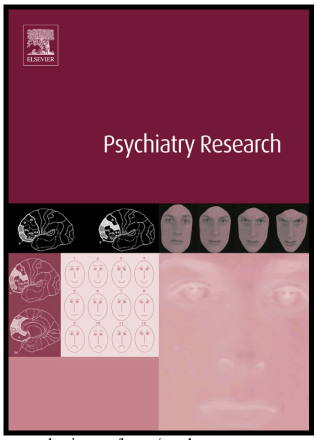
Author's Accepted Manuscript



www.elsevier.com/locate/psychres

Reduced susceptibility to the sound-induced flash fusion illusion in schizophrenia

Lucy D. Vanes, Thomas P. White, Rebekah L. Wigton, Dan Joyce, Tracy Collier, Sukhi S. Shergill

PII: S0165-1781(16)30103-2
DOI: http://dx.doi.org/10.1016/j.psychres.2016.08.016
Reference: PSY9887

To appear in: *Psychiatry Research*

Received date: 18 January 2016
Revised date: 12 July 2016
Accepted date: 5 August 2016

Reduced susceptibility to the sound-induced flash fusion illusion in schizophrenia

Lucy D. Vanes[a], Thomas P. White[a,b], Rebekah L. Wigton[a], Dan Joyce[a], Tracy Collier[a], and Sukhi S. Shergill[a]

[a]Institute of Psychiatry, Psychology and Neuroscience, de Crespigny Park, London, SE5 8AF, United Kingdom

[b]School of Psychology, University of Birmingham, Edgbaston, Birmingham, B15 2TT, United Kingdom

Abstract

Schizophrenia is characterised by the presence of abnormal complex sensory perceptual experiences. Such experiences could arise as a consequence of dysfunctional multisensory integration. We used the sound-induced flash illusion paradigm, which probes audiovisual integration using elementary visual and auditory cues, in a sample of individuals with schizophrenia (n=40) and matched controls (n=22). Signal detection theory analyses were performed to characterise patients' and controls' sensitivity in distinguishing 1 and 2 flashes under varying auditory conditions. Both groups experienced significant fission illusions (whereby one visual flash, accompanied by two auditory beeps, is misperceived as two flashes) and fusion illusions (whereby two flashes, accompanied by one beep, are perceived as one flash). Patients showed significantly lower fusion illusion rates compared to HC, while the fission illusion occurred similarly frequently in both groups. However, using an SDT approach, we compared illusion conditions with unimodal visual conditions, and found that illusory visual perception was overall more strongly influenced by auditory input in HC compared to patients for *both* illusions. This suggests that multisensory integration may be impaired on a low perceptual level in SZ.

Key words: schizophrenia; multisensory integration; fission illusion; fusion illusion

1. Introduction

The richness of human experience is based on the integration of different sensory stimuli. This integration is largely implicit and serves to facilitate perception and understanding in a complex sensory environment. Schizophrenia (SZ) is characterised by a disintegration of common multimodal experiences (Postmes et al., 2014) and has been conceptualised as a disorder of the normal connectivity and integration within the brain (Friston and Frith, 1995; Roiser et al., 2013; Stephan et al., 2009; White et al., 2010). The research on perceptual processing of low-level stimuli in SZ and other non-affective psychotic disorders has naturally focussed on single sensory systems, identifying subtle dysfunctions in visual, auditory and sensorimotor processing (Butler and Javitt, 2005; Javitt and Freedman, 2015; Näätänen and Kähkönen, 2009; Shergill et al., 2014), with a relative neglect of the integration of these elementary processes into the multimodal context in which the sensory experiences in psychosis often arise. Such multimodal experiences make it probable that symptoms such as hallucinations and delusions are not primarily a function of the failure within a single sensory system, but rather of an inappropriate interaction between different sensory modalities. This suggests that an understanding of multisensory integration (MSI) in non-affective psychosis may contribute to a better understanding of the clinical symptoms. In using low-level stimuli, perceptual deficits can be disentangled more readily from higher level dysfunctions, e.g. involving social cognition or learning processes.

MSI can also give rise to illusory phenomena in healthy perception. Multisensory illusions are characterised by a binding of information from different modalities in order to create a coherent unified percept, which is typically inconsistent with the true sensory input. For example, in the Ventriloquist illusion, an auditory stimulus is misattributed to the wrong source if that source provides temporally contingent visual information (Howard and Templeton, 1966). Similarly in the McGurk illusion, incongruent visual and auditory phonetic information is fused into the percept of an alternative, illusory phoneme (McGurk and MacDonald, 1976). Importantly, while non-veridical in nature, these illusions can be seen as optimal percepts given the ambiguous sensory input and learned expectations about multimodal stimuli (Alais and Burr, 2004; Körding et al., 2007). They therefore constitute markers for intact multimodal processing as seen in healthy individuals, while offering a framework in which to investigate aberrant sensory fusion in psychosis (White et al., 2014).

Outside of the illusion literature, a dysfunction of multisensory integration in SZ has been shown in the form of reduced facilitation of speech processing by visual articulatory motion (Ross et al., 2007), and on a more elementary level with reduced reaction time facilitation in response to bimodal detection targets as compared to unimodal targets (Williams et al., 2010). Emotional voice information has further been shown to exert a reduced influence on a visual emotion categorization task in SZ (de Gelder et al., 2005). Recently, Zvyagintsev et al. (2013) demonstrated that patients with SZ show an interference effect of auditory information on a visual discrimination task both in an emotional and basic perceptual paradigm, whereas controls did not show this effect on the basic perceptual level. It is interesting to note that in the interference paradigm, an increased multimodal interaction, or "leakage" between modalities, is observed in patients compared to controls when using low-level audiovisual stimuli. Importantly, this is an attentional effect which influences the response rather than perception. In contrast, illusion paradigms predict a *reduction* in multisensory interactions in patients. In these paradigms, perception itself is altered by multimodal information, but when MSI breaks down, perception remains veridical, albeit non-optimal. Accordingly, within the multisensory illusion paradigm, patients with SZ have been shown to exhibit reduced susceptibility to the McGurk effect (de Gelder et al., 2003; White et al., 2014). However, research on susceptibility to multimodal illusions in SZ is sparse and typically makes use of high-level social or emotional stimuli. There is therefore little evidence on MSI in SZ from illusion paradigms using low-level stimulus integration.

