Psychological Medicine (2012), **42**, 2301–2311. © Cambridge University Press 2012
doi:10.1017/S0033291712000384

ORIGINAL ARTICLE

The influence of prior expectations on facial expression discrimination in schizophrenia

G. Barbalat[1]*, M. Rouault[2], N. Bazargani[1], S. Shergill[3] and S.-J. Blakemore[1]

[1] *UCL Institute of Cognitive Neuroscience, London, UK*
[2] *Ecole Normale Supérieure de Lyon, France*
[3] *Department of Psychiatry, Downing Site, University of Cambridge, Cambridge, UK*

Background. Belief inflexibility is a thinking style observed in patients with schizophrenia, in which patients tend to refute evidence that runs counter to their prior beliefs. This bias has been related to a dominance of prior expectations (prior beliefs) over incoming sensory evidence. In this study we investigated the reliance on prior expectations for the processing of emotional faces in schizophrenia.

Method. Eighteen patients with schizophrenia and 18 healthy controls were presented with sequences of emotional (happy, fearful, angry or neutral) faces. Perceptual decisions were biased towards a particular expression by a specific instruction at the start of each sequence, referred to as the context in which stimuli occurred. Participants were required to judge the emotion on each face and the effect of the context on emotion discrimination was investigated.

Results. For threatening emotions (anger and fear), there was a performance cost for facial expressions that were incongruent with, and perceptually close to, the expression named in the instruction. For example, for angry faces, participants in both groups made more errors and reaction times (RTs) were longer when they were asked to look out for fearful faces compared with the other contexts. This bias against sensory evidence that runs counter to prior information was stronger in the patients, evidenced by a group by context interaction in accuracy and RTs for anger and fear respectively.

Conclusions. Overall, the present data suggest an overdependence on prior expectations for threatening stimuli, reflecting belief inflexibility, in schizophrenia.

Received 13 September 2011; Revised 8 February 2012; Accepted 9 February 2012; First published online 12 March 2012

Key words: Belief inflexibility, bias against disconfirmatory evidence, emotion perception, prior expectations, schizophrenia.

Introduction

Several theories have suggested that cognitive biases contribute to the symptoms of schizophrenia. One of the most influential is related to problems with probabilistic reasoning, that is an abnormal integration of new evidence into prior beliefs (Blackwood *et al*. 2001; Moritz & Woodward, 2005; Freeman, 2007). An example of such a bias is the 'jumping to conclusions' bias, in which patients with schizophrenia tend to make decisions based on less evidence than healthy controls (Garety *et al*. 1991, 2005; Moritz & Woodward, 2005; Averbeck *et al*. 2011). Another example is belief inflexibility (Woodward *et al*. 2008), a thinking style in which patients show an unwillingness to modify their belief even when confronted with disconfirmatory evidence. In this case, it has been proposed that the maintenance of false beliefs, a common symptom in schizophrenia, might be related to patients giving too much weight to prior expectations (prior beliefs) as compared to incoming sensory evidence. This imbalance between prior beliefs and new sensory evidence would then give rise to the discounting of disconfirmatory evidence that runs counter to prior beliefs (Moritz & Woodward, 2006; Woodward *et al*. 2006, 2007).

The inflexible beliefs associated with schizophrenia often have an emotional component. For example, people with paranoid schizophrenia typically believe that they are under threat by others. Prominent theories of emotion processing have proposed that top-down prior expectations (prior beliefs) strongly influence emotion perception (Pessoa, 2008). However, the majority of studies that have investigated belief inflexibility have not attempted to integrate this bias with emotion perception. The goal of the current study was to investigate the influence of prior expectations on facial expression discrimination in

* Address for correspondence: Dr G. Barbalat, UCL Institute of Cognitive Neuroscience, 17 Queen Square, London WC1N 3AR, UK.
(Email: guillaumebarbalat@gmail.com)

Table 1. *Sociodemographic data and clinical ratings of the PANSS for all participants*

	Schizophrenic patients ($n=18$)	Healthy participants ($n=18$)
Age (years)	38.4 ± 7.7	35.5 ± 7.8
Number of males	10	9
Verbal IQ	90.4 ± 16.0	107.0 ± 17.9
Clinical ratings: PANSS	48.8 ± 10.4	N.A.
Positive symptoms scale (P1–P7)	11.5 ± 4.5	N.A.
Suspiciousness/persecution subscale (P6)	2.15 ± 1.3	N.A.
Negative symptoms scale (N1–N7)	14.1 ± 4.2	N.A.
General psychopathology scale (G1–G16)	22.7 ± 4.8	N.A.

PANSS, Positive and Negative Symptoms Scale; N.A., not assessed.
Values given as mean ± standard deviation or *n*.

schizophrenia. To this aim, we used a task in which perceptual decisions were biased by prior expectations. In healthy participants, it has previously been shown that prior expectations, in the form of a simple instruction asking for a presence/absence (or 'yes/no') judgment, improve the ability to discriminate incoming stimuli that specifically match expectations (Summerfield & Koechlin, 2008). In this study we adapted this paradigm to investigate whether prior expectations similarly bias facial emotion discrimination in healthy individuals, and whether such a bias is heightened in patients with schizophrenia.

We presented short sequences of photographs of faces displaying one of four emotional expressions (happy, fearful, angry or neutral) to 18 patients with schizophrenia and 18 healthy controls. At the beginning of each sequence, participants were instructed to look out for faces with a particular 'target' expression. Participants were asked to respond to each face stimulus by pressing one of two-alternative forced choice response buttons: one corresponding to the target emotion, the other to all non-target emotions. Our first hypothesis was that this manipulation of the instruction cue would provide a prior bias towards the corresponding emotion, which should result in lower performance in emotion discrimination in face of disconfirmatory evidence (Summerfield & Koechlin, 2008). Second, we predicted that patients with schizophrenia would demonstrate a stronger prior bias than do healthy controls (Fletcher & Frith, 2009), in particular for threat-related emotions (anger and fear).

