Reduced Depth Inversion Illusions in Schizophrenia Are State-Specific and Occur for Multiple Object Types and Viewing Conditions

Brian P. Keane
University of Medicine and Dentistry of New Jersey and Rutgers University

Yushi Wang
University of Medicine and Dentistry of New Jersey

Steven M. Silverstein
University of Medicine and Dentistry of New Jersey

Thomas V. Papathomas
Rutgers University

Schizophrenia patients are less susceptible to depth inversion illusions (DIIs) in which concave faces appear as convex, but what stimulus attributes generate this effect and how does it vary with clinical state? To address these issues, we had 30 schizophrenia patients and 25 well-matched healthy controls make convexity judgments on physically concave faces and scenes. Patients were selectively sampled from three levels of care to ensure symptom heterogeneity. Half of the concave objects were painted with realistic texture to enhance the convexity illusion; the remaining objects were painted uniform beige to reduce the illusion. Subjects viewed the objects with one eye while laterally moving in front of the stimulus (to see depth via motion parallax) or with two eyes while remaining motionless (to see depth stereoscopically). For each group, DIIs were stronger with texture than without, and weaker with stereoscopic information than without, indicating that patients responded normally to stimulus alterations. More importantly, patients experienced fewer illusions than controls irrespective of the face/scene category, texture, or viewing condition (parallax/stereo). Illusions became less frequent as patients experienced more positive symptoms and required more structured treatment. Taken together, these results indicate that people with schizophrenia experience fewer DIIs with a variety of object types and viewing conditions, perhaps because of a lesser tendency to construe any type of object as convex. Moreover, positive symptoms and the need for structured treatment are associated with more accurate 3-D perception, suggesting that DI may serve as a state marker for the illness.

Keywords: schizophrenia, binocular depth inversion, positive symptoms, state marker, hollow mask illusion

Supplemental materials: http://dx.doi.org/10.1037/a0032110.supp

Three-dimensional (3-D) depth inversion illusions—more simply, depth-inversion illusions (DIIs)—occur when concave objects appear convex. A hollow face, for example, is commonly perceived as if it were protruding outward, with points furthest from the observer (i.e., on the nose) being perceived as if they were closest (Gregory, 1970). Factors contributing to DII include: 1) Object familiarity, with common objects producing more DIIs; 2) misleading surface texture, which can trick the visual system into seeing a surface recede in depth (see below); and 3) a general bias to interpret any ambiguous 3-D surface as convex, regardless of its surface texture or familiarity. What makes DIIs especially interesting—and potentially useful—is that they are perceived to a lesser extent among schizophrenia patients (Dima et al., 2009; Koethe et al., 2006; Schneider et al., 2002). Patients are hypothesized to utilize poor “top-down” processing, meaning that their conceptual expectations of how the world is supposed to look is inadequate for overriding contrary sensory (viz., stereoscopic) information (Dima et al., 2009; Koethe et al., 2006; Schneider et al., 2002).

