Decision Processes and the Slowing of Simple Choices in Schizophrenia

Andrew Heathcote, Anna Suraev, Samuel Curley,
and Qinchun Gong
University of Newcastle

Jonathon Love
University of Amsterdam

Patricia T. Michie
University of Newcastle

Individuals diagnosed with schizophrenia have slowed response times (RT). We examined the role of decision processes in the slowing of simple choice responses. We updated Schatz's (1998) meta-analysis of deficits in speed and extend it to systematically examine the effects of schizophrenia on choice accuracy. We then report an experiment requiring decisions about motion direction, which we analyzed using an evidence accumulation model of choice, the linear ballistic accumulator (LBA; Brown & Heathcote, 2008). By simultaneously taking into account all aspects of behavior, the LBA was more sensitive to deficits than mean RT or accuracy alone. It also identified the 2 underlying causes of slowing: more cautious decisions (i.e., requiring more evidence before making a decision) and perceptual deficits. The schizophrenia group displayed strong sequential effects that were captured by the response on the previous trial affecting the relative amount of evidence required for choice in the LBA. These results illustrate that evidence accumulation models provide a sensitive tool that can be used to identify the cognitive mechanisms causing slowing in schizophrenia.

General Scientific Summary
Slowed responding and inaccurate choices are common in people with schizophrenia. We use a cognitive model of how choices are made to identify the underlying causes of these deficits in a motion discrimination task.

Keywords: schizophrenia, choice response time, cognitive modeling, cognitive deficits

Supplemental materials: http://dx.doi.org/10.1037/abn0000117.supp

Schizophrenia is characterized by disturbances in cognition, behavior, emotion, and perception (Frangou & Kington, 2004). Clinical symptoms are both positive (e.g., hallucinations, delusions, disorganized speech and behavior) and negative (e.g., poverty of thought and speech, social withdrawal, blunted and flat affect, loss of motivation and experience of pleasure) and are accompanied by cognitive impairments in numerous domains (Keefe & Harvey, 2012). Cognitive impairments are stronger predictors of functional outcomes and quality of life than clinical symptoms (Nuechterlein et al., 2011) and are not simply a result of clinical symptoms or medication effects (Green, Kern, Braff, & Mintz, 2000). Green et al. (2004) identified social cognition, working memory, attention or vigilance, verbal and visual learning

and memory, reasoning and problem solving, and speed of processing as showing the most significant deficits. We focus on the latter, looking at response time (RT) in tasks requiring a rapid choice among more than one option.

In particular, we studied deficits in choice RT processes in schizophrenia using an evidence accumulation model of choice processing, the linear ballistic accumulator (LBA; Brown & Heathcote, 2008).[1] This model provides a specific understanding of the causes of slowing that is not afforded by traditional psychometric approaches, which separately analyze partial summary measures such as mean RT and accuracy (e.g., Silverstein, 2008). In contrast to such traditional approaches, the LBA addresses all aspects of choice behavior, simultaneously accounting for not only mean RT and accuracy but also the entire distribution of RT (e.g., its variability and positively skewed shape). It is this rich characterization that enables insight into which several potential underlying factors cause deficits.

Andrew Heathcote, Anna Suraev, Samuel Curley, and Qinchun Gong, School of Psychology, University of Newcastle; Jonathon Love, Department of Psychology, University of Amsterdam; Patricia T. Michie, School of Psychology, University of Newcastle.

Correspondence concerning this article should be addressed to Andrew Heathcote, who is now at the Division of Psychology, University of Tasmania, Social Sciences Building, Sandy Bay 7005, Tasmania, Australia. E-mail: andrew.heathcote@utas.edu.au

[1] We also did a parallel analysis with another commonly used evidence accumulation model, the drift diffusion model (DDM; Ratcliff & McKoon, 2008). These results are presented in the supplementary materials, as the DDM produced a worse fit to our data than the LBA and was unable to accommodate key effects that were well explained by the LBA.

Slowing and Sequential Effects in Schizophrenia

Wells and Kelley (1922) summarized early research showing markedly slower simple RT (i.e., RT in tasks with only one response option) in patients with psychosis compared with healthy participants. Subsequent research found that simple RT slowing in schizophrenia is associated with poorer prognosis and poorer outcomes and occurs across many tasks (Nuechterlein, 1977). Other early work focused on the way in which simple RT is affected by prior events, particularly a shift in a target's sensory modality relative to the last trial (e.g., Mowrer, 1941; Sutton, Hakerem, Zubin, & Portnoy, 1961). Later research found that heightened *sequential effects*—that is, carryover effects of stimuli and responses from the previous trial on responding in the present trial—are a pervasive characteristic of schizophrenia in a range of tasks, particularly a tendency to perseverate (e.g., Abbruzzese, Ferri, & Scarone, 1996; Lyon & Gerlach, 1988; Lyon, Mejsholm, & Lyon, 1986; Yogev, Hadar, Gutman, & Sirota, 2003).

Although there are some exceptions, slowing is generally found to be disproportionally greater for choice than simple RT (Benton, Jentsch, & Wahler, 1959; Hintze, Bebenek, Kuhn-Dymecka, Wronska, & Wciorka, 2006; Karras, 1967; Krieger, Lis, & Gallhofer, 2001; Krieger, Lis, Cetin, Gallhofer, & Meyer-Lindenberg, 2005; Zahn et al., 1998). Rapid choice paradigms, where RT is typically less than 2 to 3 s, have been widely used to study impairment in a range of perceptual and cognitive processes in schizophrenia because such paradigms are relatively free from the strategies and multiple-stage higher order judgments required in more complex decision-making tasks. Schatz (1998) provided a meta-analysis of mean RT in 196 mainly choice conditions drawn from 40 papers and found pervasive slowing for schizophrenia relative to control groups.

In the first part of this paper, we provide the results of a meta-analysis that updates the work of Schatz (1998). In the second part of the paper, we applied the LBA to data from a new experiment. In the experiment, patients with a diagnosis of schizophrenia (hereafter *patients*) and control participants decided the direction of motion in random-dot kinematogram (RDK) stimuli—clouds of mostly randomly moving dots with a proportion that move coherently either left or right. In light of their pervasive nature in schizophrenia, we analyzed these data for sequential effects by using the last choice as a covariate in the LBA analysis. We also analyzed the data for sequential effects using autocorrelations between choices (i.e., correlations between the choice made on the present trial and the last trial). This analysis does not require the assumptions made by the LBA model and so provides an independent test of whether sequential effects differ between patients and controls. In order to set the stage for the analysis of sequential effects using the LBA model, we report the results of the autocorrelation-based test first, along with traditional separate analyses of RT and accuracy. First, however, we give an overview of evidence accumulation models.

Evidence Accumulation Models

The LBA and a number of other similar models (e.g., Ratcliff & Smith, 2004) share the idea that the process underlying choice is the accumulation of information (evidence) over time. Response selection occurs when a threshold amount of evidence favoring a particular choice has accrued. A core prediction implied by this mechanism is that choice involves a speed–accuracy trade-off; as the threshold is increased, choices become more accurate because they are based on more evidence, but they are also slower because it takes longer to collect the evidence. Model parameters accounting for speed–accuracy trade-offs have been linked to specific neural structures (Forstmann et al., 2008, 2010; Mansfield, Karayanidis, Jamadar, Heathcote, & Forstmann, 2011).