Method

Participants

Eighteen healthy, right-handed volunteers [nine males; mean (S.D.) age = 35.5 (7.8) years] and 18 clinically stable, age- and gender-matched, right-handed patients with DSM-IV schizophrenia [10 males; mean (S.D.) = 38.4 (7.7) years] took part in the current study (Table 1). All patients were on stable doses of second-generation antipsychotic medication. Symptom severity was assessed with the Positive and Negative Syndrome Scale (PANSS; Kay *et al*. 1987; see Table 1). Exclusion criteria for the healthy participants were the presence of any neurological or psychiatric disorder for the participant or their first-degree relatives. After the procedure was fully explained, written consent was obtained from all participants according to the declaration of Helsinki. Participants were paid £10 per hour. The study was approved by the Institute of Psychiatry Research Ethics Committee.

A measure of verbal IQ (vIQ), as an estimate of pre-morbid intelligence, was obtained using the Vocabulary subset of the Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999). There was a significant difference in vIQ between patients and controls [$t(34) = 2.9$, $p < 0.01$; see Table 1]. However, in a meta-analysis, Woodberry *et al*. (2008) concluded that: 'Years before the onset of psychosis symptoms, individuals with schizophrenia, as a group, demonstrate mean IQ scores approximately one-half of a standard deviation below that of healthy comparison subjects. … A significant decline in the IQ of individuals with schizophrenia, relative to comparison subjects, was associated with the onset of frank psychosis'. Therefore, a representative sample of schizophrenia patients would necessarily be associated with a lower IQ as compared to healthy controls. In addition, we analysed the association between vIQ and task performance in each group to ensure that vIQ was not associated with our dependent measure of bias. Finally, we analysed subgroups ($n = 12$ in each group) matched on vIQ to ensure that their bias results did not differ from the results of the whole group.

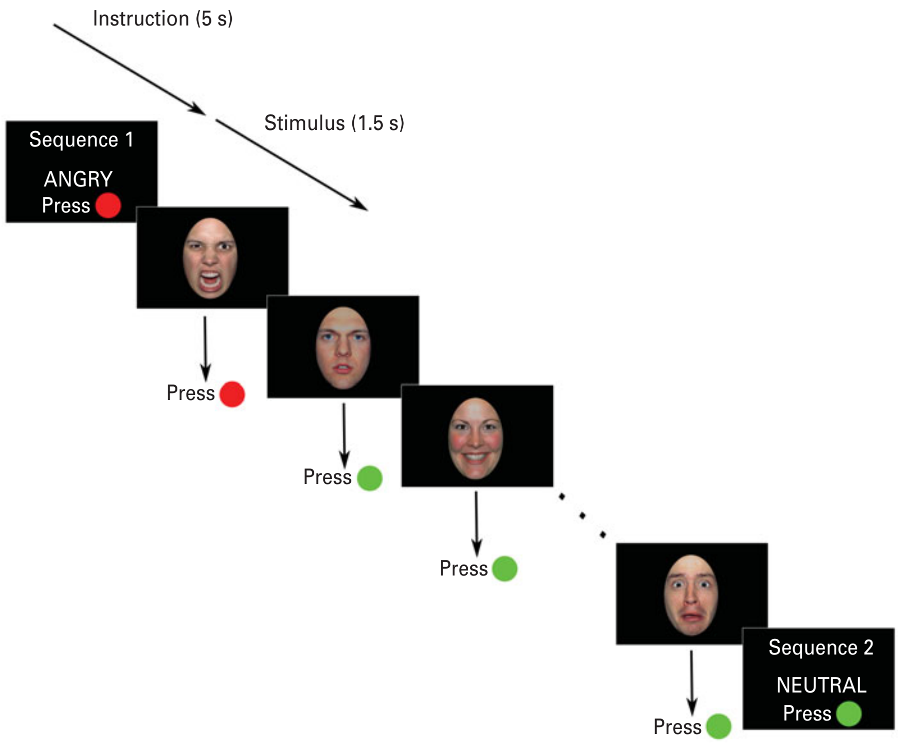

Fig. 1. Experimental paradigm. According to the instructions, participants had to press the red (target response) button for every angry expression displayed (the target expression) and to press the green button for every other expression displayed (non-target expressions). Details are provided in the main text.

Emotion task

The face emotion discrimination task consisted of 76 sequences, divided into four experimental runs of ~6 min each, in which faces showing different emotions (happy, fearful, angry or neutral) were presented sequentially. The design was $2 \times 4 \times 4$ factorial, with group as the between-subjects factor and two within-subjects factors: 'emotion' shown on each face (four levels) and 'context' set by the specific instruction at the beginning of each sequence (four levels). Each sequence began with a 5-s instruction cue, followed by eight successive emotional faces and ending with a rest period of 2 s. Each face stimulus was presented for 1.5 s on a uniform black background, and was followed by a fixation cross in the centre of the screen presented for 0.5 s. The instruction (which corresponded to the context factor) informed participants about the nature of the target facial expression (happy, fearful, angry or neutral) and the corresponding target response button (Fig. 1). Specifically, participants were told to press the target button on the keyboard if a face stimulus was concordant with the instruction, and the non-target button if it was not. For example, in happy context sequences, participants were instructed to press the target button if the stimulus was a happy face and the non-target button if the face displayed a fearful, angry or neutral expression. The order of happy, fearful, angry or neutral context sequences, along with the response keys (right index *versus* middle finger) used for target/non-target emotions, was counterbalanced within and between participants.

Within each sequence of eight faces, there were four target and four non-target face stimuli. This enabled an equal number of right index and middle finger button presses within each sequence. For each target emotion, there were 76 non-target emotion stimuli (19 sequences for each target emotion $\times$ 4 non-target stimuli) across the whole experiment, which were approximately equally distributed between the three different possible non-target expressions (25, 25, 26, counterbalanced between participants).