In the current study, we used the sound-induced flash illusion (Shams et al., 2000) to examine auditory-visual integration on a very basic perceptual level in non-affective psychosis. Our sample consisted largely of patients with a diagnosis of SZ, but also included a small number of patients with schizoaffective disorder. These subgroups did not differ in terms of their symptoms, thus we shall henceforth refer to SZ as a whole for simplicity. In the sound-induced flash illusion paradigm, different numbers of brief flashes and beeps are presented contemporaneously. The classic illusion occurs when one flash accompanied by two beeps is erroneously perceived as two flashes (*fission* illusion). Conversely, two flashes are often perceived as one when they are accompanied by a single beep (*fusion* illusion; Andersen et al., 2004). Crucially, this task avoids potential confounds due to biological motion, phonetics, or social aspects often present in audiovisual paradigms (de Gelder et al., 2003; de Gelder et al., 2005; Ross et al., 2007; White et al., 2014). Previous research has shown that modulation of visual cortex activity in early stages of sensory processing is at least partially involved in eliciting the sound-induced flash illusion (Shams et al., 2005), suggesting that low level sensory processes play an important role for this illusion. It has been suggested that auditory input modulates visual cortical activation via direct pathways, resulting in the perception of an illusion. This paradigm is therefore useful in order to examine whether auditory input modulates early visual cortical processing to the same degree in SZ as it does in healthy perception.

We implemented analyses from Signal Detection Theory (SDT) in order to disentangle perceptual sensitivity from more general perceptual biases in characterising the sound-induced flash illusions. The choice of SDT measure was made in order to remain consistent with existing literature on the sound-induced flash illusion. However, there has been recent criticism of the use of SDT measures in this context (Witt et al., 2015); specifically, it has been pointed out that the criterion measure (also known as response bias) is frequently inaccurately interpreted as an internal decision criterion. We therefore emphasize that the criterion measure as used in the current study does not necessarily reflect a decisional response rule, but may indeed reflect a perceptual process. Consequently, when we use the term "tendency", we do not exclude the possibility that this may be perceptually driven – indeed we believe this to be more likely than an internal decision criterion.

Based on previous research showing that patients with SZ show reduced susceptibility to the McGurk effect (White et al., 2014) and reduced MSI (de Gelder et al., 2003; Williams et al., 2010), we hypothesised that patients would exhibit attenuated fission and fusion illusions relative to HC in our study. The sensitivity index d' was used to indicate whether illusions were due to attenuated sensitivity which was specific to illusion trials, and the criterion measure ln(β) was used to indicate whether auditory information created a general (likely perceptual) bias which was unspecific to the illusion trials.

2. Material and Methods

2.1 Participants

Forty individuals diagnosed with a diagnosis of non-affective psychosis according to ICD-10 and 22 HC matched for age, sex, handedness, and socioeconomic background participated. Thirty-six patients had a diagnosis of schizophrenia and 4 patients had a diagnosis of schizoaffective disorder. Patients diagnosed with schizoaffective disorder did not differ from patients diagnosed with schizophrenia in terms of symptoms, and exclusion of these patients from analysis did not alter the statistical results of this study. Intelligence quotient (IQ) was measured with the two-item Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999). Thirty-eight patients were taking prescribed anti-psychotic medication at the time of the study and two patients were stable off their medication for over six weeks. Chlorpromazine (CPZ) equivalent doses of medications were calculated using conversion tables (Bazire, 2005; Woods, 2003). Demographic characteristics are presented in Table 1. Exclusion criteria for all subjects were a history of neurological illness, current major physical illness, and drug dependency over the last six months. Exclusion criteria for HC were a history of psychiatric illness and a first-degree relative currently or previously suffering from a psychotic illness. All subjects had normal hearing and normal or corrected-to-normal vision. Ethical approval was provided by Central London Research and Ethics Committee. All participants provided informed written consent and were compensated for their time and travel.

2.2 Stimuli

The task was conducted as part of a larger magnetic resonance imaging (MRI) study. The fMRI data is not presented here, as it captures a different set of processes involved in this task. Visual stimuli were presented on a screen viewed via a head-mounted mirror (refresh rate 60Hz). Each flash (F) was a white disk (diameter: 1.6°; duration: 16.66ms; eccentricity: 4° right off-centre) presented against a black background. Beeps (B) were presented via headphones at a volume permitting dissociation of the tone from background noise (frequency: 480Hz; duration 16ms). 0, 1 or 2 flashes were presented alongside 0, 1 or 2 beeps, resulting in nine possible stimulus combinations (trial types). Each trial type was presented 24 times in a randomized order, resulting in 216 trials. On congruent trials (F1B1 and F2B2), flashes and beeps had identical onsets. On incongruent trials (F1B2 and F2B1), the single stimulus in one modality was presented symmetrically in between the first and second onset of the stimuli in the other modality. The inter-stimulus-interval (ISI) between two stimuli of the same modality was determined on a single-subject level based on unimodal visual performance.

