including epilepsy, any dementia diagnosis, Parkinson's disease, traumatic brain injury and current psychiatric disorders. All older participants had a Mini Mental State Exam (standardised MMSE; Molloy & Standish, 1997) higher than 23 (mean score = 28, range 24 to 30). The MMSE is used to test for symptoms of dementia or cognitive impairment. A MMSE greater than 23 suggests they did not have symptoms of dementia or cognitive impairment.

Ethical approval for this study was obtained by the UCC School of Applied Psychology ethics committee. Participants provided written informed consent for this study and they were not compensated for their participation.

Materials and Stimuli

The experiment was delivered with Neurobehavioral Systems Presentation software (version 17.2) on a 15.6" laptop (Fujitsu Lifebook AH532) with a 1366x768 screen resolution and 60Hz refresh rate. Participants sat at an approximate distance of between 57 and 70 centimetres from the laptop screen. The visual stimulus was a white circular disk, subtending 2° of visual angle. This disk was placed 8° of visual angle below the fixation cross. The presentation duration of the disk was 16 ms. The auditory stimulus consisted of a 16 ms, 3500 Hz pure tone with a total rise- and decay-time of 20 μs at a sound pressure level of 68 dBA.

Procedures

In auditory-lead trials, the first auditory beep was presented; following a variable SOA the second auditory beep was presented, accompanied by a simultaneously presented visual flash. In vision-lead trials, the first beep was accompanied by the visual flash. After a variable SOA, the second auditory beep was presented. The SOAs used were ±70, ±110, ±150, ±190, and ±230, where + indicates vision-lead trials and – indicate auditory-lead trials. There were additional multisensory conditions where 2-beeps and 2-flashes (2 beeps/2 flashes) were presented simultaneously. In the 2 beeps/2 flashes condition, the same SOAs were used. As well as a 1-beep and 1-flash condition, presented simultaneously.

Furthermore, additional unimodal trials were presented: 1-beep, 2-beeps, 1-flash, and 2-flashes. In the 2-beeps and 2-flashes conditions the SOAs between stimuli were the same as the previous conditions. For all conditions, there were four trial repetitions.

The experiment was separated into two separate blocks ('short' and 'long'). All participant took part in both blocks, with the order of presentation counter-balanced between participants. Within each block the presentation of order of auditory-lead, vision-lead and control trials (unimodal and multisensory) was randomized. In the 'short'-block, only three SOAs were presented (±70, ±150, and ±230). In the 'long'-block, all five SOAs were presented. The participant's task was to indicate how many flashes were presented. Each of the blocks was preceded by a short practice.

After each block (short or long) there was a separate auditory-only block, where the beep-only trials were presented. In this block, the participants' task was to indicate how many beeps they perceived. Participant's responses were made via the computer keyboard placed directly in front of them. These trials were presented separately because the participants' task was different compared to the other blocks and may cause task confusion if this task was embedded in the other blocks.

Results

In order to determine if there were sensory differences between the age groups, analyses were conducted on the proportion correct for the unimodal and multisensory control conditions, separately. The unimodal control conditions were analysed using a 2x2x6 mixed-design ANOVA, with Age Group (young vs. older), Duration (long vs. short), and Condition (1-beep; 1-flash; 2-flashes 70 ms; 2-beeps, 70 ms; 2-beeps 150ms; and 2-beeps 230 ms) as factors. There was a significant main effect of Condition [$F(5,245) = 214.66$, $\eta^2 = 0.814$, $p < 0.0001$]. A tukey HSD posthoc revealed performance in the 2 flashes, 70 ms condition (mean proportion correct = 0.14) was significantly worse compared to the 1 beep (mean proportion correct = 0.94) and 1-flash conditions (mean proportion correct = 0.97) (all *p*s < 0.0001). The 2-beeps, 70 condition (mean proportion correct = 0.66) was significantly different from all other unimodal control conditions (see Figure 1; all *p*s < 0.0001). There was a significant interaction between Condition and Group [$F(5, 245) = 5.325$, $\eta^2 = 0.1$, $p = 0.0001$]. Participants, regardless of age group, were significantly worse in the 2-flashes, 70 ms conditions compared to all other conditions, including the older 2-flashes, 70 ms condition (all *p*s < 0.0001; see Figure 1). Also, young adults were significantly worse in the 2 beeps-70 ms condition (mean proportion correct = 0.58) compared to the older adults (mean proportion correct = 0.74; p < 0.0001). In fact, for both age groups, performance was significantly worse in the 2-beeps, 70 ms condition compared to the other unimodal sound-only conditions at longer SOAs (all *p*s < 0.0001). Importantly, there was no main effect of Age Group.