For the top-down explanation to be taken seriously, other possibilities need to be ruled out. It is at least possible, for example, that patients are impaired not at top-down processing but at forming or deploying face representations (Sachs, Steger-Wuchse, Kryspin-Exner, Gur, & Katschnig, 2004). To our knowledge, DI reductions in schizophrenia have been shown only with faces
or only by averaging data deriving from both face and nonface stimuli where the averaging appears to reduce the effect (Koethe et al., 2009, p. 197). It thus remains unclear whether the perceptual aberration is face specific. More speculative explanations for reduced DIIs also need to be considered. For instance, patients may be worse at utilizing (misleading) depth-relevant texture information. These depth cues include foreshortening, in which a uniform texture changes predictably in appearance as it recedes in depth; or perspective, as when parallel lines appear to converge as they stretch away from the eye. Alternatively, reduced DIIs may owe to a greater reliance on stereoscopic depth. This shape-revealing property (but not others) may take precedence when conflicting with an otherwise ordinary conceptual schema of object shape. To investigate these possibilities, we had schizophrenia patients and healthy controls judge the convexity of 4 concave objects (see Figure 1). Two were 3-D scenes (“reverse perspectives”) and two were hollow faces (masks). One scene and one face were painted with texture cues to magnify the illusion; the remaining were painted matte beige to mitigate the illusion. There were two viewing conditions. In the first, subjects examined objects monocularly (with one eye) while rocking left and right. This movement allowed 3-D structure to be recovered via motion parallax, a phenomenon in which points at further distances appear to move more slowly relative to points closer. For instance, as an observer moves laterally in front of a stationary concave mask, the nose appears to move less than the surrounding parts of the face, indicating that the nose is further away and hence concave. In the second half of the experiment, subjects remained motionless and observed objects binocularly, so that stereoscopic depth could disambiguate an object’s structure (Papathomas & Bono, 2004). If reduced illusions in schizophrenia turn out to be stimulus specific—for example, if patients only have a problem in forming and accessing face representations or if patients only have abnormal stereoscopic perception—then the reduced DIIs should occur only in those cases. By contrast, if the illusion does not depend on texture, object category, or viewing condition, then the reduced illusions can be attributed to a reduced reliance on a generic convexity assumption. Another aim was to consider whether patients’ resistance to DIIs generalizes to physical 3-D objects. Prior studies induced DIIs by presenting two 2-D images to each eye, where each image was presented as either a normal stereo pair (eliciting true depth) or with the left and right eye images reversed (pseudoscopically) to invert depth perception. A major difference between this technique and normal object viewing is that in the latter, accommodation cues (how much the eyes lenses bend to see a stimulus) and vergence cues (how much the eyes turn inward to focus on an object) provide congruent and veridical information about an object’s depth structure; this is rarely the case with pseudoscopic viewing. It is therefore useful to probe whether patients witness ordinary 3-D physical objects differently, as it will bear more directly on everyday visual functioning. Finally, we sought to determine the extent to which reduced DIIs depend on the illness state. In a small longitudinal study, Schneider et al. (2002) found that DIIs became more frequent and symptoms less severe as patients transitioned from hospital admission to discharge. However, in a large cross-sectional study, Koethe et al. (2009) found that prodromal patients, neuroleptic-naïve first-episode patients, and treated acute patients differed from healthy controls, but not from one another. To shed light on whether DIIs are state- or trait-related, we selectively sampled from three treatment programs that we knew beforehand would yield highly distinct symptom profiles (see Subjects and Table S1; supplemental material available online). State-dependence predicts that patients who have fewer symptoms or who have less of a need for structured treatment will experience stronger DIIs.

**Figure 1. Stimuli.** Subjects observed concave faces and scenes that were shown with or without misleading texture. Because of the concavity, the green fixation points were further from the observer than the surrounding regions (cheeks or landscape). A beige face was convex and served as a catch. (See supplemental online materials for movies and additional pictures of stimuli.)
Subjects

The subject sample consisted of 25 healthy control subjects and 30 schizophrenia patients (Table 1). Special effort was made to match controls and patients at the group level on age, education level (father/mother/self), ethnicity, full-scale IQ, gender, and handedness (all ps > .07). Patients were recruited from the same vertically integrated treatment facility at University Behavioral HealthCare (Smith, Hull, Hedway-Harris, Ryder, & Berger, 1999) and consisted of three groups: 10 acute partial hospital (PH) program patients who were either recently discharged from an inpatient unit (<6 months) or recently admitted to the program following a symptom exacerbation in lieu of an inpatient hospitalization; 10 extended PH program patients who were first admitted into the acute PH between 6 to 24 months ago but who were still undergoing stabilization and needed the continuing daily structure of a longer-term PH program; and 10 outpatients who were last admitted into the acute PH at least 2 years prior to testing and who required only biweekly or monthly visits with treatment providers. The three subgroups did not differ significantly on age, education level (father/mother/self), ethnicity (% Caucasian), full-scale IQ, gender, handedness, or visual acuity (all ps > .06; Table S1). Written informed consent was obtained from all subjects after explanation of the nature and possible consequences of participation. The study followed the tenets of the Declaration of Helsinki and was approved by the Institutional Review Boards at Rutgers and UMDNJ. All participants were naïve to the goals of the study and received monetary compensation for their time.