Evidence accumulation models are fit to all aspects of choice behavior simultaneously. That is, they do not separately rely on summary statistics—such as error rates or the fastest, slowest, or average RTs—that in isolation provide only a partial and potentially misleading characterization and do not map in any simple way to underlying cognitive processes. Rather, they account for all such statistics together, and they do so in terms of the psychological causes of primary interest. For example, slowing may be due to (a) lower quality evidence that accumulates at a slower rate, (b) requiring more evidence (i.e., a higher threshold), and/or (c) slowing in the nondecision process (e.g., the time to encode the stimulus or produce a response). Evidence accumulation modeling can determine which of these—either singly or in combination—is responsible for slowing.

Because evidence accumulation modeling directly addresses underlying causes rather than the often-complicated constellation of effects in different statistics summarizing behavior, it has proven fruitful in understanding a range of deficits. These include deficits associated with factors ranging from sleep deprivation and alcohol use (Ratcliff & Van Dongen, 2011; van Ravenzwaaij, Dutilh, & Wagenmakers, 2012) to depression and anxiety (Ho et al., 2014; White, Ratcliff, Vasey, & McKoon, 2010a, 2010b). To our knowledge, however, this type of modeling has not previously been applied to choice RT deficits in schizophrenia.

Given the key role played by interaction between speed and accuracy in evidence accumulation models, our update of Schatz's (1998) meta-analysis reported in the first part of this paper addressed accuracy as well as mean RT. Although it has been noted in individual studies that choice accuracy is generally reduced in schizophrenia, we are not aware of any systematic and wide-ranging analysis of this issue. If slowing in schizophrenia is due to reduced quality of evidence, we would expect accuracy to also be systematically lower. If it is due to a speed–accuracy trade-off, then accuracy should be increased. If both factors interplay to varying degrees across different tasks and conditions, a mixed pattern may emerge.

Meta-Analysis of Choice RT Studies

In a Brinley plot (Brinley, 1965; Salthouse & Somberg, 1982), results from two groups are plotted against each other (e.g., in Figure 1, the schizophrenia group on the abscissa and the control group on the ordinate). Schatz (1998) adopted Cerella's (1994) methods for the analysis of Brinley plots to study slowing in aging and applied them to slowing in schizophrenia. Although we agree with Ratcliff, Spieler, and McKoon's (2000) critique of this analysis methodology (see the supplementary materials for a detailed discussion of Ratcliff et al.'s critique), Schatz's meta-analysis remains highly informative in the way it systematically demonstrates the pervasive nature of slowing in schizophrenia across a broad range of tasks. The Brinley plot is also useful for compactly
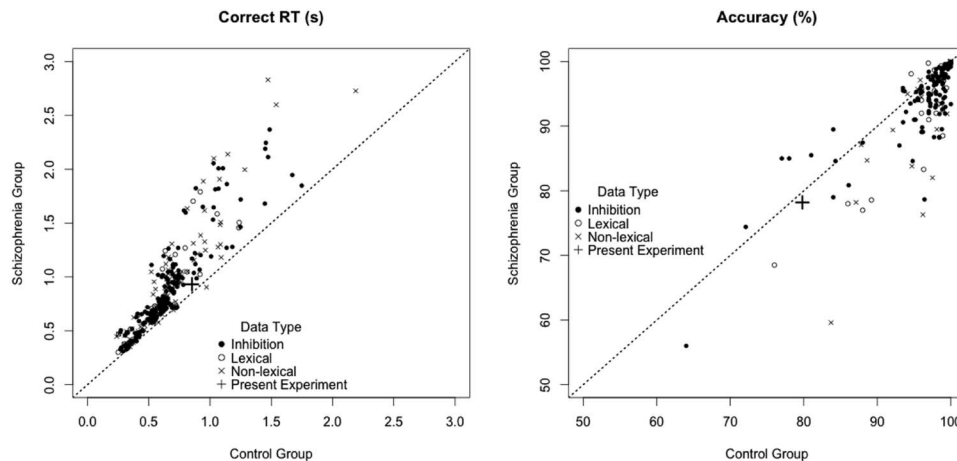
Figure 1. Average response time in seconds (RT) for correct responses and percentage of correct responses for choice RT tasks. The dotted diagonal line indicates equality between the two groups. Tasks were classified as inhibition, lexical, or nonlexical. Results for the present experiment are for the 10% coherence condition. In order to improve the resolution of other results, the RT panel excludes Schneider et al.'s (2011) inhibition task data, where SZ = 6.1 s and Control = 4.5 s.

representing results, with slowing being indicated by points falling above the main diagonal, as shown in the left-hand panel of Figure 1. In the supplementary materials, we describe how we substantially expanded Schatz's sample (except that we excluded simple RT) to 83 papers and 317 conditions. The supplementary materials also provide a detailed analysis of our findings, which we summarize here.

The left-hand panel of Figure 1 shows that slowing is a highly consistent result, with only 6/314 cases (1.9%) below the main diagonal (i.e., the dotted line indicating equal results for the two groups), although in many cases the slowing is only minor. The right-hand panel of Figure 1 shows that, in general, accuracy was lower for patients than controls (i.e., points below the main diagonal), although this effect was not as consistent as for slowing, with 32/189 cases (17%) above the main diagonal. This overall pattern of findings is not consistent with slowing being due to a speed–accuracy trade-off, which predicts increased accuracy.

Although these results do not support a speed–accuracy trade-off, they also did not reveal the strong positive correlations that would be expected if patients had deficits in the quality of the evidence on which they based their decisions, as poor quality evidence would be expected to both slow decisions and make them more error prone. A potential reason for the lack of correlation is that both factors are in play, with patients attempting to compensate for errors caused by lower evidence quality by responding more cautiously. If they failed to fully compensate, they could still be less accurate while having an even bigger disadvantage in speed. The resulting correlation between speed and accuracy differences would be reduced or even absent.

These possibilities must remain speculations with respect to the studies that contributed to the meta-analysis, as we have reached the limits of the inferences that can be made based on mean RT and accuracy data. The experiment we performed enabled us to overcome these limitations. It also allowed us to address another limitation of most of the studies in the meta-analysis—accuracy performance near ceiling—which makes it hard to detect speed–accuracy trade-off effects. Difficulty in our study was carefully calibrated to avoid both ceiling and floor effects on accuracy.

Experiment

We measured the RT and accuracy of choices about left-versus-right motion in RDK stimuli in a group of schizophrenia patients and matched controls. Overall difficulty, as determined by the proportion of coherently moving dots, was calibrated for each participant to avoid floor and ceiling effects in accuracy, which facilitates better estimation of model parameters. Tailoring difficulty to participants' abilities also encouraged engagement among patients which, along with the use of a simple task, aimed to address motivational factors that could confound comparisons (Joyce & Huddy, 2004).

We also manipulated the difficulty of the task (the degree of coherence of moving dots), with participants experiencing two types of randomly intermixed trials: easier (i.e., higher coherence) trials and harder (i.e., lower coherence) trials. This was done to test the construct validity of the LBA model. Previous research (e.g., Forstmann et al., 2008) has found that coherence manipulations selectively influence parameters related to the rate of evidence accumulation. Thus, if the LBA model is valid, it should provide a good fit to the data when only its rate parameters (and not other parameters, such as evidence thresholds) are allowed to vary as a function of the difficulty.