2.3 Determination of the inter-stimulus-interval (ISI)

A dual staircase procedure was used prior to the task to assess the minimal gap (as a multiple of screen frames, i.e. 16.66ms) needed to be able to reliably discern two flashes. This procedure was conducted outside the MRI scanner in quiet conditions. The two staircases were interleaved such that each odd trial used a randomly selected staircase and each even trial used the respective other staircase. The procedure was ended when two minimum inversions had occurred on both staircases (or if two flashes were consistently identified on three consecutive trials at the minimum staircase value, i.e. minimal gap). The threshold was set as the largest of all minimum inversion values that occurred. Two screen frames were subsequently added to this threshold, and the resulting duration set as a consistent ISI throughout the flash illusion task. We hereby aimed to maximise the chance that two flashes could correctly be identified as such in viewing conditions inside the scanner.

2.4 Procedure

On each trial, participants were required to fixate a permanently visible, central white cross, which turned red 500 ms after stimulus onset as a response cue. Upon response, it changed back to white. Participants were instructed to attend to the flashes and disregard beeps, and to indicate via button press how many flashes (0, 1, or 2) they had seen on each trial. Incongruent trials were conditions of interest for the occurrence of illusions: F1B2 trials were potential *fission illusion* trials (a fission occurring when two flashes were reported), and F2B1 trials were potential *fusion illusion* trials (a fusion occurring when one flash was reported).

2.3 Statistical analyses

Accuracy was calculated as proportion correct for each trial type and compared between groups using non-parametric Mann-Whitney tests.

For the SDT analysis, trials on which subjects responded “0” were excluded from all further analyses. This ensured that calculated response rates were based on two response alternatives of interest and summed to 1 in a given condition, which allows for comprehensive comparison of SDT measures. Furthermore, trials on which 0 flashes were presented were excluded from further analyses as these were not of primary interest.

SDT measures were calculated similarly to previous reports (Watkins et al., 2006; Whittingham et al., 2014). A baseline condition was defined as trials on which no beeps were presented (F1B0 and F2B0), the fission condition was defined as trials on which two beeps were presented (F1B2 and F2B2), and the fusion condition was defined as trials on which one beep was presented (F1B1 and F2B1). For each of these conditions, sensitivity measure d’ and criterion measure ln(β) (Stanislaw and Todorov, 1999) were calculated and baseline-illusion difference scores compared using mixed ANOVAs.

Despite the difficulties in interpreting SDT measures adequately, the use of sensitivity and criterion measures are useful in that they allow a distinction between the sensitivity between perceiving one or two flashes (under varying auditory conditions), and an overall tendency to perceive one or two flashes within a condition. Witt et al. (2015) correctly state that the sound-induced flash illusion may be reflected in the criterion measure (e.g. an overall increased tendency to perceive and respond “2” when two beeps are presented will lead to an increase in fission illusions on F1B2 trials), however this effect will not be specific to the illusion trial (i.e. there will also be increased “2” responses on F2B2 trials). In contrast, sensitivity d’ disentangles these two effects by contrasting them directly.

Sensitivity is defined as z(hit rate)-z(false alarm rate), where z is the inverse cumulative normal function. Criterion ln(β) was defined as $\frac{z(false\ alarm)^2 - z(hit)^2}{2}$. Negative values indicate a tendency towards responding "signal present", whereas positive values indicate a tendency towards responding "signal absent". A log linear transform was applied before computing hit and false alarm rates (adding 0.5 each to the number of hits and false alarms, and adding 1 each to the total number of signal trials and no-signal trials) in order to avoid extreme values of d' (e.g. when the number of hits or false alarms is 0). The definition of SDT terms for each condition was as follows:

2.3.1 Fission condition (F1B2 and F2B2 trials)

Two flashes were defined as the signal. Therefore, a response of "2" on F2B2 trials was a *hit*, while a response of "1" was a miss. Conversely, a response of "2" on F1B2 trials was a *false alarm* (or fission illusion), and a response of "1" was a correct rejection. Sensitivity d' reflects the ability to distinguish one and two flashes, under the condition that two beeps are presented. A positive ln(β) reflects the overall tendency towards responding "1".

2.3.2 Fusion condition (F1B1 and F2B1 trials)

One flash was defined as the signal. Therefore, a response of "1" on F1B1 trials was a *hit*, and a response of "2" was a miss. A response of "1" on F2B1 trials was a *false alarm* (or fusion illusion) and a response of "2" was a correct rejection. Sensitivity d' reflects the ability to distinguish one and two flashes, under the condition that one beep is presented. A positive ln(β) reflects the overall tendency towards responding "2". However for the further analysis, ln(β) in this condition was inverted to represent the tendency towards responding "1" in order to more easily compare to the fission and baseline conditions.

2.3.3 Baseline condition (F1B0 and F2B0 trials)

The definition of either one or two flashes as the signal is arbitrary. Sensitivity d', indicating the ability to distinguish one and two flashes when no beeps are presented, will be identical regardless of this definition. Criterion ln(β), reflecting the overall tendency towards one particular response, will only differ with respect to its sign. We chose to define two flashes as the signal (in line with the fission illusion); hence positive ln(β) will reflect the overall tendency to respond "1" on these trials.

Susceptibility to each illusion was quantified as the difference in d' between the baseline and respective illusion condition. Overall biasing due to the presentation of beeps was quantified as the difference in ln(β) between the baseline and illusion conditions. Finally, d'-difference and ln(β)-difference were computed between the fission and fusion conditions for completeness.

Correlation analyses were also performed for baseline-illusion d'-difference scores with clinical and behavioural variables of interest.

3. Results

3.1 ISI

The staircase procedure resulted in a marginally longer mean ISI in patients (M = 75ms, SD = 28.33ms) compared to HC (M = 65ms, SD = 13.33), $t(59.5) = -1.90$, $p = .062$.