Stimuli (coloured photographs of faces) were selected from a set validated in healthy people and patients with schizophrenia, and consisted of eight identities (four males; four females) of Caucasian ethnicity with the highest degree of emotion identification accuracy (Tottenham *et al*. 2009). All faces had a direct gaze and open mouth. We used Adobe Photoshop CS3 to remove extra-facial elements and to adjust the pixel intensity and size of the presented picture (13.97 mm in width and 20.66 mm in height). The order of the stimuli within a sequence was pseudo-randomized so that, within each sequence, no more than two consecutive pictures displayed the same expression or the same gender. Within each sequence, the eight pictures represented the eight different identities. The task was presented on a laptop using the Cogent Software Package (www.cogentsoftware.net/). Accuracy and reaction time (RT) were recorded for each trial. Participants were instructed to respond as quickly and as accurately as possible. No feedback was given. Before the experiment, participants performed a short practice session of 10 sequences.

Behavioural analysis on accuracy and RTs

Accuracy (percentage correct) and median RT from correct responses were calculated for each participant and each of the 16 (4 contexts $\times$ 4 emotions) conditions (a condition being a particular emotion in a particular context). Satisfactory sequences were defined as those completed with an accuracy $\geqslant 50\%$ (i.e. at least four correct responses in each sequence). Using this criterion, on average, 0.47 (S.D. $= 0.62$) sequences (out of 76) were excluded in the control participants, compared to an average of 4.80 (S.D. $= 4.03$) in the patients ($p < 0.001$).

A $4 \times 4 \times 2$ repeated-measures ANOVA with emotion (happy, fearful, angry or neutral) and context (happy, fearful, angry or neutral) as within-subjects factors and group (controls *versus* patients) as the between-subjects factor was performed on accuracy and RTs. To investigate in further detail the differential effect of context for each expression, a *post-hoc* 2×4 (group $\times$ context) ANOVA was performed on the same measures for each expression separately. Where appropriate, the Greenhouse–Geisser epsilon correction was performed to control for violation of the sphericity assumption. Significant group $\times$ context interactions were further interrogated using Fisher least significant difference (LSD) *post-hoc* tests.

Finally, in a preliminary analysis to investigate the relationship between the prior bias for threatening faces (i.e. the difference in performance between biased and non-biased conditions) and persecutory symptoms in schizophrenia, we computed a 'prior bias' score for threatening emotions. This was defined as the difference in performance between conditions in which the responses to the threatening emotion (anger or fear) were biased by the context (i.e. anger in an angry context; fear in a fearful context) and conditions in which the responses to the threatening emotion were not biased by the context (i.e. fear in an angry context; anger in a fearful context). We conducted non-parametric analyses to test whether patients with persecutory delusions, defined as having a score of $\geqslant 3$ (i.e. mild or greater) on the Suspiciousness/Persecution subscale (P6) of the PANSS, presented a higher prior bias for threatening emotions compared with patients without such symptoms (< 3 on P6).

Results

Analyses on accuracy and RT

The main 4 (emotion) $\times$ 4 (context) $\times$ 2 (group) ANOVA on accuracy and RT revealed that all main effects were significant. As expected, patients made more errors and were slower than controls overall (main effect of group: both F's > 16.1, $p < 0.001$). Accuracy and RT also significantly varied with emotion (main effect of emotion: both F's > 22.2, $p < 0.001$) and context (main effect of context: both F's > 20.8, $p < 0.001$). The context effect was different for each emotion, as revealed by the significant emotion $\times$ context interaction (both F's > 24.6, $p < 0.001$). This effect was further qualified by a three-way group $\times$ emotion $\times$ context interaction (both F's > 2.8, $p < 0.05$), indicating different forms of the emotion $\times$ context interaction in the two groups.

To investigate the differential effects of the context on emotion in patients as compared to controls, we then performed a 2 (group) $\times$ 4 (context) ANOVA on each expression separately. We found a significant main effect of context for all expressions, and a significant group $\times$ context interaction for angry, fearful and neutral faces.

For the angry expression, the context effect was observed for both accuracy and RT (both F's > 41.3, $p < 0.001$). Specifically, both groups' accuracy was significantly lower (Fig. 2*a*) and RTs were higher (Fig. 2*b*) when participants were asked to discriminate angry faces in a fearful context as compared to angry faces in an angry context (*post-hoc* LSD Fisher tests, all $p < 0.001$). In both groups, the ability to discriminate an angry face in an angry context was not significantly higher than the ability to discriminate an angry face in a happy or neutral context (Fig. 2*a*, *b*). The group $\times$ context interaction was driven by accuracy ($F = 6.6$, $p < 0.005$; RT: $F = 0.5$, $p =$ N.S.; Fig. 2*a*, *b*) and indicated that patients' accuracy was lower than that of controls when they had to identify an angry face in a fearful context as compared to an angry face in an angry context [2 (patients *versus* controls) $\times$ 2 (fearful *versus* angry contexts) ANOVA; interaction, $F = 10.2$, $p < 0.005$; see Fig. 2*a*).

For the fearful expression, the context effect was also observed for both accuracy and RT (both F's > 27.6, $p < 0.001$). Both groups' accuracy was significantly lower (Fig. 2*c*) and RTs were higher