Method

Participants

Informed consent was obtained prior to commencement of the study. Exclusion criteria for patients were a history of

neurological trauma, a diagnosis of intellectual disability, or current drug or alcohol dependence. For controls, the same exclusion criteria were applied, as well as exclusion based on a diagnosis of a mental illness. Twenty-six participants with schizophrenia or schizoaffective disorder were recruited from a clozapine clinic and paid AU$55 for their participation, which included an extra session completing a stop signal task that is not reported here. The treating psychiatrist made diagnoses according to the *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.; *DSM–IV*; American Psychiatric Association, 1994), the current version at the time of testing. The treating clinician or case manager confirmed the participant's diagnosis, his or her ability to provide informed consent, and whether the study would be suitable for him or her.

One patient was excluded because he or she declined to complete the clinical interview and six due to accuracy not significantly greater than chance overall in the choice task. Table 1 summarizes the remaining patients' demographics and medication taken at the time of testing as well as the characteristics of excluded participants. A brief interview was undertaken to assess exclusion criteria for potential controls, and 19 were recruited to match the patients on age and gender. Controls were recruited from three sources: an undergraduate student pool, the Hunter Medical Research Institute volunteer register, and the local community. Of those who passed the exclusion criteria, four were from a student

pool and were awarded course credit for their participation. A further six were from the volunteer register, and nine were from the local community; they all received reimbursement of $40 to cover their costs.

Experimental Apparatus and Stimuli

The RDK stimuli were presented on a personal computer in a quiet room. They consisted of 40 dots that moved randomly within a circular area 50 mm in diameter, being redrawn 30 times per second, except for a proportion that moved coherently 45° either to the top left or top right. Participants were instructed to judge the direction of the movement by pressing the *Z* key with their left hand for left movement or the *slash* key with their right hand for right movement (these keys appear at the lower left and lower right, respectively, of the standard U.S. keyboard).

To avoid floor and ceiling effects, participants were assigned a harder pair of motion coherence levels (5% and 10%) or an easier pair (10% and 20%). The assignment was based on their performance during practice trials at the beginning of the session, with less able participants assigned to the easier pair and more able participants assigned to the harder pair. In this way, we aimed to avoid the near-ceiling performance in accuracy reported in many previous studies.

Table 1
Characteristics of Participants Whose Data Were Included in Analyses

Characteristics	Patients with schizophrenia (n = 19)		Controls (n = 19)		Parametric test	
	Mean	SD	Mean	SD	t value	p value
Age (years)	37.6	6.5	38.8	9.0	.5	.6
Education[a]	2.8	1.4	5	.8	6.1	<.001
Weschler Test of Adult Reading (WTAR) scaled score	94	15.4	109.4	24.7	2.3	.03
Letter Number Sequence (LNS) scaled score	8.1	2.5	11.5	3.7	3.4	.002
Positive symptoms (SAPS)	3.4	3.5				
Negative symptoms (SANS)	10.2	3.2				
Disorganization symptoms	2.4	2.3				
	n	%	n	%		
Gender						
Male	16	84.2	16	84.2		
Female	3	15.8	3	15.8		
Diagnosis						
Schizophrenia	17	89.5				
Schizoaffective disorder	2	10.5				
Medication type						
Clozapine	18	94.7				
Other atypical antipsychotics	10	52.6				
Typical antipsychotics	1	5.3				
Antidepressants	7	36.8				
Lithium carbonate	1	5.3				
Anticonvulsants	3	15.8				

Note. Relative to the included patients, the seven excluded patients were slightly older (44.6 years) and had more negative symptoms (12.5) but had similar positive symptoms (2.8), disorganization (2.3), LNS (7.7), WTAR (94), and education (2.6) values. SAPS = Scale for the Assessment of Positive Symptoms; SANS = Scale for the Assessment of Negative Symptoms.
[a] Education codes: 0 = below Year 10, 1 = Year 10 or fourth form (school certificate or equivalent), 2 = Year 11 or fifth form (leaving certificate or equivalent), 3 = Year 12 or sixth form (HSC or equivalent), 4 = technical college or TAFE college, 5 = graduate (bachelor's degree), 6 = postgraduate (master's degree or PhD). HSC = Higher School Certificate; TAFE = Technical and Further Education.

On each trial, a blank screen preceded each stimulus for 0.25 s followed by a fixation point for 0.25 s, and the stimulus was presented for 3 s. The experimental task comprised nine blocks with 49 trials in each block. In the first block, the coherence of dots started at 65% and decreased gradually until it reached 20%. The second and third blocks contained three difficulty levels: 5%, 10%, and 20% coherence. Average accuracy was determined for the easier pair (10% and 20%) and the harder pair (5% and 10%). Whichever pair had an average accuracy closer to 75% was used as the difficulty level for the remaining blocks, which provided the data that were analyzed. Between blocks, participants were encouraged to take a rest as required and continued onto the next block by pressing the space bar. The task took approximately 20 min to complete.

Participants were encouraged to perform quickly but also to try to be accurate. For correct responses, participants were provided with feedback on their RT in order to increase motivation and reinforce rapid responding. Incorrect choices were followed by *incorrect* displayed on the screen and *too slow* if participants failed to respond within 3 s. The latter trials were removed from all analyses.

Psychometric Measures

Participants completed the Letter Number Sequence (LNS), a subtest of the Wechsler Adult Intelligence Scale—Third Edition (WAIS–III; Wechsler, 1997) as a measure of working memory. Participants also completed the Wechsler Test of Adult Reading (WTAR; Wechsler, 2001), which assesses premorbid intellectual functioning in adults.

Patients were administered the Scale for the Assessment of Positive and Negative Symptoms (SAPS/SANS; Andreasen, 1983, 1984), which took 45–60 min to administer, depending on the extent of the symptoms. The SAPS/SANS has been found to have good interrater reliability, modest internal consistency, and strong construct validity (Peralta & Cuesta, 1994). Positive symptoms (SAPS score) consisted of the combined global scores for hallucinations and delusions; disorganization symptoms were composed of global scores for bizarre behavior and positive formal thought disorder; and negative symptoms (SANS score) combined the global scores for affective flattening, alogia, avolition or apathy, and anhedonia or asociality (Andreasen, Arndt, Alliger, Miller, & Flaum, 1995). Scores on the attention subscale of the SAPS/SANS were omitted, as this is not considered to be a core component of negative symptomology (Blanchard & Cohen, 2006).

Procedure

The experiment took place over two separate sessions, typically a week, and at a minimum of 24 hr apart. The average time between the first and second sessions was 8.8 days ($SD = 3.1$) for patients and 9.4 days ($SD = 6.5$) for controls. In the first session, all participants completed a demographic questionnaire, the choice RDK task, the LNS, and the WTAR. In the second session, all participants completed a stop signal task (not reported here); patients also completed the clinical interview.

Results

We first separately analyzed RT for correct responses and accuracy using analyses of variance (ANOVAs) with factors for current stimulus (left vs. right motion), last response (left vs. right), difficulty (higher vs. lower coherence), and group (patients vs. controls). We report all effects significant at the 0.05 level. We then examined autocorrelations in order to compare sequential effects in patients and controls.