For all subjects, the inclusion/exclusion criteria included: age 18–65 years; no clinically significant head injury (loss of consciousness or overnight hospitalization) or neurological disease; no diagnosis of mental retardation or pervasive developmental disorder; no substance dependence in the past 6 months; sufficient spoken English to complete testing; normal stereoscopic vision; and was approved by the Institutional Review Boards at Rutgers and UMDNJ. All participants were naïve to the goals of the study and received monetary compensation for their time.

Table 1: Demographic and Clinical Characteristics of Participants

<table>
<thead>
<tr>
<th>Variable</th>
<th>Schizophrenia (n = 30)</th>
<th>Controls (n = 25)</th>
<th>Group comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>46.6 ± 13.6</td>
<td>45.8 ± 10.2</td>
<td>p = .80</td>
</tr>
<tr>
<td>Education, father</td>
<td>12.4 ± 4.4</td>
<td>12.1 ± 3.0</td>
<td>p = .83</td>
</tr>
<tr>
<td>Education, mother</td>
<td>11.4 ± 4.2</td>
<td>10.9 ± 3.6</td>
<td>p = .61</td>
</tr>
<tr>
<td>Education, self</td>
<td>13.2 ± 2.0</td>
<td>13.3 ± 1.9</td>
<td>p = .88</td>
</tr>
<tr>
<td>Ethnicity (% Caucasian)</td>
<td>60.0 ± NA</td>
<td>24.0 ± NA</td>
<td>p = .08</td>
</tr>
<tr>
<td>FSIQ (Shipley)</td>
<td>88.5 ± 12.9</td>
<td>89.7 ± 13.0</td>
<td>p = .73</td>
</tr>
<tr>
<td>Gender (% male)</td>
<td>70.0 ± NA</td>
<td>48.0 ± NA</td>
<td>p = .10</td>
</tr>
<tr>
<td>Handedness (% right)</td>
<td>80.0 ± NA</td>
<td>92.0 ± NA</td>
<td>p = .27</td>
</tr>
<tr>
<td>PANSS, positive</td>
<td>14.7 ± 4.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PANSS, negative</td>
<td>17.3 ± 5.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PANSS, general</td>
<td>29.8 ± 8.1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
which had the same color and size as the uniformly colored concave face.

A small green square fixation mark (side = 5 mm) was placed at the center of each object. This center was locally concave for all objects, except for the catch. Objects were mounted on a uniform matte black surface at eye level for each participant, and were lit by two pairs of floodlights positioned symmetrically on either side.

Procedure

At the beginning of the experiment, subjects were shown two sample stimuli up close—one ovoid shell painted matte beige and one miniature scene painted realistically. Each sample consisted of a thin shell that was convex on one side and concave on the other, and each had a green fixation point at its center. Participants were told that they would see similar objects from afar, and that—as while fixing the green dot—they would report whether the object appeared to be “popping out” (convex) or “caving in” (concave).

On each of 10 experimental trials, a stimulus was placed 200 cm from the subject—a distance known to produce an intermediate illusion strength (Papathomas & Bon, 2004). A trial lasted for 2 minutes, during which subjects were prompted every 12 seconds to report the object’s appearance. The experimenter entered subjects’ responses into a computer and a program quantified the fraction of total time spent in the veridical percept (proportion veridical response). Between trials, subjects received a break and were shown the two sample stimuli as a reminder.