3.1 Raw data

The response rates for each of the three possible answers (0, 1, or 2 flashes seen) in each of the nine conditions are depicted in Figure 1. Accuracy was above 70% on unimodal trials as well as trials with equal flashes and beeps. Both groups showed average performance above chance level (33%) in all of these trial types, as assessed with one-sided Mann-Whitney tests, all ps < .005. Across all trial types, mean accuracy did not differ between patients (M = 0.74, SD = 0.11) and HC (M = 0.76, SD = 0.09), $W(60) = 488$, $p > 0.05$. Average performance when excluding F1B2 and F2B1 trials also did not differ between patients (M = 0.82, SD = 0.13) and HC (M = 0.87, SD = 0.09), $W(60) = 543$, $p > 0.05$.

Looking at accuracy in each trial type separately, uncorrected Mann-Whitney tests showed reduced accuracy in patients compared to HC only in the F0B0 condition, $W(60) = 589$, $p = 0.028$. This difference does not survive Bonferroni correction for multiple comparisons.

As is evident from the Figure 1, fission rates (response of "2" on F1B2 trials) did not significantly differ between groups (HC: M = 0.54, SD = 0.25, Patients: M = 0.57, SD = 0.24), $W(60) = 412$,

$p = 0.686$. In contrast, for the fusion illusion (response of "1" on F2B1 trials), patients showed lower rates compared to HC (HC: $M = 0.55$, $SD = 0.26$, Patients: $M = 0.40$, $SD = 0.26$), $W(60) = 591$, $p = 0.028$ (uncorrected).

Having removed trials on which a response of 0 was given, and recalculating response rates for the following SDT analysis, group differences in illusion rates did not change as compared to the results reported above.

3.2 SDT analysis

3.2.1 Baseline vs. illusion

We calculated d'-difference and ln(β)-difference scores for the baseline-fission and baseline-fusion comparisons. These are depicted in Figure 2. Difference scores for d' and ln(β) were subjected to separate mixed ANOVAs with Group (HC vs. Patient) as a between-subject factor and Illusion (baseline-fission vs. baseline-fusion) as a within subject factor.

The analysis of d'-difference scores revealed a significant main effect of Group, with patients showing lower d'-difference scores overall, $F(1,60) = 4.914$, $p = 0.030$. This indicates that HC sensitivity in distinguishing one and two flashes changed more drastically in HC than patients from baseline to fission condition, and well as baseline to fusion condition. In other words, adding either one or two beeps had a greater detrimental effect on controls' ability to correctly discern one from two flashes, while patients were less affected by auditory input. There was also a significant main effect of Illusion, $F(1,60) = 15.421$, $p < 0.001$, with lower d'-difference scores in the baseline-fusion comparison compared to the baseline-fission comparison. This indicates that, across groups, sensitivity was more strongly affected when adding two beeps (fission condition) compared to when adding one beep (fusion condition). There was no significant interaction, $F(1,60) = 0.289$, $p = 0.593$.

The analysis of ln(β)-difference scores showed no significant effect of Group, $F(1,60) = 0.855$, $p = 0.359$, but a highly significant effect of Illusion $F(1,60) = 104.112$, $p < 0.001$. This indicated that while subjects showed a similar response tendency in baseline and fusion conditions, it differed between the baseline and fission condition. The positive ln(β)-difference scores indicate that subjects were more likely to respond “2” when two beeps were presented (fission condition) compared to when no beeps were presented (baseline condition). This change in criterion ln(β) was larger in HC compared to patients, as seen in a significant interaction effect, $F(1,60) = 6.266$, $p = 0.015$.

3.2.2 Fission vs. Fusion

As there has been general criticism of the comparison of illusion conditions with the no-beep condition as a baseline, but not with the condition with the alternative number of beeps (Witt et al., 2015), we conduct this comparison for completeness. Comparing the fission condition to the one-beep condition, and comparing the fusion condition to the two-beep condition, is identical to comparing the fission to the fusion condition.

There was no significant group difference in d’-difference for the fission-fusion comparison, $t(60) = 0.537$, $p = 0.593$, with subjects showing higher sensitivity in the fusion (1-beep) compared to the fission (2-beep) condition (d’-difference scores across groups were significantly smaller than 0, $t(61) = -3.95$, $p < 0.001$).

Groups differed significantly in ln(β)-difference scores for the fission-fusion comparison, $t(60) = -2.503$, $p = 0.015$, with HC showing more negative values. Though both groups had a stronger tendency to respond “1” in the fusion (1-beep) and “2” in the fission (2-beep) condition, this effect was stronger in HC. Note that this difference would remain significant after p-value correction accounting for the two previous comparisons (baseline-fission and baseline-fusion) performed in the above ANOVA.

3.4 Associations with other variables

Neither PANSS score, nor duration of illness or CPZ equivalent medication dosage were significantly correlated with d’- difference scores for any of the condition comparisons, all *p*s > 0.05.

In addition, given that all subjects were assigned an individual ISI stemming from the visual thresholding procedure, we tested for correlations between ISI and all of the reported sensitivity measures. ISI was significantly negatively correlated with both d’-difference scores for the baseline-fission and baseline-fusion comparisons, however including ISI as a covariate into the reported ANOVA did not alter any of the reported effects.

Similarly smokers and non-smokers within and across groups did not differ in their d'-difference scores, all *p*s > 0.05, and including smoking status as a further predictor in the ANOVA did not alter any of the reported effects.