---------------- Place Figure 1 about here ------------------

A 2x2x4 mixed-design ANOVA was conduct on the multisensory control conditions, with Age Group (young vs. older), Duration (short vs. long), and Condition (1-beep/1-flash, 2-beeps/2-flashes 70 ms, 2-beeps/2-flashes 150 ms, and 2-beeps/2-flashes 230 ms) as factors. There was no main effect of Age Group [$F(1,49) = 1.54$, $p = 0.22$]. There was a main effect of Duration [$F(1,49) = 9.15$, $\eta^2 = 0.16$, $p = 0.004$], with better accuracy in the long condition (mean proportion correct = 0.82) compared to the short (mean proportion correct = 0.75). There was also a main effect of Condition [$F(3,147) = 4.07$, $\eta^2 = 0.53$, $p < 0.0001$], with worse performance in the 2-beeps/2-flashes 70 ms condition (mean proportion correct = 0.49) compared to all other conditions (mean proportion correct: 1-beep/1-flash = 0.82; 2-beeps/2-flashes 150 ms = 0.82; 2-beeps/2-flashes 230 ms = 0.87; all *p*s < 0.0001). There was a significant interaction between Duration and Age Group [$F(1,49) = 6.22$, $\eta^2 = 0.11$, $p = 0.02$]. Young adults were significant better when there was an increased number of SOAs presented (long; mean proportion correct = 0.83) compared to when there were few (short: mean proportion correct = 0.70; $p = 0.003$). Finally, there was a significant interaction between Duration and Condition [$F(3,147) = 4.52$, $\eta^2 = 0.08$, $p = 0.005$] (see Figure 2). Performance in the 2-beeps/2-flashes 70 ms condition (regardless of duration) was significantly worse compared to all other conditions (all *p*s < 0.005). In the 2-beeps/2-flashes 230 ms condition, participants were significantly more accurate in the long condition compared to the short ($p = 0.005$).

--------------- Place Figure 2 about here -----------------

The perception of the SiFI (i.e. perceiving two flashes when one was actually presented) was assessed using the measure of sensitivity, whereby the proportion of illusions is considered as false alarms: d' (Signal Detection Theory). This parameter measures sensitivity as a function of the probability of hits [P(H) = hits/(hits+misses)] and the probability of false alarms [P(FA) = false alarms/(false alarms+correct rejections)], and is independent of the bias toward either category. It is defined by the formula: d'= z[P(H)]) - z[P(FA)] (Macmillan & Creelman, 1991). Specifically, sensitivity to the illusion was calculated as: d' = z[P(2-beeps/2-flashes)] – z[P(2-beeps/1-flash)]. Lower d' indicate greater susceptibility to perceiving the illusion as opposed to perceiving two real flashes when presented.

Separate 2x2x3 mixed-design ANOVAs were conducted on the vision-lead and auditory-lead conditions. The factors were Age Group (young vs. older), Duration (long vs. short), and SOA (70 ms, 150 ms, and 230 ms) as conditions.

In the auditory-lead condition, there was a main effect of Duration [$F(1,49) = 8.58$, $\eta^2 = 0.15$, $p = 0.005$] with larger d' when five SOAs were presented (long = 1.45) compared to when three SOAs were presented (short = 0.87). There was also a significant main effect of

SOA [$F(2,98) = 115.20$, $\eta^2 = 0.52$, $p < 0.0001$]. A Tukey HSD posthoc revealed significantly larger d' in the -230 ms condition (mean d' = 1.90) compared to the -150 ms (mean d' = 1.65) and -70 ms (mean d' = -0.067) conditions (all *p*s = 0.0001). There was no significant difference between the -70 ms and -150 ms conditions. There was a trend for a main effect of Age Group [$F(1,49) = 3.29$, $\eta^2 = 0.06$, $p = 0.08$], with the trend for young adults to have higher d' than the older adults (mean d's: young = 1.48; older = 0.84). There were no other significant effects or interactions. See Figure 3 for a graph of the effects.

The ratio between the 1-flash and 2-beeps/1-flash conditions remained consistent between the short and long duration conditions. However, it is possible that there was a shift in the response bias due to the increased number of SOAs, especially 190 ms which may be associated with low susceptibility to the illusion. To test for response bias, another 2x2x3 mixed-design ANOVA was conducted on the beta-values. C-criteria was calculated [$c = 0.5*(z(H)+z(FA))$]. There was no main effects of Age Group [$F(1,49) < 1$, *n.s.*] and a near main effect of Duration [$F(1,49) = 3.59$, $p = 0.06$]. There was a main effect of SOA, with increased response bias in the 70 ms condition to respond 2-flashes compared to the 230 ms conditions (mean β-values: -230 ms = -0.63; -150 ms = -0.21; -70 ms = 0.33; $p < 0.0005$). There was also a significant interaction of Duration X SOA. Participants had increased bias to respond 2-flashes in the 230 ms SOA, Long Duration condition compared to all other conditions (all ps < 0.05). There were no other significant interactions.