Mean RT was significantly slower for the lower (1.1 s) than the higher (0.93 s) coherence stimuli, $F(1, 36) = 92.5$, $p < .001$, $\eta^2 = 0.46$. The same was true for median RT (0.99 s vs. 0.86 s), $F(1, 36) = 85.8$, $p < .001$, $\eta^2 = 0.44$. RT standard deviation was also greater for lower (0.38 s) than higher (0.3 s) coherence stimuli, $F(1, 36) = 74.4$, $p < .001$, $\eta^2 = 0.37$, and this effect interacted with group, $F(1, 36) = 9.5$, $p = .003$, $\eta^2 = 0.07$, due to a larger effect in the controls. Accuracy was lower for low (71%) than high (84%) coherence stimuli, $F(1, 36) = 69.3$, $p < .001$, lower for left (75%) than right (80%) stimuli, $F(1, 36) = 4.9$, $p = .03$, $\eta^2 = 0.52$, and lower for the patients (72%) than controls (82%), $F(1, 36) = 9.7$, $p = .003$, $\eta^2 = 0.21$.

We also examined the probability of making a left-versus-right response to check for response bias. There was a relatively small but highly significant bias in favor of right responses, $p(\text{left}) = .47$, $t(37) = 3.7$, $p < .001$, $r = .51$. An ANOVA did not find any significant main effects of group or last response on bias.

The tests just described involving the group factor are confounded by our difficulty calibration procedure, which assigned higher coherence (10% and 20%)—and hence easier—pairs more often to the patients (16/19) than controls (11/19). Hence, we repeated the tests with only data from the 10% coherence condition common to all participants.

Mean RT was slower in the schizophrenia (1 s) than the control (0.93 s) group, but the difference was not significant, $F(1, 36) = 1.0$, $p = .32$, $\eta^2 = 0.026$. The same was true for the median (0.93 s vs. 0.85 s), $F(1, 36) = 1.2$, $p = .27$, $\eta^2 = 0.033$. RT standard deviation was greater in the schizophrenia (0.36 s) than the control (0.32 s) group, but the difference was not significant, $F(1, 36) = 1.0$, $p = .32$, $\eta^2 = 0.027$. Accuracy was lower in the patients (78%) than controls (83%), but again the difference was not significant, $F(1, 36) = 1.8$, $p = .18$, $\eta^2 = 0.049$. The same pattern of results about response bias found with all of the data held for data from only the 10% coherence conditions.

In order to examine sequential effects, we fit an autoregressive model[2] of order 1 to each participant's binary response choices (i.e., left or right) for each of the six experimental blocks. This yielded estimates of the proportion of variance in the current choice explained by the previous choice (i.e., lag 1 squared correlation). In an ANOVA on autocorrelation estimates with block and group as factors, there was a significantly greater squared autocorrelation for patients (0.16) than controls (0.11), $F(1, 36) = 6.4$, $p = .016$, $\eta^2 = 0.15$, but no main effect of block or interaction between block and group. When we fit higher order autoregressive models (i.e., models allowing for lag 2 and higher autocorrelations), the same pattern of results

---

[2] We used the R (R Core Team, 2014) *ar* function with the default Yule–Weker method.

was found for lag 1 autocorrelations, and no significant effects were found for higher order autocorrelations. On average, the lag 1 correlation was positive for the patient group (0.03), indicating a tendency to repeat responses (i.e., perseverance), whereas in the control group it was negative (−0.03), indicating a tendency to alternate responses.

Discussion

Although there were trends for both slower and less accurate responding in the patients than controls for 10% coherent stimuli, neither difference was significant, and corresponding effect sizes were small ($\eta^2 = 0.027$ and 0.049, respectively). As shown in Figure 1, such small affects are by no means unusual, despite the fact that slowing of RT is often accepted as a universal characteristic of patients with schizophrenia. We return to the reasons why this is the case in the General Discussion.

In contrast to the speed and accuracy results, we found a significant group difference in sequential effects ($\eta^2 = 0.15$). The analysis of lag 1 squared autocorrelations indicated a stable tendency throughout the experiment for patients' responses to be more influenced by the immediate past response than controls' responses. These sequential effects in patients mainly manifested as a tendency to repeat the last response, consistent with the tendency to perseverate found in a range of other tasks. The heightened sequential effect that we observed in patients—a finding that, to our knowledge, is novel in rapid RT tasks (both of the simple and choice types)—is further investigated in the next section.

Model-Based Analysis

We fit the LBA using quantile maximum probability estimation (Heathcote & Brown, 2004; Heathcote, Brown, & Mewhort, 2002), which minimizes a measure of misfit called the deviance. Deviance was calculated based on RT distribution quantified by the 10th, 30th, 50th (median), 70th, and 90th percentiles of the RT data. A set of 256 LBA variants was generated as described in Donkin, Brown, and Heathcote (2011). These models were simplifications of a most flexible "top" model that allowed model parameters to vary with experimental factors and with factors designating the different accumulators (Table 2; see, e.g., Heathcote & Hayes, 2012; Heathcote & Love, 2012; Rae, Heathcote, Donkin, Averell, & Brown, 2014, for applications of this method). Fitting begins with the simplest model, for which it is easy to use heuristics to generate

reasonable guesses from which to start the optimization process. The best fit of that model then provides starting guesses for slightly more complex models and so on, with more complex models being fit from multiple start points, helping to ensure optimal fits of even quite complex models.

When fitting a model, choices have to be made about how parameter estimates are influenced by experimental factors. These decisions are in part based on the meaning of the parameters and in part on convention. There are also practical considerations; allowing the parameter specification to be too complex can result in an unmanageably large number of parameters to estimate and/or *overfitting*: small improvements in fit that result in parameter estimates that vary in meaningless ways in order to accommodate small random variations in the data. Parameters governing variability, like the rate standard deviation and level of start-point noise, are harder to estimate than other parameters, and so conventionally they are allowed to vary with fewer factors (see Donkin et al., 2011, for more details).

We addressed overfitting using the Akaike information criterion (AIC; Burnham & Anderson, 2004), which adds to the misfit measure (deviance) a complexity penalty proportional to the number of model parameters. Adding parameters must improve model fit (i.e., reduce deviance), but if the improvement is small it will be outweighed by the penalty, so AIC will select the simpler model. The model variant that strikes the best balance between complexity and fit will have the smallest AIC. A smaller AIC by 10 or more indicates a strong preference, 3–10 indicates substantial evidence, and less than 3 indicates an equivocal result (Wagenmakers & Farrell, 2004). The model selected by AIC indicates which of the parameters estimated in the top model are important for accommodating the variability in data.

Figure 2 describes the LBA model, and Table 3 lists its parameters. The RT predicted by the model is the sum of decision time—the time from when evidence starts accumulating until that evidence first crosses a threshold—and nondecision time. Nondecision time is the sum of the time to encode the stimulus and the time to produce the response corresponding to the decision (see Smith & Ratcliff, 2009). There is an accumulator for each possible choice. Each independently accumulates evidence favoring its choice, and a response corresponding to that choice is made if it is the first to reach its threshold.