Finally, we tested for correlations between d'-difference scores for the baseline-fission and baseline-fusion comparison in each group. There was a significant positive correlation in the patient group, indicating that susceptibility to the fusion illusion was associated with susceptibility to the fission illusion ($R = .422$, $p = 0.007$). This correlation was absent in healthy controls, $p > 0.05$. The difference in correlation was statistically significant, as confirmed using Fisher's z-test, $z = 2.17$, $p = 0.03$.

4. Discussion

The aim of this study was to identify differences in multisensory processing between individuals with SZ and healthy individuals using the sound-induced flash illusion paradigm. We found that patients showed significantly lower fusion illusion rates compared to HC, while the fission illusion occurred similarly frequently in both groups. However, using an SDT approach, we were able to show that susceptibility to both illusions, defined as a reduced ability to distinguish one and two flashes when auditory beeps are presented concurrently (compared to when there are no beeps), was reduced in patients compared to HC in *both* illusions. In other words, although absolute fission illusion rates did not differ between groups, HC experienced a greater impact of auditory beeps on illusion rates compared to the unimodal baseline.

A potential limitation of this study is that the experiment was conducted inside an MR scanner, thus it is possible that findings would differ when conducted outside the scanner. MR noise may have distracted patients and controls differentially; however if one assumed increased distractibility in patients, we would expect this to lead to less veridical responses and therefore potentially an increase in illusory perception.

The finding of reduced multisensory illusory perception in patients with SZ is consistent with previous findings of reduced perception of the McGurk effect in SZ (de Gelder et al., 2003; White et al., 2014), which has been attributed to reduced MSI. Dysfunctional MSI in SZ has also been shown in non-illusory contexts (de Gelder et al., 2005; Ross et al., 2007; Williams et al., 2010). Hyper-integration of multisensory stimuli has also been reported in the literature, particularly using interference paradigms (de Gelder et al., 2005; Zvyagintsev et al., 2013), suggesting that deficits in MSI manifest differently depending on the experimental context. In this particular setup, we were able to show disrupted integration of the auditory and visual channel in low-level perceptual processes in SZ.

The patients in our sample showed lower absolute fusion illusion rates compared to HC. As the occurrence of the fusion illusion naturally depends on the temporal resolution of visual perception (and, thus, the ability to reliably identify two rapid flashes), it stands to reason that patients may have superior visual perceptual resolution. This, however, is contradicted by the fact that in our experiment the patient group on average required longer ISIs than controls as determined by a visual thresholding procedure, indicating in fact a coarser temporal resolution in patients' visual perception. In turn, while the length of the ISI was associated with individual sensitivity to the illusions, it did not account for the reported group differences. Due to the association of ISI with illusory susceptibility, we cannot exclude the possibility that this contributed in part to the findings; this therefore constitutes a limiting factor. However, if it were the driving force behind the findings we would expect group differences to disappear when including ISI as a covariate into the ANOVA, which was not the case. We therefore conclude that a potential difference in temporal resolution with respect to visual perception did not cause the effects, and favour the interpretation that audiovisual interactions indeed differ between groups.

We also found that the fission illusion was in general a stronger effect than the fusion illusion, relative to unimodal viewing conditions. Some fundamental differences have been established between the fission and fusion illusion in previous research: Increased susceptibility to the fission illusion, relative to the fusion illusion, has been reported in healthy samples in previous studies (Andersen et al., 2004;

Bolognini et al., 2011; Innes-Brown and Crewther, 2009). Distinct neural mechanisms are suggested to underlie the fusion and fission illusions, as noted in studies using electroencephalography (Mishra et al., 2008; Mishra et al., 2007; Shams et al., 2001), functional magnetic resonance imaging (Watkins et al., 2007; Watkins et al., 2006), and magnetoencephalography (Shams et al., 2005). Specifically, while both illusion types are associated with analogous activation in multimodal superior temporal cortex, primary visual cortex shows increased activation during illusory fissions but decreased activation during illusory fusions relative to physically identical trials which do not result in illusory perception. A study which investigated the influence of transcranial direct current stimulation (tDCS) on the sound-induced flash illusions showed that the fission illusion was modulated by application of tDCS to temporal or occipital cortex, whereas the fusion illusion remained unaltered by this regional stimulation (Bolognini et al., 2011). This suggests that not only is sound-induced flash fusion a more difficult illusion to elicit, but it is also harder to perturb or modulate. These findings support the notion that the fusion illusion is a perceptually more challenging illusion. As a result, abnormal perception is likely to surface more readily in this illusion rather than the fission illusion, particularly in a population which is known to show reduced susceptibility to complex perceptual illusions as a whole (such as patients with schizophrenia). Crucially, however, we found that those patients who were least susceptible to the fusion illusion also tended to be less susceptible to the fission illusion, suggesting that a single underlying mechanism may govern abnormal illusory perception.

In terms of group differences, it is essential to note that these are unlikely to be a result of generally lower task performance in patients. First, patients showed overall high performance on trials which do not evoke illusory perception. Second, and more importantly, reduced illusion rates are in fact a result of more correct responses, indicating that patients were performing the task well and in part with higher accuracy than controls. Seeing as patients with schizophrenia frequently perform worse on wide-ranging experimental tasks, making it difficult to dissociate general task performance effects from more specific cognitive and perceptual processes, this work demonstrates that illusion paradigms offer a valuable way to study psychosis – whereby patients respond veridically, albeit non-optimally.

Besides the group difference in sensitivity, HC also showed a greater overall bias in the two-beep (fission) condition compared to patients. This demonstrates that the addition of two beeps not only reduced HC's ability to distinguish one and two flashes more than patients, but it also caused a greater overall bias towards perceiving two flashes. This effect thus includes a facilitation of perceiving two veridical flashes when two beeps are presented contemporaneously, providing additional evidence for a stronger MSI effect in HC.