In the vision-lead condition, once again there was a main effect of Duration [$F(1,49) = 12.49$, $\eta^2 = 0.20$, $p < 0.005$], with higher d' in the five SOA condition (long; mean d' = 1.80) compared to the three SOA condition (short; mean d' = 1.04). There was no main effect of Age Group [$F(1,49) = 2.97$, $p = 0.09$]. There was also a significant main effect of SOA [$F(2,98) = 48.60$, $\eta^2 = 0.50$, $p < 0.0001$]. Overall, participants were significantly more accurate in the 230 ms condition (mean d' = 2.278) compared to the 150 ms (mean d' = 1.86) and 70 ms (mean d' = 0.12) conditions (all *p*s = 0.0001). There was a significant interaction between the factors of Duration and SOA [$F(2,98) = 4.06$, $\eta^2 = 0.08$, $p = 0.02$]. A Tukey HSD posthoc revealed no significant difference between the two duration conditions at the 70 ms SOA. However, participants were significantly more accurate in the long duration block compared to the short duration block, when the SOA was 150 ms ($p = 0.001$) and 230 ms ($p = 0.0002$) (see Figures 3 & 4).

Once again, in order to determine if there was response bias between conditions and groups, another 2x2x3 mixed design ANOVA was conducted on the c-criteria values. There was a main effect of Age Group, with older adults having and overall increased bias to respond 2-flashes compared to young adults (mean β-values: older = -0.35; young = 0.23) [$F(1,49) = 7.00$, $p = 0.01$]. There were no other significant main effects or interactions suggesting that their responses biases did not change between Duration conditions.

--------------- Place Figures 3 and 4 about here -----------------

Discussion

Previous studies have suggested that the SiFI can be explained by Bayesian probability (Cuppini et al.; Shams, 2012). Studies have also demonstrated that the SiFI is an early sensory illusion (Mishra, Martinez, Sejnowski, & Hillyard, 2007; Shams, 2012; Shams, Iwaki, Chawla, & Bhattacharya, 2005) that is resistant to direct feedback training, suggesting that it is largely resistant to explicit top-down influence (Rosenthal et al., 2009; cf. Setti & Chan, 2011; Setti et al., 2014), although some transfer of learning effects have been shown by Setti et al. (2014) and Powers et al. (2016). While providing direct feedback on the SiFI may not affect the number of perceived illusions, in this study we find that varying the number of SOAs does affect the susceptibility to it.

These data demonstrate that the number of variable SOAs presented affects the perceived illusion. According to our hypotheses, while the sensitivity to illusions remained relatively consistent at the 70 ms SOA, as the number of SOAs increased participants perceived fewer illusions when a larger number of SOAs was available (‘long’ block). We argue that this occurs because participants are using their perceptions from the additional SOAs to acquire information, which, then can be used to reduce their perception of the illusion when appropriate, i.e. at longer SOAs. When there were relatively few SOAs, participants had fewer SOAs to use as a comparison. This occurs specifically with visual-lead stimuli, the reason why that is the case remains to be established, and it is likely to be related to the fact that vision leading events are more likely to occur in the real world, therefore requiring prompt adjustment by the senses. In fact, Powers, Hillock and Wallace (2009) found that audio-visual perceptual training was more effective in reducing the temporal binding window for vision-led compared to audio-led stimuli. It is to note that while the effect of perceptual training, either synchrony judgement or temporal order judgement, on the susceptibility to the illusion has provided mixed evidence of success (Setti et al., 2014; Powers et al., 2016), the positive result of this study in reducing susceptibility in the long vs. the short block could be due to the different information provided in the perceptual training tasks and in the present study. In the perceptual training tasks, the participant is provided with feedback on the synchrony/asynchrony of one beep and one flash, while in our study, participants are exposed to different synchronous and asynchronous unisensory and multisensory stimuli (1 real flash, 2 real flashes, 2 real flashes with 2 beeps, 1 flash with two beeps). While they do not receive explicit feedback on their performance, we would argue that they have a larger amount of input available to gather model evidence.

The present findings have relevant theoretical and practical implications. From a theoretical point of view, we show that both younger and older adults are able to gather information during the ‘long’ block, and refine their multisensory perceptual abilities accordingly. While Bayesian model is not directly tested here, we suggest that the effects shown here are due to the gathering of model evidence utilised to predict the occurrence of the illusion with different SOAs (Cuppini et al.; Kayser & Shams, 2015; Shams, 2012; Shams, Ma, et al., 2005).