The five objects were shown for each of the two viewing conditions. In the motion parallax condition, subjects wore an eye patch over their nondominant eye (Crovitz & Zener, 1962) and laterally swayed back and forth between two vertical planks of wood—one on the left and one on the right. The planks were separated by 35.6 cm, which greatly exceeds the average separation between the eyes (interpupillary distance) (6.3 cm; Dodgson, 2004). This ensured that—during movement—subjects would acquire substantially different views (retinal images) of the objects, and thereby obtain a robust motion parallax depth signal. In the stereoscopy viewing condition, subjects binocularly examined objects from a stationary chinrest. Stereoscopic viewing always occurred in the second half of the experiment because the visual system recovers depth better with stereopsis than motion parallax (Sherman, Papathomas, Jain, & Keane, 2012), and because the aim was to minimize the chances that subjects would use their knowledge of a stimulus when making a response. The sequence of five objects was counterbalanced across participants and viewing conditions.

Analysis

There were two sets of analyses. In the first, patients were collapsed across treatment group and compared to healthy controls. Performance on the catch stimulus was evaluated with a 2 (subject group) × 2 (viewing condition) analysis of variance (ANOVA). Performance on the noncatch stimuli was assessed with a 2 (subject group) × 2 (viewing condition) × 2 (object category) × 2 (texture type) ANOVA. In the second set of analyses, the same ANOVAs were performed as above, except that controls were excluded and patient subgroup was the first factor (outpatient, extended PH, and acute PH). Clinical stability was measured as the number of days (log units) between the last admission to the acute PH program (representing beginning of the stabilization phase after the last relapse) and the date of perceptual testing. Pearson or Spearman correlations (as required by normality tests) determined the relation between DIIs, symptoms, and clinical stability. All tests were two-tailed and equal variances were assumed, unless noted.

Results

The patient and control groups identified the catch stimulus as convex on >98% of the occasions and did not differ in this regard (all ps > .4). Across all subjects, there were fewer illusions when objects were viewed stereoscopically rather than with motion parallax, $F(1, 53) = 52.0, p < .001, \eta^2_p = .495$ (Sherman et al., 2012). DIIs were also reduced when the realistic texture was replaced by a blank surface ($F(1, 53) = 56.2, p < .001, \eta^2_p = .515$), which was expected since the texture created a more familiar appearance and provided misleading depth information (Sherman et al., 2012). The effect of texture modulation was amplified with binocular viewing, $F(1, 53) = 11.8, p = .001, \eta^2_p = .182$. Most importantly, patients experienced DIIs less than healthy controls ($F(1, 53) = 5.1, p = .028, \eta^2_p = .088$) and this group difference did not depend on viewing condition, texture, or object category (all ps > .15). Notably, the group difference arose when the analyses were restricted only to monocular viewings or only to scenes (ps < .03). For all subjects, performance did not correlate with age, education (mother, father, or self), full-scale IQ, or visual acuity (all ps > .06).

Next, we compared DIIs among patients only (outpatient, extended, acute). The subgroups performed similarly on the catch stimulus (ps > .3). Patients exhibited less illusion with stereopsis than without ($F(1, 27) = 22.8, p = .00006, \eta^2_p = .458$), and more illusion with texture than without ($F(1, 27) = 21.4, p = .00008, \eta^2_p = .442$). Importantly, the illusion strength varied as a function of patient subgroup ($F(2, 27) = 3.8, p = .036, \eta^2_p = .218$), which itself did not depend on the viewing condition, object category, or texture (all ps > .17). Follow-up t-tests revealed that, overall, the acute patients differed marginally from the extended patients (t(14.9) = 1.88, unequal variances; p = .08) and significantly from the outpatients, t(18) = 3.0, p = .009, but that the latter two groups did not differ from one another (p > .4). Patient performance did not correlate with age, education level (mother/father/self), full-scale IQ, or visual acuity (all ps > .10).

To further consider whether group differences depended on stimulus parameters, we compared only the two most extreme groups, control and acute. They differed enormously in overall DII (p = .0003, \eta^2_p = .32), but not in any other way (ps > .11, \eta^2_p < .08), suggesting that even when group differences are stark, they do not strongly depend on stimulus properties.