In this study, we cannot conclusively state whether the differences in MSI are due to purely perceptual aberrations. As the paradigm is free of complex high-level stimuli, we can largely exclude effects of higher level cognitive processes. However, it is unclear whether more basic cognitive processes such as attention, potentially interacting with early sensory processing, contribute to disruptions in MSI. Further research is necessary in order to clarify this point.

We found a positive correlation between fusion and fission illusion susceptibility in patients. In contrast, susceptibility to the illusions appeared to be independent in healthy individuals. On these grounds, we argue that healthy perception of the two illusions is driven by distinct underlying mechanisms, whereas abnormal perception of both illusions in psychosis is caused by the same underlying dysfunction in multisensory integration. It is possible that this common dysfunction is driven either on the perceptual or cognitive level; however this will need to be addressed in future research. This study is novel in that it demonstrates on an elementary level that this MSI deficit in SZ can manifest more strongly under certain perceptual conditions than others. In summary, patients with schizophrenia demonstrate a deficit in audio-visual integration which results in reduced multisensory illusions and perceptual biasing.

S.S.S. is supported by a European Research Council Consolidator Award (grant number 311686) and some of this work was supported by a Medical Research Council New Investigator award to S.S.S (grant number G0901868), and developed by the National Institute for Health Research (NIHR) Mental Health Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London (PI: Simon Lovestone; 2013). The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

References

Alais, D., Burr, D., 2004. The ventriloquist effect results from near-optimal bimodal integration. Curr. Biol. 14(3), 257-262.
Andersen, T.S., Tiippana, K., Sams, M., 2004. Factors influencing audiovisual fission and fusion illusions. Brain Res. Cogn. Brain Res. 21(3), 301-308.
Bazire, S., 2005. Psychotropic drug directory. The professionals' pocket handbook and aide memoire, Drug treatment options in psychiatric illness. Dementia, 40-48.
Bolognini, N., Rossetti, A., Casati, C., Mancini, F., Vallar, G., 2011. Neuromodulation of multisensory perception: a tDCS study of the sound-induced flash illusion. Neuropsychologia 49(2), 231-237.
Butler, P.D., Javitt, D.C., 2005. Early-stage visual processing deficits in schizophrenia. Current opinion in psychiatry 18(2), 151.
de Gelder, B., Vroomen, J., Annen, L., Masthof, E., Hodiamont, P., 2003. Audio-visual integration in schizophrenia. Schizophr. Res. 59(2), 211-218.
de Gelder, B., Vroomen, J., de Jong, S.J., Masthoff, E.D., Trompenaars, F.J., Hodiamont, P., 2005. Multisensory integration of emotional faces and voices in schizophrenics. Schizophr. Res. 72(2-3), 195-203.
Friston, K.J., Frith, C.D., 1995. Schizophrenia: a disconnection syndrome. Clin. Neurosci. 3(2), 89-97.
Howard, I.P., Templeton, W.B., 1966. Human spatial orientation. Wiley, New York.
Innes-Brown, H., Crewther, D., 2009. The impact of spatial incongruence on an auditory-visual illusion. PLoS One 4(7), e6450.
Javitt, D.C., Freedman, R., 2015. Sensory processing dysfunction in the personal experience and neuronal machinery of schizophrenia. Am. J. Psychiatry 172(1), 17-31.
Körding, K.P., Beierholm, U., Ma, W.J., Quartz, S., Tenenbaum, J.B., Shams, L., 2007. Causal inference in multisensory perception. PLoS One 2(9), e943.
McGurk, H., MacDonald, J., 1976. Hearing lips and seeing voices. Nature 264, 746-748.
Mishra, J., Martinez, A., Hillyard, S.A., 2008. Cortical processes underlying sound-induced flash fusion. Brain Res. 1242, 102-115.
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. J. Neurosci. 27(15), 4120-4131.
Näätänen, R., Kähkönen, S., 2009. Central auditory dysfunction in schizophrenia as revealed by the mismatch negativity (MMN) and its magnetic equivalent MMNm: a review. Int. J. Neuropsychopharmacol. 12(1), 125-135.
Postmes, L., Sno, H., Goedhart, S., van der Stel, J., Heering, H., de Haan, L., 2014. Schizophrenia as a self-disorder due to perceptual incoherence. Schizophr. Res. 152(1), 41-50.
Roiser, J.P., Wigton, R., Kilner, J.M., Mendez, M.A., Hon, N., Friston, K.J., Joyce, E.M., 2013. Dysconnectivity in the frontoparietal attention network in schizophrenia. Frontiers in psychiatry 4.