Even under identical conditions, human RT varies considerably from trial to trial (Luce, 1986). In the LBA, RT varies because of the effects of two kinds of noise. The mean rate of evidence accumulation varies from trial to trial according to a normal

Table 2

Definitions of Model Factors

Name	Type	Level
Difficulty (D)	Experimental	Easy = lower coherence; hard = higher coherence
Stimulus (S)	Experimental	Left = left motion; right = right motion
Last response (L)	Experimental	Left = left motion; right = right motion
Response (R)	Accumulator	Left = left motion; right = right motion
Match (M)	Accumulator	True = matches stimulus; false = mismatches stimulus

Note. Experimental factors describe the experimental design; accumulator factors differentiate the two accumulators, either according to the response they represent or the match between the stimulus and the response they represent.
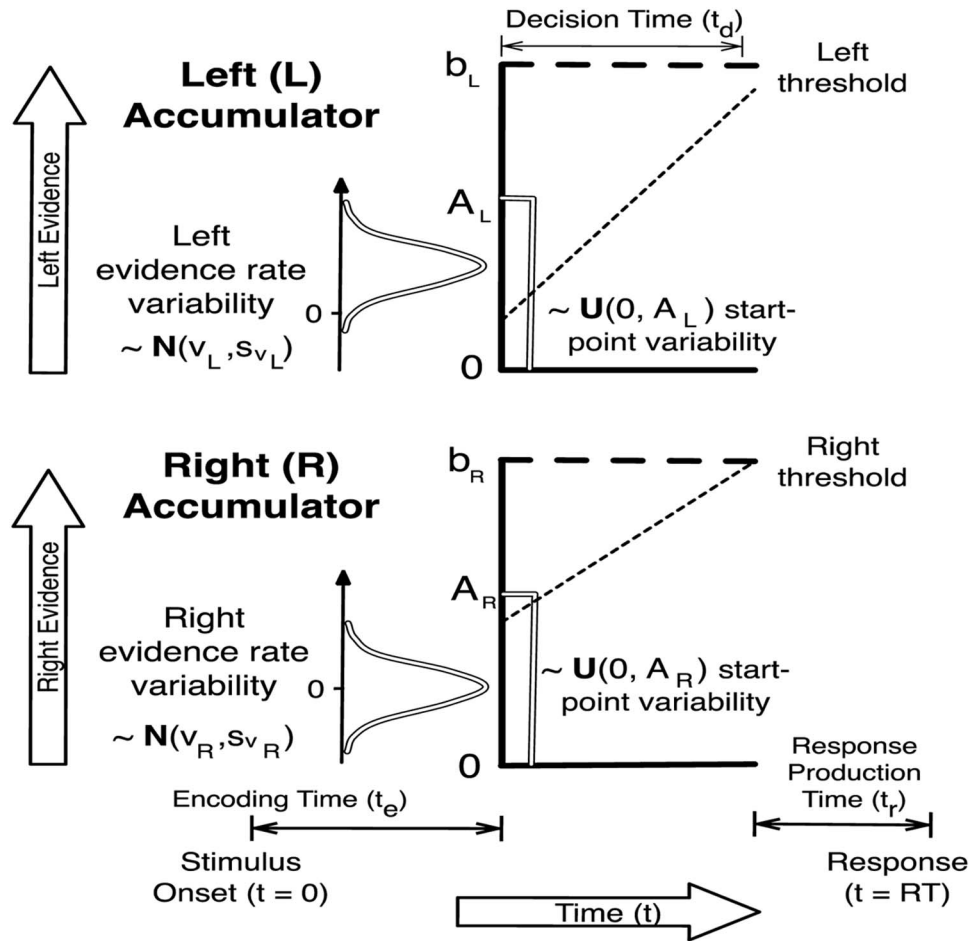
Figure 2. The linear ballistic accumulator (LBA) model for a left-versus-right decision. $U()$ indicates a uniform distribution and $N()$ indicates a normal distribution. In this example, the stimulus is leftward motion, so the left accumulator is the matching accumulator, with a larger mean rate, and the right accumulator is the mismatching accumulator, with a smaller mean rate. Nondecision time $= t_{er} = t_e + t_r$.

distribution, reflecting factors such as fluctuations in attention or stimuli between nominally equivalent trials. The starting point varies from trial to trial according to a uniform distribution, reflecting bias caused by factors such as the response made on the last trial. Accumulating more evidence by setting a higher threshold can reduce errors caused by start-point biases, but errors caused by rate variability remain, consistent with the fact that often even quite slow decisions are not perfectly accurate.

Each accumulator has its own parameters for the rate mean and standard deviation, range of start-point noise, and threshold. Threshold differences (specified by the response factor in Table 2) cause response bias, with unbiased responding corresponding to equal thresholds across accumulators. When accuracy is above chance, the rate for the matching accumulator (i.e., the left accumulator when viewing a left-moving stimulus or the right accumulator when viewing a right-moving stimulus) is greater than for the mismatching accumulator. The match factor (see Table 2) is used to specify such differences. The rate for an accumulator determines how quickly its corresponding response is made, whereas differences in rates between accumulators mainly affect

accuracy (i.e., accurate responding occurs when there is a large rate difference between matching and mismatching accumulators).

To allow for sequential effects, we fit the data broken down by the last response type (i.e., left or right) and allowed this factor to bias the LBA threshold. A lower threshold for the accumulator corresponding to the last response causes a bias toward repeating responses, whereas a bias in favor of the other accumulator causes a tendency toward alternating responses.

Table 3 also shows the mappings between experimental factors and LBA parameters in the top model, which has 29 parameters. The threshold can vary with the response factor, allowing for response bias, and the last trial factor, allowing for sequential effects ($2 \times 2 = 4$ parameters). In parameterizing rates, we used a match factor that allows for differences between the accumulators that do and do not match the stimulus. For example, if the stimulus is left motion, the left-response accumulator is the matching accumulator with a higher mean rate, and the right-response accumulator is the mismatching accumulator with a lower mean rate (see Figure 2). Better motion discrimination corresponds to a larger advantage in the mean rate for the matching accumulator

Table 3
Definition of Linear Ballistic Accumulator (LBA) Parameters (Two Left-Hand Columns)

Symbol	Name	Top model	AIC model
b	Response threshold	R, L	R, L
a	Proportional start-point variability range	1	1
v	Mean rate of evidence accumulation	L, S, D, M	D, M
s_v	Rate standard deviation	S, D, M	D, M
t_{er}	Mean nondecision time	1	1

Note. AIC = Akaike information criterion; R = response; L = last response; S = stimulus; D = difficulty; M = match. The variable $a = A/b$, where A is the range of start-point variability. The "top model" column indicates that parameter estimates are allowed to vary with the factors (see Table 2) in the top model. An entry of 1 in the final column indicates that a single value was estimated for all combinations of factor levels. The number of model variants estimated by Donkin et al.'s (2011) method equals $2^{n-1} = 256$, where $n = 9$ is the number of factors in the "top model" column. One is subtracted from n because even in the simplest model, the mean rate varied with M to allow for above-chance responding. The final column specifies factors in the model selected by the AIC summed over participants.

over the mismatching accumulator. The mean rate varied with stimulus type (left vs. right motion), difficulty (easy vs. hard), the last response (left vs. right) factor, and the match factor ($2 \times 2 \times 2 \times 2 = 16$ parameters).