PANS symptoms (positive, negative, and general) increased linearly from outpatient to acute patients (all ps < .004, one-tailed; see also ANOVA results, Table S1). Perhaps because of this heterogeneity, we were able to identify a positive correlation between proportion veridical response and positive symptoms (r = .42, p = .021). DIIs did not correlate with negative symptoms (p = .22), general symptoms (p = .11), cognitive disorganization (p = .06) or Depression (the sum of G1, G2,
G3, and G6; $p > .4$). Exploratory analyses revealed a strong negative correlation between inappropriate affect and DII ($r = .592$, $p = .0006$). A significant correlation was observed between DII and log days since last acute PH admission, a measure of clinical stability ($r = .50$, $p = .006$; Figure 2C).

**Discussion**

When judging the appearance of physically concave objects, schizophrenia patients experienced fewer DII than healthy controls who were matched at the group level on eight different variables. Illusion resistance was more pronounced for patients who had a more recent acute PH admission date, more of a need for structured treatment, more positive symptoms, or less appropriate affect. Differences between patients and controls or between subgroups of patients did not significantly depend on object category (faces/scenes), viewing condition (parallax/stereopsis) or texture type (realistic/blank). These results could not be explained in terms of a stimulus-independent concavity bias among patients, since each patient subgroup almost always identified the catch stimulus as convex. Together, these data suggest that i) schizophrenia patients more accurately perceive the 3-D structure of concave objects; ii) the likelihood of veridical perception increases monotonically with state severity; and iii) between-groups differences in convexity perception occur for a variety of object types and viewing conditions.

These data provide important clues as to why weaker DII are experienced in schizophrenia. As noted in the Introduction, prior studies examined DII with faces (at least in part), but because faces may be processed abnormally in schizophrenia (Sachs et al., 2004), there remained the possibility that group differences could be explained in terms of this dysfunction rather than poor top-down modulation. Our data disconfirm this view. None of the patient (or control) effects depended on whether objects were scenes or faces, and group differences arose even when the scene stimuli were considered alone. Second, since the late 1980s, it has been argued that schizophrenia involves reduced “binocular depth inversion” (Emrich, 1989), but our data show that binocularity is not necessary for
the group difference. When veridical depth information derived purely from monocular motion parallax, patients still experienced fewer DIIs. Finally, patients and controls responded similarly to alterations in texture and depth, which indicates that the processing of these features is not obviously disturbed in schizophrenia. The constancy of the group (or subgroup) differences across stimulus conditions is noteworthy since it shows that the overall level of illusion can shift dramatically without large changes in the processing of any other stimulus attribute. This suggests that schizophrenia may entail a weakened bias to construe ambiguous objects as convex, regardless of their stimulus properties, although larger subject samples will be needed to further confirm this conclusion.

The relation between DII, treatment program (subgroup), and duration since last AP hospital admission indicates that abnormal seeing predicts level of functioning. This is so because—in the current study—program assignments were made within a single, vertically integrated system that: makes level of care decisions on the basis of standard criteria (as set by external review agencies), facilitates rapid transfers between programs, and maintains continuity of treatment (Smith et al., 1999). Thus, the program assignment—and the duration since last acute PH admission—reflect the need for daily structure and intensive treatment in a highly standardized way. A goal in future studies will be to probe whether DIIs can predict future need for treatment and, in particular, future inpatient and emergency room visits.

Our illness-state effects are consistent with the longitudinal inpatient study of Schneider et al. (2002), but inconsistent with the larger, cross-sectional study of Koethe et al. (2009), which documented equivalent performance between prodromal and treated schizophrenia patients. Why the discrepancy? We hypothesize that the subject groups of Koethe et al. were, on average, asymptomatic schizophrenia patients. Why the discrepancy? We hypothesize that schizophrenia may entail a weakened bias to construe ambiguous objects as convex, regardless of their stimulus properties, although larger subject samples will be needed to further confirm this conclusion.