Unlikely, previous studies (DeLoss et al., 2013; McGovern et al., 2014; Setti et al., 2011) we did not find a main effect of Age Group between young and older participants (although there is a trend). In the studies by McGovern et al. (2014) and Setti et al. (2011), six or seven SOAs were used, respectively. To determine if these data could be replicated, a separate 2x5 mixed design ANOVA was performed on the ‘long’ duration condition only as this is most akin to the previous studies. Indeed, there is a main effect of Age Group, with older adults perceiving more illusions (i.e., lower d’s; mean = 1.41) compared to younger adults (mean d’ = 2.19) [$F(1,49) = 5.19$, $p = 0.03$]. This replicates the previous work showing older adults perceiving more illusions compared to young adults. A similar analysis was conducted on the ‘short’ duration condition and revealed no main effect of age Group [$F(1,49) < 1$, *n.s.*]. These results further highlight the importance of the present study. This discrepancy may in fact explain some differences found between studies (e.g. Setti et al., 2011 and Stapleton et al., 2015) where a larger or smaller number of SOAs was utilised. It is important to note that older adults were significantly worse when responding to the 2-flashes, 70 ms condition. Furthermore, older adults were more accurate in the 2-beeps, 70 ms condition compared to young adults. These results help to explain why older adults perceive illusions. Older adults have increased their temporal reliability to the auditory modality, relative to the visual modality. These results can help to explain why older adults perceived more illusions compared to young adults. However, it is important to note that the unimodal results do not explain the difference in performance between the two Durations conditions (short vs. long). Similarly, response bias towards responding to having perceived two flashes was higher in older than in younger, but this difference did not interact with the block (long or short).

On applied grounds, the SiFI has been liked to many cognitive and physical deficits such as: general ageing (DeLoss & Andersen, 2015; DeLoss et al., 2013), falling (Setti et al., 2011; Setti et al., 2014), balance maintenance (Merriman et al., 2015; Stapleton et al., 2014), autism spectrum disorder (Foss-Feig et al., 2010), and mild cognitive impairment (Chan et al., 2015). We argued that it is reasonable to assume that the SiFI is linked to a fundamental mechanism that underlies these disorders with different patterns of comorbidities. It is possible that with the further understanding of these disorders the SiFI can be used as part of a non-verbal diagnostic tool for inefficient multisensory deficits. To do that, clear guidelines must be established so that normative results can be determined. However, to establish diagnostic guidelines a defined fixed-set experimental context should be established or, at the very least, potential differences between different experimental contexts should be taken into account when interpreting results. These differences may also offer insight into the specific deficits captured by the SiFI in different populations; for example the partial difference in findings found by Setti et al. (2011), where seven SOAs were used, and Merriman et al. (2015) were three SOA were used, in relation to falls and susceptibility to the illusion could be explained in light of the present results.

The main limitation of the present study is constituted by the fact that we can only infer indirectly that the change in susceptibility to the SiFI between the two versions of the test is due to the acquisition of more information in the longer than in the shorter version allowing for more refined predictions on whether vision and hearing should be merged into a unified percept. More direct evidence would be constituted by a design allowing for Bayesian modelling. In addition, with a larger number of trials, one could explore whether more accurate inferences build over time, as one would expect to find wider differences between the ‘long’ and ‘short’ duration blocks towards the end, more than at the beginning of the test. Recording of neurophysiological activity during the two tasks would also constitute an important tool to assess whether this evidence is built over time and determines lower susceptibility to the SiFI when 5 rather than 3 SOAs are available.

In sum, the present findings contribute in showing that the SiFI is robust to experimental manipulations and reliable to test/re-test within the same individual, both for younger and older, with short SOAs clearly falling within the temporal window of integration. They also show that information on the perceptual characteristics of the experiment can be gathered online during the test of the illusion, producing a refinement of multisensory integration, i.e. fewer illusions, when more information is available in terms of a larger number of SOAs. This implies that healthy older adults maintain the ability to update their perceptual abilities according to the situation at hand. It also shows that in establishing normative values in view of utilising the SiFI as tool to assess temporal multisensory integration, it is necessary to standardise the experimental context in relation to the optimal number of SOAs to highlight the deficits to be explored.

References

Bedard, G., & Barnett-Cowan, M. (2016). Impaired timing of audiovisual events in the elderly. *Experimental Brain Research, 234*(1), 331-340. doi:10.1007/s00221-015-4466-7

Chan, J. S., Kaiser, J., Brandl, M., Matura, S., Prvulovic, D., Hogan, M. J., & Naumer, M. J. (2015). Expanded temporal binding windows in people with mild cognitive impairment. *Current Alzheimer Research, 12*(1), 61-68.

Cuppini, C., Magosso, E., Bolognini, N., Vallar, G., & Ursino, M. A neurocomputational analysis of the sound-induced flash illusion. *NeuroImage*(0). doi:http://dx.doi.org/10.1016/j.neuroimage.2014.02.001

de Gelder, B., & Bertelson, P. (2003). Multisensory integration, perception and ecological validity. *Trends in Cognitive Sciences, 7*(10), 460-467.

DeLoss, D. J., & Andersen, G. J. (2015). Aging, Spatial Disparity, and the Sound-Induced Flash Illusion. *PLoS ONE, 10*(11), e0143773. doi:10.1371/journal.pone.0143773

DeLoss, D. J., Pierce, R. S., & Andersen, G. J. (2013). Multisensory Integration, Aging, and the Sound-Induced Flash Illusion. *Psychology and Aging*.