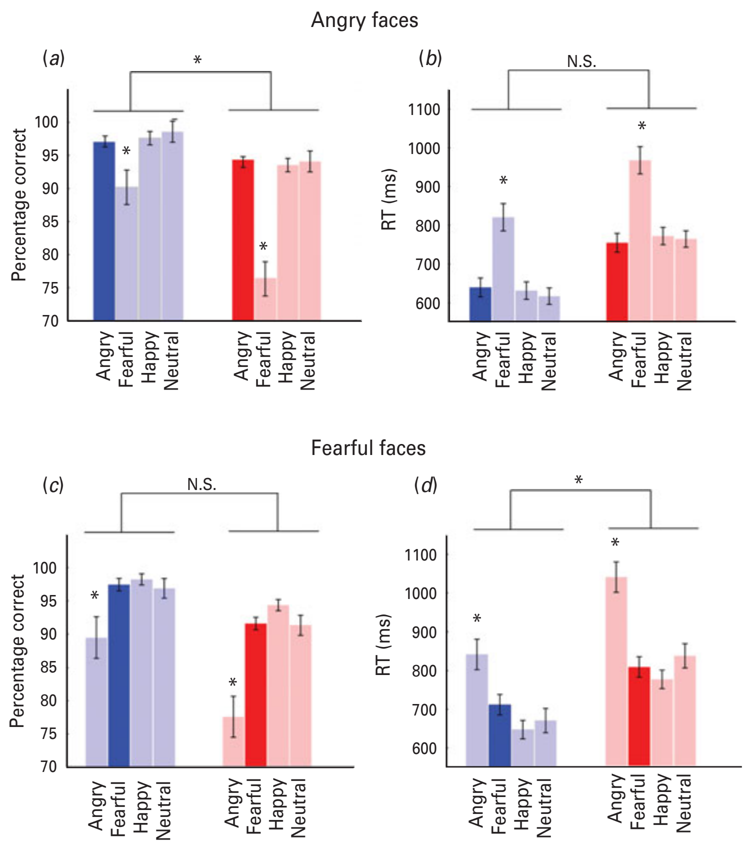

Fig. 2. Results from analyses on accuracy and reaction times (RTs) for the threatening emotions. (*a*) Accuracy (percentage correct) and (*b*) RTs (ms) for the conditions showing angry faces and (*c*) accuracy (percentage correct) and (*d*) RTs (ms) for the conditions showing fearful faces, for each of the four contexts (angry, fearful, happy, neutral). Data for congruent and incongruent contexts are represented by dark and light colours respectively. Data from the patients are in red and those from the controls are in blue. There was a group × condition interaction for accuracy for angry faces and RT for fearful faces. This was driven by a greater performance cost for facial expressions that were incongruent with, and perceptually close to, the expression named in the instruction for patients relative to controls. Error bars represent the standard errors.

(Fig. 2*d*) when participants were asked to discriminate fearful faces in an angry context as compared to fearful faces in a fearful context (*post-hoc* LSD Fisher tests, all $p<0.001$). In both groups, the ability to discriminate a fearful face in a fearful context was not significantly higher than the ability to discriminate a fearful face in a happy or neutral context (Fig. 2*c*, *d*). The group × context interaction was driven by RT ($F=3.0$, $p=0.05$; accuracy: $F=2.6$, $p=$ N.S.; see Fig. 2*c*, *d*) and indicated that patients were slower than controls to identify a fearful face in an angry context than a fearful face in a fearful context (2 × 2 ANOVA, interaction effect, $F=6.9$, $p<0.01$; see Fig. 2*d*).

For the happy expression, the context effect was only observed for RT ($F=24$, $p<0.001$; for accuracy, $F=1.3$, $p=$ N.S.; Fig. 3*a*, *b*). Both groups' RTs were faster when participants were asked to discriminate happy faces in a happy context as compared to happy faces in all other contexts (*post-hoc* LSD Fisher tests, all $p<0.05$; Fig. 3*b*). We observed no significant group × context interaction (both F's <2.4, $p=$ N.S.; Fig. 3*a*, *b*).

For the neutral expression, the context effect was observed for both accuracy and RT (both F's >5.4, $p<0.005$). However, we did not find a clear pattern of results regarding the effects of prior information on decision making. For accuracy, we did not find any significant difference from one context to another in controls (all $p=$ N.S., Fig. 3*c*), and the only significant difference involving the neutral context in patients was a higher accuracy when patients were asked to discriminate neutral faces in a neutral context as compared to neutral faces in a fearful context ($p<0.05$, Fig. 3*c*). For RTs, controls were faster when participants were asked to discriminate neutral faces in a fearful context as compared to neutral faces in all other contexts (*post-hoc* LSD Fisher tests, all $p<0.005$; Fig. 3*d*). Patients were faster when participants were asked to discriminate neutral faces in a neutral context as compared to neutral faces in all other contexts

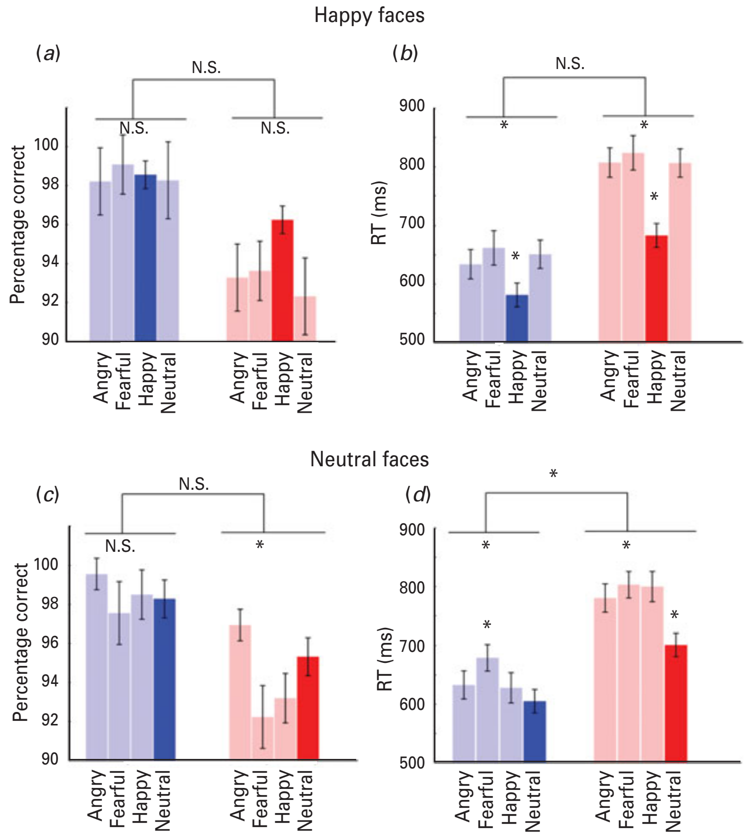

Fig. 3. Results from analyses on accuracy and reaction times (RTs) for the happy and neutral expressions. (*a*) Accuracy (percentage correct) and (*b*) RTs (ms) for the conditions showing happy faces and (*c*) accuracy (percentage correct) and (*d*) RTs (ms) for the conditions showing neutral faces, for each of the four contexts (angry, fearful, happy, neutral). Data for congruent and incongruent contexts are represented by dark and light colours respectively. Data from the patients are in red and those from the controls are in blue. For the happy expression, we found a main effect of context only for RTs, but no group × context interaction. For the neutral expression, the context effect, observed for both accuracy and RT, did not yield any clear pattern regarding the dependence on prior information. Error bars represent the standard errors.