One accumulation process parameter must take a fixed value in at least one condition in order to identify parameter estimates (Donkin, Brown, & Heathcote, 2009). The rate standard deviation was allowed to vary with all of the same factors as the rate mean, except last response, but is fixed at 1 for one condition [so $(2 \times 2 \times 2) - 1 = 7$ parameters]. The same values of the start-point variability and nondecision time parameters were estimated for all experimental conditions and accumulators (two parameters). Note that in Figure 2, A denotes the range of start-point variability. As shown in Table 3, we estimated the range as a proportion of the threshold: $a = A/b$ (note that $0 \le A \le b$, so $0 \le a \le 1$ for both accumulators), so start-point variability increases in proportion to the threshold. This proportionality assumption largely held even when we did not enforce it, so we chose this parameterization to simplify the model.

Results

Deviance values can be used to test between variants within a model, where a difference in deviance has a χ^2 distribution with degrees of freedom equal to the difference in the number of parameters. These χ^2 tests confirmed the AIC selection, with the decrease in fit from the top model to the AIC model not being significant, $\chi^2(646) = 383$, $p = 1$. Also, there was clear evidence of overfitting in the top model, which had a number of implausible parameter estimates. In light of these findings, we focus on the AIC-selected model (13 parameters per subject) for further analysis.

The fit of the AIC-selected model is displayed in Figure 3. The left-hand panels show the observed and predicted error rates for controls (top row) and patients (bottom row), and the right-hand panels show the distribution of correct RT by displaying the 10th percentile (i.e., reflecting fast responses), the 50th percentile (i.e., the median), and the 90th percentile (i.e., reflecting slow responses). The 30th and 70th percentiles used in fitting showed the same trends but are omitted for clarity. The AIC-selected model provides an accurate and detailed description, capturing, for example, the positive skew characteristic of RT distributions as

evidenced by a greater gap between the median and the 90th percentile than between the median and the 10th percentile. The model also clearly captures the effect of difficulty, validating the LBA's explanation of this manipulation in terms of only rate parameter differences.

We do not display analogous results for error RTs, as errors were less frequent than correct responses, and consequently error RT percentiles had quite wide 95% confidence intervals. However, we note that the model's predictions mostly fell within these intervals and that the model captured the general tendency for slower incorrect than correct responses. Details are provided in the supplementary materials.

Parameter Analysis

The AIC-selected model displayed a large set of significant effects, including group effects in both rates and thresholds. We address the group effects after first discussing the effects common to both groups. As expected, the rate of the accumulator that matches the stimulus (1.7) was greater than the rate of the accumulator that does not match the stimulus (-1.2), $F(1, 36) = 70$, $p < .001$, $\eta^2 = 0.62$. The rate standard deviation for the mismatching accumulator (1.1) was greater than that for the matching accumulator (0.8), $F(1, 36) = 9.4$, $p = .004$, $\eta^2 = 0.15$, as has been found in other LBA applications (e.g., Heathcote & Love, 2012; Rae et al., 2014). The difference between matching and mismatching accumulators was larger for easy (3.7) than hard (2.2) stimuli, $F(1, 36) = 9.4$, $p = .004$, $\eta^2 = 0.094$, consistent with higher quality evidence, and hence greater accuracy, for easy than hard stimuli.

As evidence quality is also a function of the rate standard deviation, it is useful to calculate a quantity analogous to the signal detection theory sensitivity measure (d'). A higher rate standard deviation will more often cause errors due to increasing the number of trials on which mismatching evidence exceeds matching evidence. The d' measure takes this into account by expressing the difference between matching and mismatching means in units of the pooled standard deviation:

$$d' = (v|match - v|mismatch)/\left(s_v^2|match + s_v^2|mismatch\right)/2$$
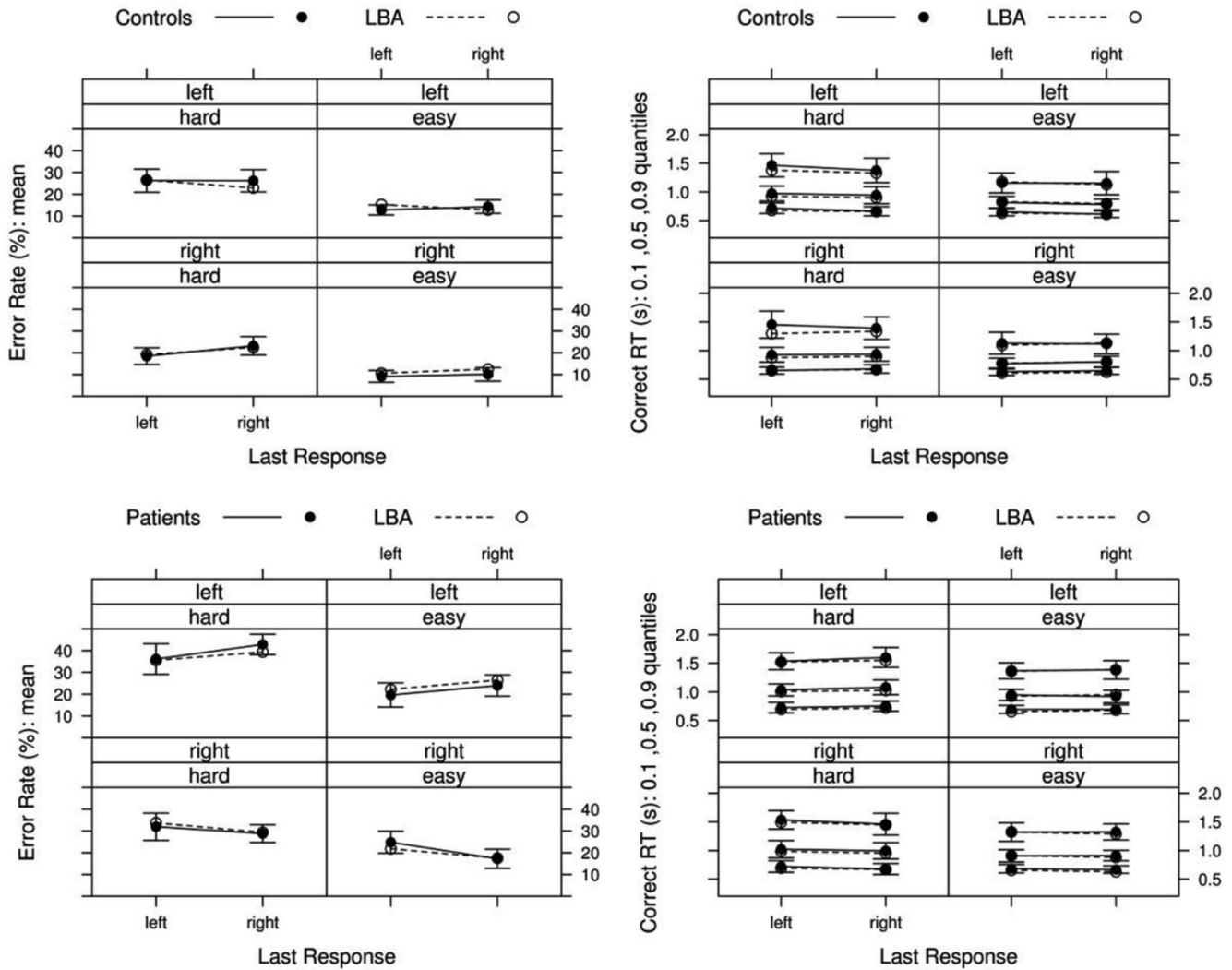
Figure 3. Error rates (left column) and response time (RT) distribution (10th, 50th, and 90th percentiles for correct responses (right column) as a function of difficulty and current and last stimulus (left or right motion) with superimposed linear ballistic accumulator (LBA) model fits. Error bars indicate 95% within-subject confidence intervals (Morey, 2008).