Diederich, A., & Colonius, H. (2012). Modeling Multisensory Processes in Saccadic Responses: Time-Window-of-Integration Model. In M. M. Murray & M. Wallace (Eds.), *The Neural Bases of Multisensory Processes*. Boca Raton (FL): CRC Press.

Diederich, A., Colonius, H., & Kandil, F. I. (2016). Prior knowledge of spatiotemporal configuration facilitates crossmodal saccadic response. *Experimental Brain Research, 234*(7), 2059-2076. doi:10.1007/s00221-016-4609-5

Diederich, A., Colonius, H., & Schomburg, A. (2008). Assessing age-related multisensory enhancement with the time-window-of-integration model. *Neuropsychologia, 46*(10), 2556-2562. doi:10.1016/j.neuropsychologia.2008.03.026

Foss-Feig, J. H., Kwakye, L. D., Cascio, C. J., Burnette, C. P., Kadivar, H., Stone, W., & Wallace, M. T. (2010). An extended multisensory temporal binding window in autism spectrum disorders. *Experimental Brain Research, 203*(2), 381-389.

Fujisaki, W., Shimojo, S., Kashino, M., & Nishida, S. y. (2004). Recalibration of audiovisual simultaneity. *Nature Neuroscience, 7*(7), 773-778.

Groh, J. M., & Werner-reiss, U. (2002). Visual and Auditory Integration. In V. S. R. Editor-in-Chief: (Ed.), *Encyclopedia of the Human Brain* (pp. 739-752). New York: Academic Press.

Harvey, C., Van der Burg, E., & Alais, D. (2014). Rapid temporal recalibration occurs crossmodally without stimulus specificity but is absent unimodally. *Brain Research, 1585*, 120-130. doi:http://dx.doi.org/10.1016/j.brainres.2014.08.028

Kayser, C., & Shams, L. (2015). Multisensory causal inference in the brain. *PLoS Biology, 13*(2), e1002075. doi:10.1371/journal.pbio.1002075

Lupo, J., & Barnett-Cowan, M. (2017). Perceived timing of a postural perturbation. *Neuroscience Letters, 639*, 167-172. doi:http://dx.doi.org/10.1016/j.neulet.2016.12.055

Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide*. Cambridge, UK: Cambridge University Press.

McGovern, D. P., Roudaia, E., Stapleton, J., McGinnity, T. M., & Newell, F. N. (2014). The sound-induced flash illusion reveals dissociable age-related effects in multisensory integration. *Frontiers in Aging Neuroscience, 6*(250). doi:10.3389/fnagi.2014.00250

Merriman, N. A., Whyatt, C., Setti, A., Craig, C., & Newell, F. N. (2015). Successful balance training is associated with improved multisensory function in fall-prone older adults. *Computers in human behavior, 45*, 192-203. doi:10.1016/j.chb.2014.12.017

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. *Journal of Neuroscience, 27*(15), 4120-4131.

Molloy, D. W., & Standish, T. I. (1997). A guide to the standardized Mini-Mental State Examination. *Int Psychogeriatr, 9*(suppl 1), 87-94.

Powers, A. R., III, Hillock-Dunn, A., & Wallace, M. T. (2016). Generalization of multisensory perceptual learning. *Scientific Reports, 6*.

Powers, A. R., III, Hillock, A. R., & Wallace, M. T. (2009). Perceptual Training Narrows the Temporal Window of Multisensory Binding. *Journal of Neuroscience, 29*(39), 12265-12274. doi:10.1523/jneurosci.3501-09.2009

Rosenthal, O., Shimojo, S., & Shams, L. (2009). Sound-Induced Flash Illusion is Resistant to Feedback Training. *Brain Topography, 21*(3), 185-192. doi:10.1007/s10548-009-0090-9

Setti, A., Burke, K. E., Kenny, R. A., & Newell, F. N. (2011). Is inefficient multisensory processing associated with falls in older people? *Experimental Brain Research, 2209*, 375-284.

Setti, A., & Chan, J. S. (2011). Familiarity of objects affects susceptibility to the sound-induced flash. *Neuroscience Letters, 492*(1), 19-22.

Setti, A., Stapleton, J., Leahy, D., Walsh, C., Kenny, R. A., & Newell, F. N. (2014). Improving the efficiency of multisensory integration in older adults: Audio-visual temporal discrimination training reduces susceptibility to the sound-induced flash illusion. *Neuropsychologia, 61*, 259-268. doi:http://dx.doi.org/10.1016/j.neuropsychologia.2014.06.027

Shams, L. (2012). Early Integration and Bayesian Causal Inference in Multisensory Perception. In M. M. Murray & M. Wallace (Eds.), *The Neural Bases of Multisensory Processes*. Boca Raton (FL): CRC Press.

Shams, L., & Beierholm, U. R. (2010). Causal inference in perception. *Trends in Cognitive Sciences, 14*(9), 425-432.