(*post-hoc* LSD Fisher tests, all $p<0.001$; Fig. 3*d*). Finally, we found a group × context interaction driven by RT ($F=4.2$, $p<0.01$; for accuracy: $F=1.4$, $p=$ N.S.; Fig. 3*c*, *d*), confirming that patients' RTs were different from those of the controls according to the context.

Table 2 presents a summary of all the statistics, including the *F* and *p* values, and the degrees of freedom (df), taking into account the correction for sphericity.

To summarize the results, we found a differential effect of prior expectations on emotion discrimination for the different facial expressions. For angry and fearful faces, there was a performance cost for facial expressions that were incongruent with, and perceptually close to, the expression named in the instruction (i.e. fearful and angry contexts respectively). By contrast, for happy faces, emotion discrimination was improved when the emotion displayed in the face stimulus was congruent with the instruction as compared to when it was incongruent. We did not find a clear dependence on prior information for the neutral expression. Finally, the between-group results indicated that the conditional expectation effect was stronger in the patients for the threatening emotions.

It is noteworthy that we found the same pattern of results (a group × context interaction in accuracy and RT for anger and fear respectively) when rerunning the analyses after having included blocks in which performance was unsatisfactory (both *F*'s >2.9, $p<0.05$). Moreover, we observed no correlation between the biased *versus* unbiased contexts for anger and fear and: (1) the overall performance for angry faces and fearful faces respectively ($r<0.1$, $p=$ N.S. in each group); or (2) general cognitive abilities (as measured by vIQ) ($r<0.2$, $p=$ N.S. in each group). To ensure that the results did not change when vIQ was matched, we reran the analysis on a subset of participants in both groups matched for vIQ [$n=12$ in each group; for controls: mean (S.D.) vIQ score = 97.5 (13.7); for patients with schizophrenia: mean (S.D.) vocabulary *T* score = 98.6 (11.3); $t(22)=0.2$, $p=$ N.S.]. There

Table 2. *Results from all ANOVAs on accuracy and reaction times (RTs)*

Effect	*F* value	Sphericity (*p* value)	df[a]	*p* value
From the 2 (group) × 4 (emotions) × 4 (contexts) ANOVA on accuracy				
Group	16.1	N.A.	1, 34	< 0.001
Emotion	22.2	< 0.001	1.7, 56.5	< 0.001
Emotion × group	1.8	< 0.001	1.7, 56.5	0.2
Context	20.8	0.8	3, 102	< 0.001
Context × group	3.7	0.8	3, 102	0.02
Emotion × context	24.6	< 0.001	3.9, 132.3	< 0.001
Emotion × context × group	3.2	< 0.001	3.9, 132.3	0.02
From the 2 (group) × 4 (emotions) × 4 (contexts) ANOVA on RT				
Group	21.5	N.A.	1, 34	< 0.001
Emotion	36.1	< 0.001	2.1, 69.9	< 0.001
Emotion × group	0.2	< 0.001	2.1, 69.9	0.8
Context	67.1	< 0.001	2.2, 74.3	< 0.001
Context × group	1.1	< 0.001	2.2, 74.3	0.4
Emotion × context	49	< 0.001	4.9, 165.1	< 0.001
Emotion × context × group	2.9	< 0.001	4.9, 165.1	0.02
From the 2 (group) × 4 (contexts) ANOVA on accuracy (one for each emotion)				
Angry faces				
Group	14.9	N.A.	1, 34	< 0.001
Context	41.3	< 0.001	1.8, 60.9	< 0.001
Group × context	6.6	< 0.001	1.8, 60.9	0.003
Fearful faces				
Group	15.7	N.A.	1, 34	< 0.001
Context	27.6	< 0.001	1.5, 49.5	< 0.001
Group × context	2.6	< 0.001	1.5, 49.5	0.1
Happy faces				
Group	7.5	N.A.	1, 34	0.01
Context	1.3	0.09	3, 102	0.3
Group × context	1.3	0.09	3, 102	0.4
Neutral faces				
Group	9.6	N.A.	1, 34	0.004
Context	5.4	0.07	3, 102	0.002
Group × context	1.4	0.07	3, 102	0.2
From the 2 (group) × 4 (contexts) ANOVA on RT (one for each emotion)				
Angry faces				
Group	18.3	N.A.	1, 34	< 0.001
Context	85	< 0.001	1.8, 62.1	< 0.001
Group × context	0.5	< 0.001	1.8, 62.1	0.6
Fearful faces				
Group	14.8	N.A.	1, 34	< 0.001
Context	63.8	0.003	2.2, 73.7	< 0.001
Group × context	3	0.003	2.2, 73.7	0.05
Happy faces				
Group	23.6	N.A.	1, 34	< 0.001
Context	24	0.01	2.5, 83.8	< 0.001
Group × context	2.4	0.01	2.5, 83.8	0.09
Neutral faces				
Group	20.4	N.A.	1, 34	< 0.001
Context	21.8	0.8	3, 102	< 0.001
Group × context	4.2	0.8	3, 102	0.008

RT, Reaction time; df, degrees of freedom; N.A., not assessed.

[a] After having taken into account the Greenhouse–Geisser epsilon correction for df to control for violation of the sphericity assumption (where appropriate).

https://doi.org/10.1017/S0033291712000384 Published online by Cambridge University Press

remained a significant group × context interaction for anger (accuracy: $F_{3,66}=5.1$, $p<0.005$) and fear (RTs: $F_{3,66}=3.0$, $p<0.05$). Therefore, this higher conditional expectation effect for the fearful and angry expressions that we observed in patients is probably not due to a more general cognitive deficit in processing emotional faces (Miller *et al*. 1995).