Ross, L.A., Saint-Amour, D., Leavitt, V.M., Molholm, S., Javitt, D.C., Foxe, J.J., 2007. Impaired multisensory processing in schizophrenia: deficits in the visual enhancement of speech comprehension under noisy environmental conditions. Schizophr. Res. 97(1), 173-183.
Shams, L., Iwaki, S., Chawla, A., Bhattacharya, J., 2005. Early modulation of visual cortex by sound: an MEG study. Neurosci. Lett. 378(2), 76-81.
Shams, L., Kamitani, Y., Shimojo, S., 2000. Illusions: What you see is what you hear. Nature 408(6814), 788-788.
Shams, L., Kamitani, Y., Thompson, S., Shimojo, S., 2001. Sound alters visual evoked potentials in humans. Neuroreport 12(17), 3849-3852.
Shergill, S.S., Samson, G., Bays, P.M., Frith, C.D., Wolpert, D.M., 2014. Evidence for sensory prediction deficits in schizophrenia. Am. J. Psychiatry 162(12), 2384-2386.
Stanislaw, H., Todorov, N., 1999. Calculation of signal detection theory measures. Behav. Res. Methods Instrum. Comput. 31(1), 137-149.
Stephan, K.E., Friston, K.J., Frith, C.D., 2009. Dysconnection in schizophrenia: from abnormal synaptic plasticity to failures of self-monitoring. Schizophr. Bull. 35(3), 509-527.
Watkins, S., Shams, L., Josephs, O., Rees, G., 2007. Activity in human V1 follows multisensory perception. Neuroimage 37(2), 572-578.
Watkins, S., Shams, L., Tanaka, S., Haynes, J.D., Rees, G., 2006. Sound alters activity in human V1 in association with illusory visual perception. Neuroimage 31(3), 1247-1256.
Wechsler, D., 1999. WASI manual. San Antonio, Psychological Corporation.
White, T.P., Joseph, V., Francis, S.T., Liddle, P.F., 2010. Aberrant salience network (bilateral insula and anterior cingulate cortex) connectivity during information processing in schizophrenia. Schizophr. Res. 123(2), 105-115.
White, T.P., Wigton, R.L., Joyce, D.W., Bobin, T., Ferragamo, C., Wasim, N., Lisk, S., Shergill, S.S., 2014. Eluding the illusion? Schizophrenia, dopamine and the McGurk effect. Front. Hum. Neurosci. 8, 565.
Whittingham, K.M., McDonald, J.S., Clifford, C.W., 2014. Synesthetes show normal sound-induced flash fission and fusion illusions. Vision Res. 105, 1-9.
Williams, L.E., Light, G.A., Braff, D.L., Ramachandran, V.S., 2010. Reduced multisensory integration in patients with schizophrenia on a target detection task. Neuropsychologia 48(10), 3128-3136.
Witt, J.K., Taylor, J.E.T., Sugovic, M., Wixted, J.T., 2015. Signal detection measures cannot distinguish perceptual biases from response biases. Perception 44(3), 289-300.
Woods, S.W., 2003. Chlorpromazine equivalent doses for the newer atypical antipsychotics. J. Clin. Psychiatry 64(6), 663-667.
Zvyagintsev, M., Parisi, C., Chechko, N., Nikolaev, A.R., Mathiak, K., 2013. Attention and multisensory integration of emotions in schizophrenia.

Table 1. *Sample sizes (N), means (M), and standard deviations (SD) of demographic and illness characteristics of the study sample*

	HC			Patients			Between-group test statistics	
	N	*M*	*SD*	*N*	*M*	*SD*		
Age	22	36.4	9.1	40	37.0	8.8	$t(60) = -0.24$	$p = 0.811$
NS-SEC	22	2.4	1.7	40	2.8	1.6	$t(60) = -0.85$	$p = 0.400$
WASI	22	107.8	14.2	40	97.8	14.1	$t(60) = 2.66$	$p = 0.010$
Female (%)		14			20		$\chi^2 = 0.08$	$p = 0.779$
Smokers (%)		27			75		$\chi^2 = 11.39$	$p < 0.001$
Onset age (years)				39	23.6	5.6		
Illness duration (years)				39	13.4	8.8		
CPZ equivalents				38	487.0	389.0		
PANSS score								
Positive symptoms				38	16.2	4.5		
Negative symptom				38	18.1	5.5		
General symptoms				38	31.2	7.1		
Total score				38	65.5	14.3		