Shams, L., Iwaki, S., Chawla, A., & Bhattacharya, J. (2005). Early modulation of visual cortex by sound: an MEG study. *Neuroscience Letters, 378*(2), 76-81. doi:DOI: 10.1016/j.neulet.2004.12.035

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

Shams, L., Kamitani, Y., & Shimojo, S. (2002). Visual illusion induced by sound. *Cognitive Brain Research, 14*, 147-152.

Shams, L., Ma, W. J., & Beierholm, U. R. (2005). Sound-induced flash illusion as an optimal percept. *Neuroreport, 16*(17), 1923-1927.

Stapleton, J., Setti, A., Doheny, E. P., Kenny, R. A., & Newell, F. N. (2014). A standing posture is associated with increased susceptibility to the sound-induced flash illusion in fall-prone older adults. *Experimental Brain Research, 232*(2), 423-434. doi:10.1007/s00221-013-3750-7

Stevenson, R. A., Segers, M., Ferber, S., Barense, M. D., Camarata, S., & Wallace, M. T. (2015). Keeping time in the brain: Autism spectrum disorder and audiovisual temporal processing. *Autism Res*. doi:10.1002/aur.1566

Stevenson, R. A., & Wallace, M. T. (2013). Multisensory temporal integration: task and stimulus dependencies. *Experimental Brain Research, 227*(2), 249-261. doi:10.1007/s00221-013-3507-3

Van der Burg, E., Alais, D., & Cass, J. (2013). Rapid Recalibration to Audiovisual Asynchrony. *The Journal of Neuroscience, 33*(37), 14633-14637. doi:10.1523/jneurosci.1182-13.2013

Vanes, L. D., White, T. P., Wigton, R. L., Joyce, D., Collier, T., & Shergill, S. S. (2016). Reduced susceptibility to the sound-induced flash fusion illusion in schizophrenia.

Psychiatry Research, 245, 58-65. doi:http://dx.doi.org/10.1016/j.psychres.2016.08.016

Vatakis, A., & Spence, C. (2008). Evaluating the influence of the 'unity assumption'on the temporal perception of realistic audiovisual stimuli. *Acta Psychologica, 127*(1), 12-23.

Vroomen, J., & De Gelder, B. (2000). Crossmodal integration: A good fit is no criterion. *Trends in Cognitive Science, 4*(2), 37-38.

Vroomen, J., & De Gelder, B. (2004). Temporal ventriloquism: Sound modulates the flash-lag effect. *Journal of Experimental Psychology: Human Perception & Performance, 30*(3), 513-518.

Vroomen, J., & Keetels, M. (2010). Perception of intersensory synchrony: A tutorial review. *Attention, Perception & Psychophysics, 72*(4), 871-884.

Acknowledgements

We would like to thank Sean Kelleher, Maeve McLoughlin, Amy Jordan for conducting the experiment with the young participants. Part of this work was presented as the final year project for the BA in Applied Psychology by Shannon K. Connolly and it was presented as poster at the meeting of Irish Geriatrics Society 2016.

Figure Captions

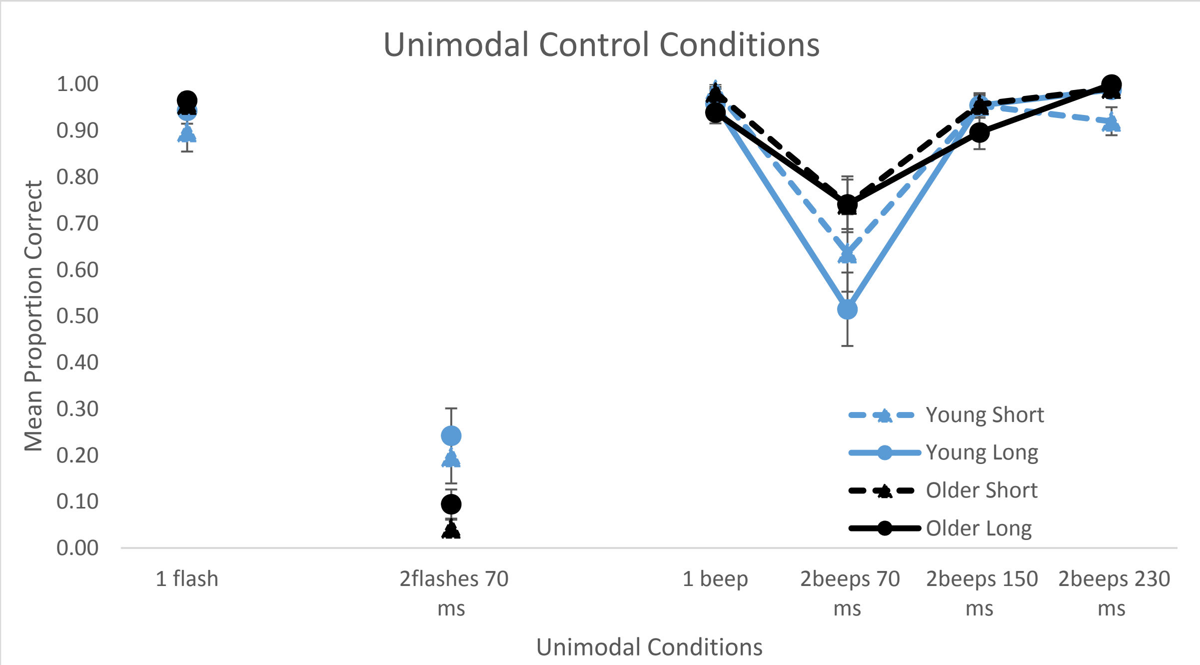
1. Mean proportion correct in the unimodal control conditions for young and older adults. This figure represents only the SOAs that were presented in both the 'short' and 'long' conditions. The error bars represent the standard error of the means.