The d' for easy stimuli (5.5) was significantly greater than for hard stimuli (3.6), $F(1, 36) = 15.2$, $p < .001$, $\eta^2 = 0.44$. Finally, consistent with the overall right-response bias noted earlier, there was a lower threshold estimate for the right (2.1) than left (2.3) accumulator, $F(1, 36) = 5.1$, $p = .03$, $\eta^2 = 0.12$.

For group effects,[3] the matching versus mismatching rate difference was larger for controls (3.8) than patients (2.1), $F(1, 36) = 6.7$, $p = .014$, $\eta^2 = 0.11$. This was mainly due to the mismatching rate being significantly smaller for controls than patients (-2.3 vs. 0.2), $t(35.8) = 1.57$, $p = .01$, $r = .25$. Controls also had a much larger d' (6.2) than patients (2.9), $F(1, 36) = 8.8$, $p = .005$, $\eta^2 = 0.2$. Patients had a larger threshold than controls (2.8 vs. 1.6), $F(1, 36) = 4.45$, $p = .04$, $\eta^2 = 0.11$. Patients and controls did not differ significantly in nondecision time (0.24 s vs. 0.22 s), $F(1, 36) = 0.31$, $\eta^2 = 0.01$.

The top LBA model allowed the last response to potentially affect both rates and thresholds, but AIC indicated that the last-response effect was mediated only by threshold. To form a model-based index of perseverance (i.e., the tendency to repeat the last response), we first calculated indices of response bias. A bias toward the left response occurs when the left accumulator has a lower threshold than the right accumulator, and vice versa for a right bias. Left-response bias (i.e., $b_R - b_L$) when the last response was left was summed with right bias (i.e., $b_L - b_R$) when the last response was right to form a measure of perseverance. For both components, positive values indicate a response bias toward the last response, so their sum is the overall strength of the model's tendency to be biased toward the last response.

---

[3] The supplementary materials report individual parameter estimates and follow-up analyses of the group effect, removing potentially outlying participants. There was no substantive change to conclusions about group differences, with conclusions in most cases being strengthened.

Recall that lag 1 autocorrelations are positive when responses tend to be repeated. In the overall sample, the model-based perseverance index correlated highly with the lag 1 autocorrelations, $r = .73$, $t(36) = 6.4$, $p < .001$. The correlation was almost entirely due to patients, $r = .83$, $t(17) = 6.1$, $p < .001$, and was nonsignificant in controls, $r = .16$, $t(17) = 0.67$, $p = .16$.

Discussion

The relatively simple 13-parameter AIC-selected LBA model provided an accurate and detailed description of important aspects of the data, ranging from large effects like the shape of correct RT distribution to more subtle effects like those of the last response on accuracy. AIC model selection confirmed mediation of the difficulty effect by rate mean and variability parameters and of stimulus type (left vs. right motion) effects through response bias (i.e., the difference between the left vs. right accumulator thresholds). Both findings are plausible in terms of past findings and the meanings conventionally attributed to these parameters (Donkin et al., 2011), confirming the validity of our application of the LBA model in the present case. Mediation of the last-response effects through threshold changes also provides a mechanism that explains the perseveration seen in patients.

We quantified the combined effects of the rate means and standard deviations using a d' measure of sensitivity that indexes the quality of the evidence derived from the stimulus. That is, a higher value of d' indicates that this evidence is better able to support discrimination of the direction of motion. The group effect on sensitivity was highly significant ($\eta^2 = 0.2$). Indeed, it was actually larger than the analogous difficulty effect on sensitivity. Difficulty significantly affected accuracy, and this effect was clear because it was not masked by a change in the evidence threshold. However, the group effect on sensitivity was not reflected in a significant effect on accuracy because it was masked by a threshold that was larger for patients than controls ($\eta^2 = 0.11$). The higher threshold reduces the effects of random biases caused by start-point noise and so increases accuracy, partially countering the deleterious effects of the patients' deficits in sensitivity.

The sensitivity deficit in patients was in large part due to a markedly increased mean rate for the accumulator corresponding to the incorrect response. For controls, the mean rate was more than 2 standard deviations below zero, so on most trials, evidence for the incorrect response decreased over the course of accumulation. For patients, however, the mean was slightly positive, so evidence for the incorrect response increased over the course of accumulation, although still at a lesser rate than correct evidence. This result indicates that patients were more prone to interpreting random motion in the stimulus as supporting the incorrect response, whereas they were about equally good at extracting evidence for the correct response from the coherent motion in the stimulus.

General Discussion

Slow and error-prone choices in simple tasks are pervasive and heavily researched features of schizophrenia, but these effects are quite variable, particularly with respect to accuracy (see Figure 1), and their underlying psychological causes remain unclear. We

used evidence accumulation modeling, particularly the LBA model (Brown & Heathcote, 2008), to examine the causes. This approach provides a unified account of all aspects of choice behavior in terms of the cognitive mechanisms that determine both RT and accuracy.

We used the LBA to analyze the performance of schizophrenia patients and matched controls in a motion direction binary (left vs. right) classification task. Patients were a little slower and a little more error prone than controls, but neither difference by itself was significant. We did, however, find significantly stronger sequential effects in patients, who showed an increased tendency to repeat the last response. Although perseverance has been identified as a feature of schizophrenia in complex tasks (e.g., Abbruzzese et al., 1996), as well as guessing about random sequences (e.g., Lyon & Gerlach, 1988; Lyon et al., 1986), it has not, to our knowledge, previously been identified as a feature of responding in simple choice tasks.

Our LBA analysis provided clear answers about a set of interrelated causes of slowing, error proneness, and perseverance in the motion direction classification task. Perseverance had a single cause: a bias to require less evidence for the response made on the last trial. Error proneness had two causes: the bias toward the last response (which also caused perseverance) and deficits in the quality of perceptual evidence. Slowing also had two underlying causes: the perceptual deficits (which also affected accuracy) and patients using a higher evidence threshold (i.e., requiring a greater quantity of evidence before making a decision) than controls. Setting a higher threshold was likely a strategy used by patients to reduce the inaccuracy caused by the lower quality of their perceptual evidence (i.e., a speed–accuracy trade-off).

The LBA was not only able to uncover the causes of these deficits; it also identified what did not cause group differences. In particular, patients did not display any deficit in the combined time to begin sampling evidence from a stimulus after it first appeared or the time to produce a response after it was selected (i.e., nondecision time). The mechanisms uncovered by the LBA also suggest an explanation for why error proneness is not always observed in simple choice tasks (see Figure 1): variability in the degree to which patients used a speed–accuracy trade-off strategy (e.g., if the speed–accuracy trade-off was large enough, there would be no accuracy difference, only slowing).

These findings demonstrate that evidence accumulation modeling has substantial promise for improving the detection of deficits in choice tasks displayed by people with schizophrenia, and for explaining their underlying causes, just as it has for other deficits (e.g., Ho et al., 2014; Ratcliff & Van Dongen, 2011; van Ravenzwaaij et al., 2012; White et al., 2010a, 2010b). Clearly, however, our work represents only a first step that is limited to one type of perceptual choice. Much research on schizophrenia has been based on choice tasks using other types of perceptual and lexical information, as well as choice paradigms addressing inhibition and interference phenomena (see Schatz, 1998, for a taxonomy). Performance in all of these tasks could also be investigated using the same approach that we applied to our task. In the following, we discuss the implications and limitations of our results and identify issues that could be pursued in future research.