Relationship with persecutory delusions

In a preliminary analysis conducted to relate the prior bias for threatening faces to persecutory symptoms in schizophrenia, we computed a 'prior bias' score as the difference between accuracy in conditions in which the threatening emotion was biased by the context (i.e. anger in an angry context; fear in a fearful context) from conditions in which the threatening emotion was unbiased (i.e. fear in an angry context; anger in a fearful context). We found that patients with persecutory delusions ($n=9$) presented a higher dependence on prior information for anger than patients without persecutory symptoms ($n=9$) (non-parametric Mann–Witney test, $Z=2.4$, $p<0.05$; we observed a marginal significant result for fear, $Z=1.8$, $p=0.07$). This was driven by higher performance accuracy in patients with persecutory delusions compared with patients without persecutory delusions for the unbiased context only (i.e. fear in an angry context; anger in a fearful context; both Z's >2.2, $p<0.05$; for the biased context, both Z's <1.7, $p=$ N.S.). There was no difference between the two symptom groups in terms of their dependence on prior information for happy or neutral expressions, nor in terms of their performance for simple expression discrimination ($Z<1.0$, $p=$ N.S.). There was no significant difference between patients with and those without persecutory delusions with respect to overall symptoms severity (according to the PANSS) and vIQ (both Z's <1.2, $p=$ N.S.). We note that the numbers of patients in both groups is relatively low and therefore these results are preliminary and need replicating in a larger sample.

Discussion

In this study, we used an emotion discrimination task, where decisions were biased by an instruction, to investigate the influence of prior expectations on facial expression discrimination in schizophrenia. Healthy control participants tended to overvalue prior expectations (the instruction) when faced with an emotional stimulus (facial expression) that was incongruent with the expectation. For threatening emotions (fear and anger), this dependence on prior expectations was heightened in patients with schizophrenia, and a preliminary analysis revealed that this was even more pronounced in patients with persecutory delusions.

In the current study, we biased participants to form prior expectations by simply instructing them to discriminate one particular facial expression among others (Summerfield *et al*. 2006; Summerfield & Koechlin, 2008). The instruction effectively creates an internal 'template' (reflecting conditional expectations) against which to match the incoming sensory information (Dayan *et al*. 1995; Dosher & Lu, 1999) and therefore favours the anticipation of one perceptual alternative. The data support our first hypothesis, demonstrating that participants from both control and patient groups tended to overvalue conditional expectations when faced with an unexpected emotional stimulus (Summerfield & Koechlin, 2008). This effect was different for the different facial expressions. For threatening stimuli (angry and fearful faces), there was a performance cost for facial expressions that were incongruent with, and perceptually close to, the expression named in the instruction. Specifically, for angry face stimuli, participants made more errors (Fig. 2*a*) and RTs were longer (Fig. 2*b*) when they were asked to look out for fearful faces compared with the other contexts. Similarly, for fearful faces, participants made more errors (Fig. 2*c*) and RTs were longer (Fig. 2*d*) when they were asked to look out for angry faces compared with the other contexts. By contrast, for happy faces, emotion discrimination was improved when the emotion displayed in the face stimulus was congruent with the instruction as compared to when it was incongruent (no perceptual challenge was involved; see Fig. 3*a*, *b*). Finally, we did not find a clear dependence on prior information for the neutral expression (Fig. 3*c*, *d*). This could be related to the fact that neutral faces are often interpreted as emotional stimuli (Carré *et al*. 2010). Therefore, it might have been difficult to clearly categorize neutral faces in our task. In general, our results are consistent with previous reports showing that prior expectations influence decision making (Summerfield & Egner, 2009; Evans *et al*. 2011). However, previous studies have shown this effect with gabor patches (Summerfield & Koechlin, 2008) or non-emotional faces (Summerfield *et al*. 2006), whereas we reproduced such an effect, for the first time, with emotions.

Prior expectations in schizophrenia

Our second finding showed that, for threatening emotions, patients were more biased by prior expectations provided by the instruction than were controls. For angry faces, participants with schizophrenia made more errors than controls when they were asked to look out for fearful faces compared with the other contexts (Fig. 2*a*). Similarly, for fearful faces, RTs of participants with schizophrenia were longer than RTs of controls when they were asked to look out for angry faces compared with the other contexts (Fig. 2*d*). This result is in line with previous findings suggesting that patients afford too much weight to prior expectations, leading to a tendency to discount evidence that runs counter to a prior belief and thus to the inappropriate maintenance of this belief (i.e. belief inflexibility) (Fletcher & Frith, 2009; Stephan *et al*. 2009). To measure the ability to integrate disconfirmatory evidence into a prior belief, previous studies have used scenarios designed to lure participants into a belief that then became implausible when additional information was presented (Woodward *et al*. 2008). Other studies have used the Maudsley Assessment of Delusions Scale (Wessely *et al*. 1993), a standardized assessment of delusions that inquires about the evidence for the delusion and includes an item about the possibility of being mistaken (Garety *et al*. 2005). In both studies, patients with schizophrenia demonstrated an unwillingness to change their strong belief (Garety *et al*. 2005; Woodward *et al*. 2008). Our results reveal that a bias towards prior expectations may also extend to the domain of emotion perception. Note that we did not observe such an increased bias in patients for happy emotions, which supports previous findings showing that this probabilistic reasoning impairment is more prominent for salient stimuli such as threatening emotions (Blackwood *et al*. 2001).

Typically, poor performances may confound the interpretation of specific cognitive functions thought to be impaired in schizophrenia. Indeed, patients may show impairment on one task as an artefact of a more general cognitive deficit and not because of a specific cognitive deficit. In the current study, as would be expected, the patients performed worse overall, but the question is whether this general performance decrement caused the more specific pattern of results seen when contrasting the conditions. Contrary to this possibility, we found that the higher prior expectation effect for the fearful and angry expressions in patients was not related to their overall performance on the task, suggesting that this cognitive dysfunction is unlikely to result from a general deficit.