NS-SEC: National Statistics Socio-Economic Classification

WASI: Wechsler Abbreviated Scale of Intelligence

CPZ: Chlorpromazine

PANSS: Positive and Negative Syndrome Scale for Schizophrenia

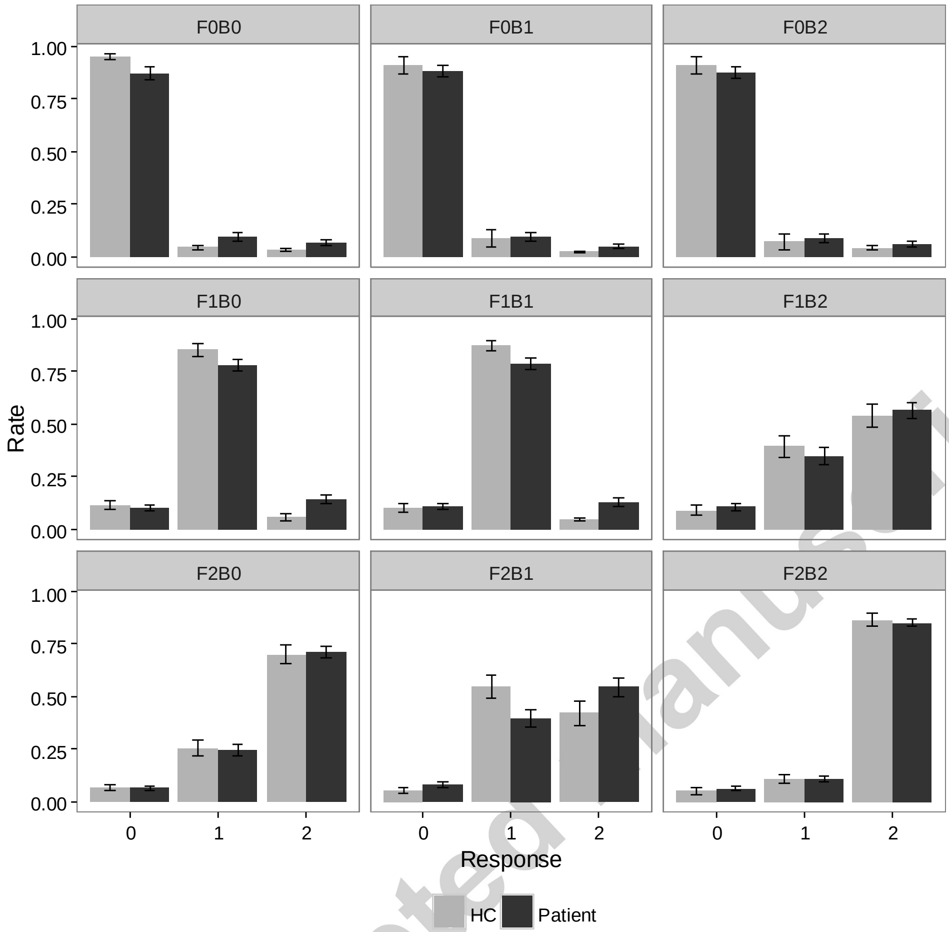


Figure 1. Response rates for each of the three possible answers (0, 1, or 2 flashes seen), in each of the nine trial types per group. Error bars represent the standard error of the mean.

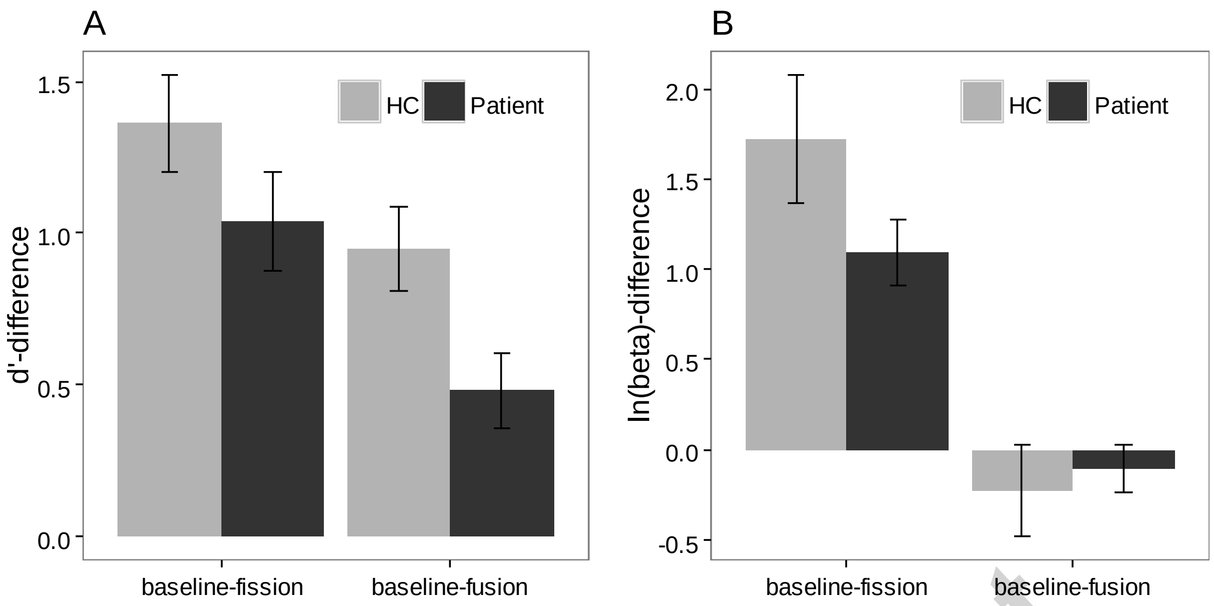


Figure 2. Mean sensitivity d' difference score (A) and mean ln(β) difference score (B) for the baseline-fission and baseline-fusion comparisons per group. Error bars represent the standard error of the mean.

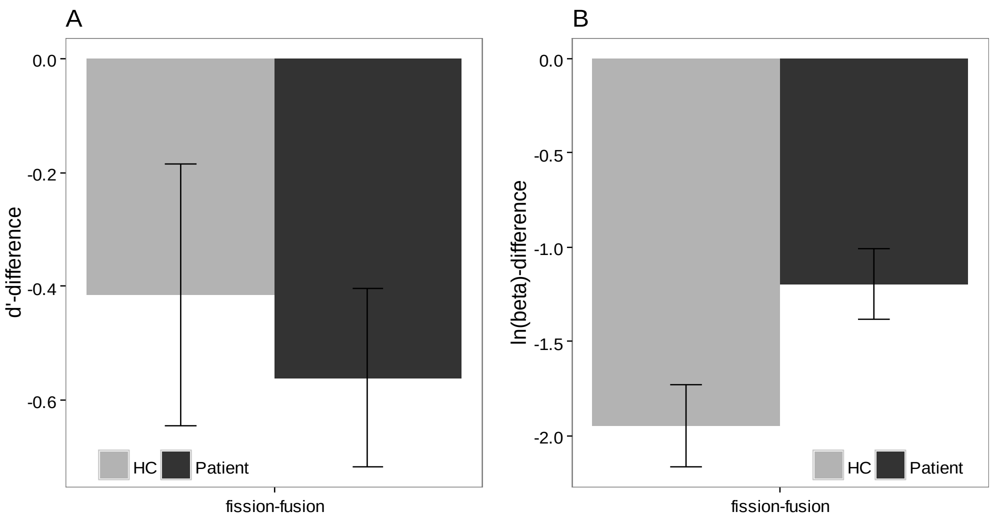


Figure 3. Mean sensitivity d' difference score (A) and mean ln(β) difference score (B) for the fission-fusion comparisons per group. Error bars represent the standard error of the mean.

Highlights

- Patients with schizophrenia are less sensitive to the sound-induced flash fusion illusion
- This is not accounted for by a more general perceptual bias or visual temporal resolution
- Reduced capacity for multisensory integration in schizophrenia is evident on an elementary perceptual level