The relation between DII, treatment program (subgroup), and duration since last AP hospital admission indicates that abnormal seeing predicts level of functioning. This is so because—in the current study—program assignments were made within a single, vertically integrated system that: makes level of care decisions on the basis of standard criteria (as set by external review agencies), facilitates rapid transfers between programs, and maintains continuity of treatment (Smith et al., 1999). Thus, the program assignment—and the duration since last acute PH admission—reflect the need for daily structure and intensive treatment in a highly standardized way. A goal in future studies will be to probe whether DIIs can predict future need for treatment and, in particular, future inpatient and emergency room visits.

Our illness-state effects are consistent with the longitudinal inpatient study of Schneider et al. (2002), but inconsistent with the larger, cross-sectional study of Koethe et al. (2009), which documented equivalent performance between prodromal and treated schizophrenia patients. Why the discrepancy? We hypothesize that the subject groups of Koethe et al. were, on average, asymptomatic schizophrenia patients. Why the discrepancy? We hypothesize that schizophrenia may entail a weakened bias to construe ambiguous objects as convex, regardless of their stimulus properties, although larger subject samples will be needed to further confirm this conclusion.

State-related reductions of DII in schizophrenia may begin to be explained by disturbances in prediction-error monitoring, the statistical process by which the brain updates expectations about future states of the world on the basis of how past predictions matched with past experience. This Bayesian belief-updating process provides one way of characterizing the at times nebulous notion of “top-down influence” (Dima et al., 2010). The idea in its simplest form is that past experiences with mostly convex objects (faces) leads us to construe newly encountered ambiguous objects as also being convex. This experientially established assumption is typically strong enough to effectively veto conflicting (e.g., stereoscopic) information. People with schizophrenia and particularly those individuals with positive symptoms are especially poor at detecting or employing prediction errors (Fletcher & Frith, 2009), and this may provide some reason why they are also less susceptible to DII.

The presented data also offer methodological advancement. The DII task typically involves judging local and global convexity with a 5-item subjective rating scale, ranging from clearly convex to clearly concave. A problem that Schneider et al. (2002) pointed out is that different subject groups could use the scale in different ways (e.g., a patient may need less evidence than a control to categorize an object as clearly concave rather than possibly concave). Our version of the task avoids confounds with response bias by having subjects make simple, qualitative, binary judgments. Additionally, all prior studies utilized pseudoscopy, which requires a computer, spatially offset 2-D images, and a device for stereoscopic presentation (e.g., shutter glasses or Wheatstone stereoscope). The present study, on the other hand, has shown that appropriately placed, uniform concave objects are perfectly acceptable alternatives, and that—in principle—no electronic equipment is even necessary to study the effect.

Finally, the current results show that abnormal convexity perception potentially occurs outside of the laboratory setting and applies to objects that could realistically be encountered in everyday life (i.e., any concave surface). Future research will need to consider the practical ramifications of this perceptual difference, especially for more symptomatic or more recently hospitalized patients.

Limitations
A limitation of our approach was that—similar to prior studies—the interviewer knew a subject’s patient status when recording a subject’s response, introducing the potential for experimenter bias. This confound is not especially troubling since many of the significant findings were unanticipated, and since the interviewer’s role was merely to periodically prompt a binary response. Another limitation was the small size of the patient subgroups. Null results involving this factor indicate only the absence of a large effect; smaller effects cannot be discounted until larger samples are examined. However, it is worth noting that the overall patient sample size allowed for easy detection of certain within-group effects (e.g., texture). Third, a standard interview for level of functioning was not employed. A caveat is that these measures do not always correlate highly with each other, and may not accurately reveal the need for daily structure and intensive treatment as do the “real-world” metric of program assignment (Silverstein, All, & Jaeger, 2011). Notwithstanding the foregoing limitations, the data show that schizophrenia patients have less DIIs with physical objects, perhaps because of a weakened visual convexity assumption. This impairment is strongly linked to the presence of positive symptoms and need for structure, and may therefore represent a biomarker of psychosis and recovery.

References


Received August 28, 2012
Revision received January 22, 2013
Accepted January 28, 2013