The higher dependence on prior expectations as compared to the incoming sensory evidence in schizophrenia fits well with findings showing impaired performance in schizophrenia patients in the interference condition of the Stroop test, in which participants are required to name the colour of a stimulus word while ignoring its conflicting meaning (e.g. the word 'red' printed in blue ink) (Henik & Salo, 2004). In our task, this conflict between a cue and a stimulus occurs when the face stimulus is incongruent with the general context provided by the instruction cue (e.g. in case of an angry face in a fearful context). Therefore, the higher instruction-induced bias against sensory evidence observed in the present study and the increased Stroop-like interference effect between the instruction and the incongruent stimulus could arise, at least in part, from the same core dysfunctional mechanism (i.e. a dysfunction of the anterior cingulate cortex and left inferior frontal cortex during conflict monitoring) (Carter *et al*. 1997, 2001; Kerns *et al*. 2005; Krabbendam *et al*. 2009). This hypothesis should be further investigated in the future.

The hyper-sensitivity to context is the opposite of what would be predicted by the prevailing view on context processing and schizophrenia, which proposes a hypo-sensitivity to context (Servan-Schreiber *et al*. 1996; Barch *et al*. 2001; Hemsley, 2005; Chambon *et al*. 2008; Barbalat *et al*. 2009). Evidence for this latter theory comes from studies using paradigms such as the continuous performance task (CPT), in which participants are presented with a sequence of letters and are instructed to respond to a prespecified probe (X) only if it follows a particular contextual cue (A) (Servan-Schreiber *et al*. 1996). The different results (i.e. a hyper-sensitivity to context in our task *versus* a hyposensitivity in the CPT) may arise from the difference in the basic nature of the paradigms. As pointed out previously (Park *et al*. 2003; Hemsley, 2005), context is a composite construct with various dimensions referring to separate processes, which may be impaired differentially. In a broad sense, context refers to an internal representation of any task-relevant information that can be used to mediate an appropriate behavioural response (Servan-Schreiber *et al*. 1996). In the cognitive tasks that result in reduced context processing in schizophrenia, the contextual signal refers to a prior stimulus indicating the specific tasks-sets that subsequently need to be performed by the participant (e.g. if the contextual cue is the letter A, then respond to the target only if it is the letter X) (Servan-Schreiber *et al*. 1996; Barch *et al*. 2001; Hemsley, 2005). By contrast, in the present study, the contextual signal is a prior instruction that biases the participant to predict the proceeding stimuli. In these latter types of task, patients tend to show increased sensitivity to contextual information (Blackwood *et al*. 2001; Freeman, 2007; Fletcher & Frith, 2009; Summerfield & Egner, 2009).

Relationship between prior expectations and persecutory delusions: a preliminary finding

When we split the patient group into two subgroups based on the presence of persecutory delusions, we found that patients with persecutory delusions demonstrated a heightened overall dependency on prior expectations as compared to patients without such symptoms. Persecutory delusions were not related to a prior bias effect alone, as we found no differences in the prior bias for happy or neutral expressions between patients with and without persecutory delusions. Nor were they related to emotion processing impairment alone, as we found no differences in overall RTs or accuracy for all emotional expressions between patients with and without persecutory delusions. Instead, our data support the notion that an interaction between a dependency on prior expectations and threatening emotion processing might specifically underlie delusions of persecution (Blackwood *et al*. 2001). However, because of the small sample size of the subgroups ($n=9$ in each), this result should be considered as a preliminary finding.

Conclusions

Because the information entering the visual system is inherently ambiguous and abundant, and the processing of visual information is limited by computational capacity, the brain needs to prioritize stimulus processing. In this context, rapid detection of threatening stimuli is crucial for self-protection and, as such, could be understood as a routine evolutionary process. The neurocognitive processes responsible for fast and appropriate threat detection could be dependent on two mechanisms: attention, which prioritizes stimulus processing on the basis of motivational relevance (e.g. threatening as compared to neutral stimuli); and expectation, which constrains visual interpretation on the basis of prior likelihood (e.g. being more efficient at detecting a threatening face when the general context in which the person is acting is threatening) (Summerfield & Egner, 2009). We suggest that this dependence on prior expectations is heightened in schizophrenia, reflecting belief inflexibility for threatening stimuli, and could underlie persecutory delusions. This finding may warrant the investigation of psychological approaches focused on correcting this bias through cognitive remediation training.

Acknowledgements

We thank T. Bobin for assistance with patient recruitment and I. Dumontheil and J. Roiser for commenting on previous versions of this manuscript. This work was supported by the Royal Society, UK and the Fyssen Foundation, France. S.J.B. has a Royal Society University Research Fellowship. G.B. was supported by a grant from the Fyssen Foundation.

Declaration of Interest

None.

References

Averbeck BB, Evans S, Chouhan V, Bristow E, Shergill SS (2011). Probabilistic learning and inference in schizophrenia. *Schizophrenia Research* **127**, 115–122.

Barbalat G, Chambon V, Franck N, Koechlin E, Farrer C (2009). Organization of cognitive control within the lateral prefrontal cortex in schizophrenia. *Archives of General Psychiatry* **66**, 377–386.

Barch DM, Carter CS, Braver TS, Sabb FW, MacDonald 3rd A, Noll DC, Cohen JD (2001). Selective deficits in prefrontal cortex function in medication-naive patients with schizophrenia. *Archives of General Psychiatry* **58**, 280–288.

Blackwood NJ, Howard RJ, Bentall RP, Murray RM (2001). Cognitive neuropsychiatric models of persecutory delusions. *American Journal of Psychiatry* **158**, 527–539.

Carré JM, Morrissey MD, Mondloch CJ, McCormick CM (2010). Estimating aggression from emotionally neutral faces: which facial cues are diagnostic? *Perception* **39**, 356–377.