2. Mean proportion correct in the multisensory control conditions for young and older adults. This figure represents only the SOAs that were presented in both the 'short' and 'long' conditions. The error bars represent the standard error of the means.

3. Interaction between Duration and SOA in the 2-beeps/1-flash, vision-lead condition. Participants were significantly more accurate in the 'long' condition compared to the 'short', when the SOAs were 150 ms or 230 ms.

4. Mean d' for the 2-beeps/1-flash for young and older adults across all SOAs. To note that 5 SOAs are presented for the 'long' condition, while only the 3 available SOAs are presented for the 'short' condition.

Figure 1.



Unimodal Control Conditions
Mean Proportion Correct
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
1 flash
2flashes 70 ms
1 beep
2beeps 70 ms
2beeps 150 ms
2beeps 230 ms
Unimodal Conditions
Young Short
Young Long
Older Short
Older Long

Figure 2.

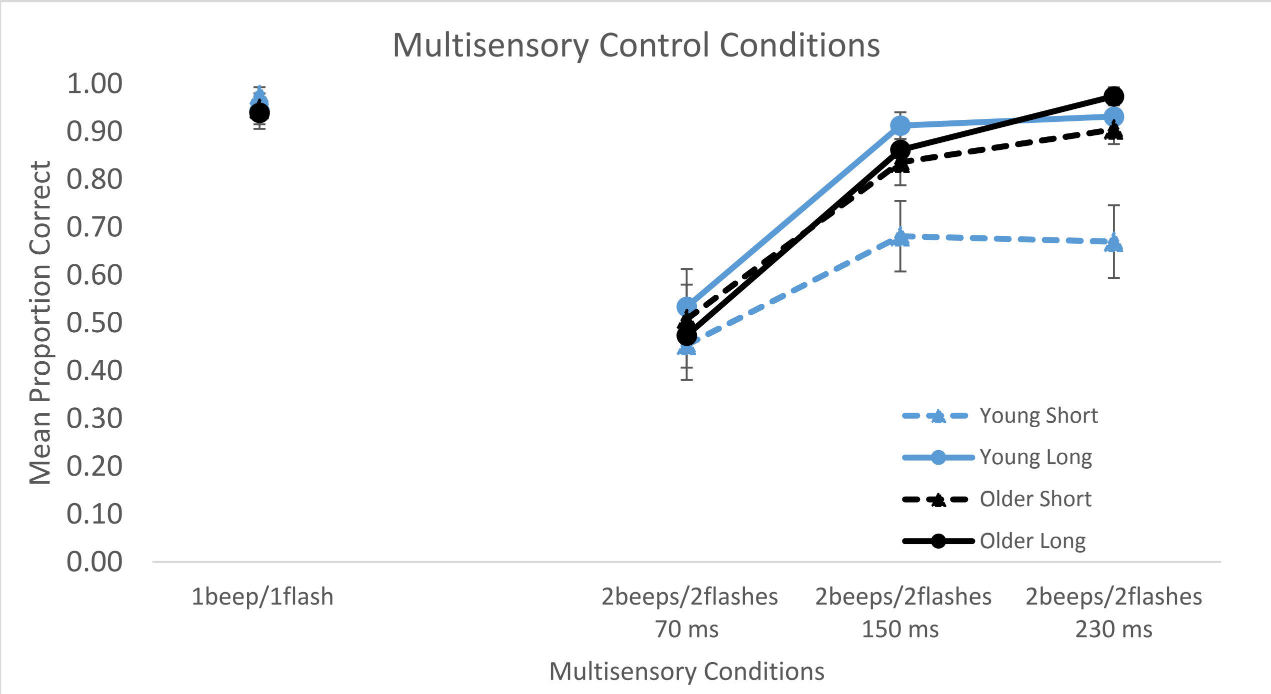

Multisensory Control Conditions
Mean Proportion Correct
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
1beep/1flash
2beeps/2flashes 70 ms
2beeps/2flashes 150 ms
2beeps/2flashes 230 ms
Multisensory Conditions
Young Short
Young Long
Older Short
Older Long

Figure 3.