Carter CS, MacDonald AW, Ross LL, Stenger VA (2001). Anterior cingulate cortex activity and impaired self-monitoring of performance in patients with schizophrenia: an event-related fMRI study. *American Journal of Psychiatry* **158**, 1423–1428.

Carter CS, Mintun M, Nichols T, Cohen JD (1997). Anterior cingulate gyrus dysfunction and selective attention deficits in schizophrenia: [^{15}O]H_2O PET study during single-trial Stroop task performance. *American Journal of Psychiatry* **154**, 1670–1675.

Chambon V, Franck N, Koechlin E, Fakra E, Ciuperca G, Azorin J, Farrer C (2008). The architecture of cognitive control in schizophrenia. *Brain* **131**, 962–970.

Dayan P, Hinton GE, Neal RM, Zemel RS (1995). The Helmholtz machine. *Neural Computation* **7**, 889–904.

Dosher BA, Lu ZL (1999). Mechanisms of perceptual learning. *Vision Research* **39**, 3197–3221.

Evans S, Shergill SS, Chouhan V, Bristow E, Collier T, Averbeck BB (2011). Patients with schizophrenia show increased aversion to angry faces in an associative learning task. *Psychological Medicine* **41**, 1471–1479.

Fletcher PC, Frith CD (2009). Perceiving is believing: a Bayesian approach to explaining the positive symptoms of schizophrenia. *Nature Reviews. Neuroscience* **10**, 48–58.

Freeman D (2007). Suspicious minds: the psychology of persecutory delusions. *Clinical Psychology Review* **27**, 425–457.

Garety PA, Freeman D, Jolley S, Dunn G, Bebbington PE, Fowler DG, Kuipers E, Dudley R (2005). Reasoning, emotions, and delusional conviction in psychosis. *Journal of Abnormal Psychology* **114**, 373–384.

Garety PA, Hemsley DR, Wessely S (1991). Reasoning in deluded schizophrenic and paranoid patients. Biases in performance on a probabilistic inference task. *Journal of Nervous and Mental Disease* **179**, 194–201.

Hemsley DR (2005). The schizophrenic experience: taken out of context? *Schizophrenia Bulletin* **31**, 43–53.
Henik A, Salo R (2004). Schizophrenia and the Stroop effect. *Behavioral and Cognitive Neuroscience Reviews* **3**, 42–59.
Kay SR, Fiszbein A, Opler LA (1987). The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophrenia Bulletin* **13**, 261–276.
Kerns JG, Cohen JD, MacDonald 3rd AW, Johnson MK, Stenger VA, Aizenstein H, Carter CS (2005). Decreased conflict- and error-related activity in the anterior cingulate cortex in subjects with schizophrenia. *American Journal of Psychiatry* **162**, 1833–1839.
Krabbendam L, O'Daly O, Morley LA, van Os J, Murray RM, Shergill SS (2009). Using the Stroop task to investigate the neural correlates of symptom change in schizophrenia. *British Journal of Psychiatry* **194**, 373–374.
Miller MB, Chapman JP, Chapman LJ, Collins J (1995). Task difficulty and cognitive deficits in schizophrenia. *Journal of Abnormal Psychology* **104**, 251–258.
Moritz S, Woodward TS (2005). Jumping to conclusions in delusional and non-delusional schizophrenic patients. *British Journal of Clinical Psychology* **44**, 193–207.
Moritz S, Woodward TS (2006). A generalized bias against disconfirmatory evidence in schizophrenia. *Psychiatry Research* **142**, 157–165.
Park S, Lee J, Folley B, Kim J (2003). Schizophrenia: putting context in context. *Behavioral and Brain Sciences* **26**, 98–99.
Pessoa L (2008). On the relationship between emotion and cognition. Nature Reviews. *Neuroscience* **9**, 148–158.
Servan-Schreiber D, Cohen JD, Steingard S (1996). Schizophrenic deficits in the processing of context. A test of a theoretical model. *Archives of General Psychiatry* **53**, 1105–1112.
Stephan KE, Friston KJ, Frith CD (2009). Dysconnection in schizophrenia: from abnormal synaptic plasticity to failures of self-monitoring. *Schizophrenia Bulletin* **35**, 509–527.
Summerfield C, Egner T (2009). Expectation (and attention) in visual cognition. *Trends in Cognitive Sciences* **13**, 403–409.
Summerfield C, Egner T, Greene M, Koechlin E, Mangels J, Hirsch J (2006). Predictive codes for forthcoming perception in the frontal cortex. *Science* **314**, 1311–1314.
Summerfield C, Koechlin E (2008). A neural representation of prior information during perceptual inference. *Neuron* **59**, 336–347.
Tottenham N, Tanaka JW, Leon AC, McCarry T, Nurse M, Hare TA, Marcus DJ, Westerlund A, Casey BJ, Nelson C (2009). The NimStim set of facial expressions: judgments from untrained research participants. *Psychiatry Research* **168**, 242–249.
Wechsler D (1999). *Wechsler Abbreviated Scale of Intelligence* (*WASI*). Harcourt Assessment: San Antonio, TX.
Wessely S, Buchanan A, Reed A, Cutting J, Everitt B, Garety P, Taylor PJ (1993). Acting on delusions. I: Prevalence. *British Journal of Psychiatry* **163**, 69–76.
Woodberry KA, Giuliano AJ, Seidman LJ (2008). Premorbid IQ in schizophrenia: a meta-analytic review. *American Journal of Psychiatry* **165**, 579–587.
Woodward TS, Buchy L, Moritz S, Liotti M (2007). A bias against disconfirmatory evidence is associated with delusion proneness in a nonclinical sample. *Schizophrenia Bulletin* **33**, 1023–1028.
Woodward TS, Moritz S, Cuttler C, Whitman JC (2006). The contribution of a cognitive bias against disconfirmatory evidence (BADE) to delusions in schizophrenia. *Journal of Clinical and Experimental Neuropsychology* **28**, 605–617.
Woodward TS, Moritz S, Menon M, Klinge R (2008). Belief inflexibility in schizophrenia. *Cognitive Neuropsychiatry* **13**, 267–277.