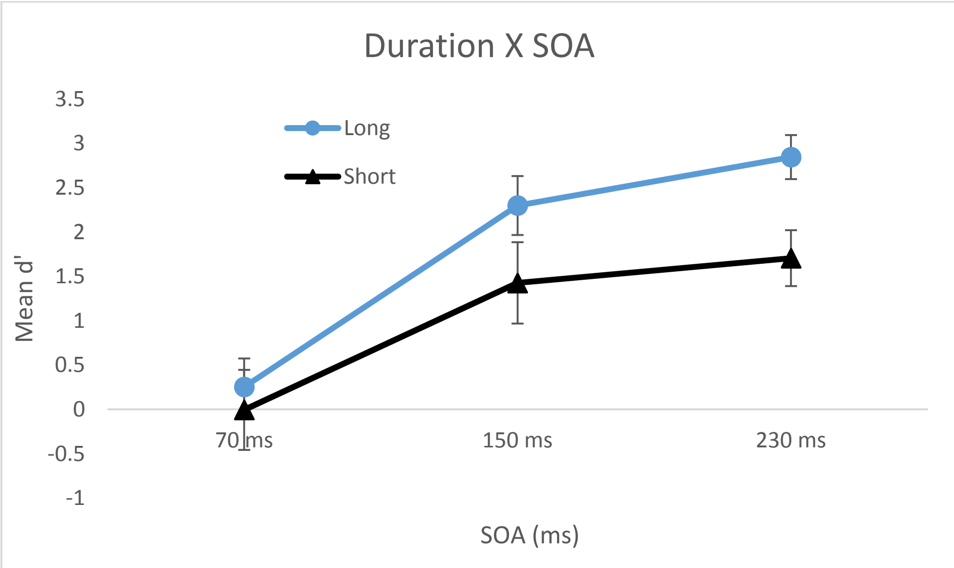
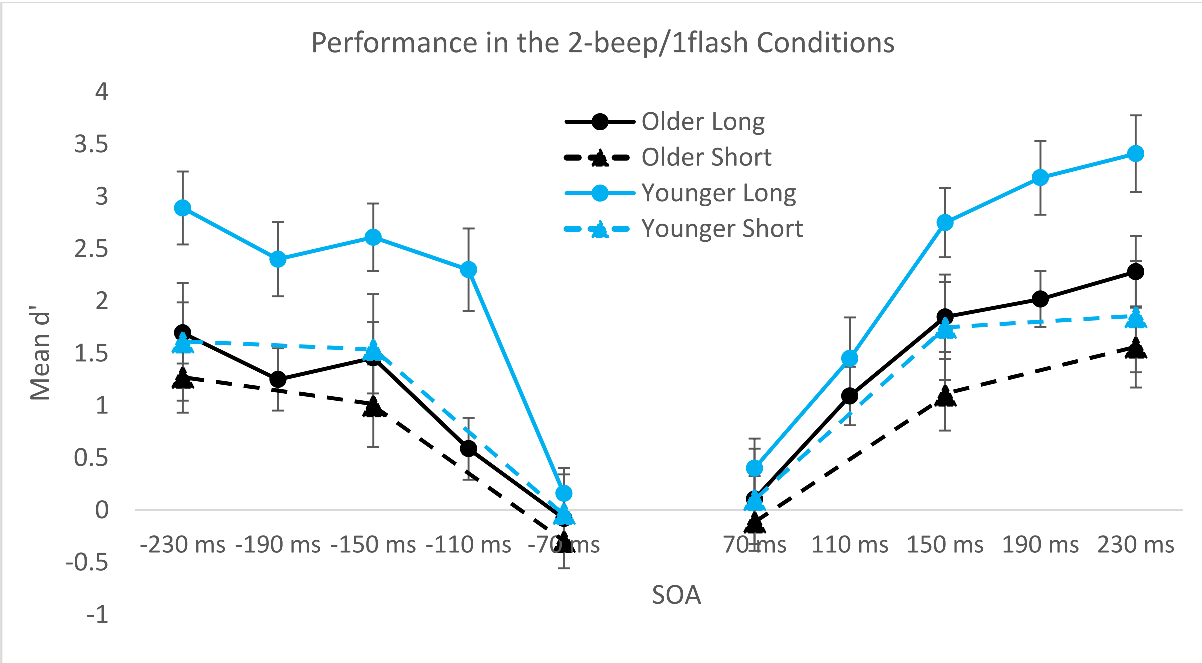

Duration X SOA
Long
Short
3.5
3
2.5
2
1.5
1
0.5
0
-0.5
-1
Mean d'
70 ms
150 ms
230 ms
SOA (ms)

Figure 4.


Performance in the 2-beep/1flash Conditions
Older Long
Older Short
Younger Long
Younger Short
Mean d'
4
3.5
3
2.5
2
1.5
1
0.5
0
-0.5
-1
-230 ms
-190 ms
-150 ms
-110 ms
-70 ms
70 ms
110 ms
150 ms
190 ms
230 ms
SOA