Limitations and Future Directions

Evidence accumulation rates. Schizophrenia is associated with sensitivity deficits in a variety of sensory processes (Butler, Silverstein, & Dakin, 2008), including RDK motion detection (Chen, Bidwell, & Holzman, 2005), consistent with our findings about deficits in the quality of perceptual evidence. However, evidence accumulation rates can also be affected by attentional and working memory capacity (Eidels, Donkin, Brown, & Heathcote, 2010; Schmiedek, Oberauer, Wilhelm, Süß, & Wittmann, 2007). Impairments in these capacities have long been an important explanatory concept of the cognitive deficits in schizophrenia (e.g., Nuechterlein & Dawson, 1984). Hence, further research should investigate whether these factors also contribute to group differences in evidence accumulation rates.

In stimulus–response compatibility and flanker paradigms, Kappenman et al. (2012) concluded, based on event-related potentials, that patients failed to activate the correct response rather than failing to suppress the incorrect response, whereas we found a failure to suppress evidence for the incorrect response. Further research will be required to resolve whether task differences are responsible for these different conclusions.

Nondecision time. Although schizophrenia is associated with psychomotor slowing in complex sequence production, this is largely in the planning component (Jogems-Kosterman, Zitman, Van Hoof, & Hulstijn, 2001; Kim, Lee, Choi, & Goh, 2009). Our finding of there being no nondecision time deficit is consistent with planning not being a factor in the simple button-press response required in our study. An important limitation of current evidence accumulation modeling approaches is that they do not differentiate stimulus encoding time from response production time, so trade-offs between these components cannot be detected. A reviewer suggested that encoding time might be slowed in schizophrenia. If this were the case and production time were also speeded to an equal degree, our methods would not detect any effect.

RT variability. As discussed in the supplementary materials, our meta-analysis found an increased RT standard deviation in patients (e.g., Fassbender, Scangos, Lesh, & Carter, 2014; Kieffaber et al., 2006; Rentrop et al., 2010), which is consistent with the LBA threshold increase we observed, which increases RT variability. Smyrnis et al. (2009) reported a similar effect in eye movement times in a simple RT paradigm and fit their data with the LATER model (Carpenter & Williams, 1995), which corresponds to a single LBA accumulator, except that it has no start-point noise. They found a significant increase in the evidence accumulation rate standard deviation for their schizophrenia group, with no significant difference in the mean rate from controls. In contrast, our LBA fits revealed no significant group effect on the rate standard deviation. The divergent pattern of results could be due to task differences (simple vs. choice RT, saccadic vs. manual responses). However, it would be interesting to also investigate the possibility that start-point noise and sequential effects may have a role to play in the paradigm.

Sequential effects. There is a long history of studies on healthy participants performing simple choice tasks like the one used in our experiment that have examined sequential effects related to stimulus and response repetition and alternation (e.g., Bertelson, 1961; Kirby, 1976). They found a small but reliable bias to repeat responses, or make repeat responses faster, when the response-to-stimulus interval is less than 0.5 s and a bias to alternate responses, or make alternate responses faster, when it is longer (Soetens, Boer, & Hueting, 1985). Our response-to-stimulus interval was around two seconds on average, so the general tendency we found in our controls toward an alternation bias is to be expected.

The early history of the study of slowing in schizophrenia focused on the effect on simple RT of prior events, with Zubin (1975) proposing that prior events leave behind facilitatory or inhibitory neuronal traces that tend to last longer in schizophrenia patients. In a description reminiscent of Zubin's (1975) explanation, Soetens et al. (1985) attributed repetition effects in healthy participants—occurring when the gap between trials is much shorter than in our experiment—to "decaying memory traces related to the structural pathway of the reaction process" (p. 598). With a compatible stimulus–response mapping such as the one we used, repetition effects have been localized to response-related stages of processing (Soetens, 1998). Thus, our finding of strong repetition (i.e., perseverance) supports the decay of memory traces in the response stage being abnormally slowed in schizophrenia. However, it is also possible that perseverance was more strategic in nature, with patients deliberately setting lower thresholds because they believed that repetitions are more likely than alternations. Future research might contrast these explanations by manipulating the response-to-stimulus interval, as longer intervals are associated with strategic effects, and shorter intervals are associated with automatic effects (Soetens et al., 1985).

Sample size, power, and speed–accuracy trade-offs. Our study consisted of a relatively small sample of treatment-resistant patients from a clozapine clinic who were chronically ill. These characteristics, and others specific to our sample (see Table 1; e.g., they were predominantly male), may limit the generalizability of our findings. Although our sample size was relatively small, it was sufficient to produce a significant group difference in the correlation between choices on successive trials, but group differences in speed and accuracy failed significance. However, our LBA analysis was able to reveal underlying differences in evidence quality and caution because it simultaneously took into account all of the differences between patient and control performance. Hence, it was not a lack of power, but rather the trade-off between speed and accuracy, that resulted in nonsignificant effects in separate analyses of mean RT and accuracy. White, Ratcliff, Vasey, and McKoon (2009) also found that their dysphoric and nondysphoric groups did not differ significantly in mean RT and accuracy but that there were significant underlying differences in evidence quality (see also White et al., 2010a, 2010b).

Our results demonstrate the utility of evidence accumulation model analyses as a sensitive method of identifying the underlying causes of deficits in choice processes in abnormal populations. However, a limitation of this approach is that a fairly large number of trials in each experimental condition have to be collected for each participant. For future research, it is worth noting that our design probably represents a lower bound in this regard and that, if possible, more trials per participant, as well as larger and more varied samples of patients, are desirable in order to realize the full benefits of evidence accumulation modeling.

References

Abbruzzese, M., Ferri, S., & Scarone, S. (1996). Performance on the Wisconsin Card Sorting Test in schizophrenia: Perseveration in clinical subtypes. *Psychiatry Research, 64,* 27–33. http://dx.doi.org/10.1016/0165-1781(96)02927-7

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Andreasen, N. C. (1983). *Scale for the Assessment of Negative Symptoms (SANS).* Iowa City: University of Iowa.

Andreasen, N. C. (1984). *Scale for the Assessment of Positive Symptoms (SAPS).* Iowa City: University of Iowa.

Andreasen, N. C., Arndt, S., Alliger, R., Miller, D., & Flaum, M. (1995). Symptoms of schizophrenia: Methods, meanings, and mechanisms. *Archives of General Psychiatry, 52,* 341–351. http://dx.doi.org/10.1001/archpsyc.1995.03950170015003

Benton, A. L., Jentsch, R. C., & Wahler, H. J. (1959). Simple and choice reaction times in schizophrenia. *A.M.A. Archives of Neurology and Psychiatry, 81,* 373–376.

Bertelson, P. (1961). Sequential redundancy and speed in a serial two-choice responding task. *The Quarterly Journal of Experimental Psychology, 13,* 90–102. http://dx.doi.org/10.1080/17470216108416478

Blanchard, J. J., & Cohen, A. S. (2006). The structure of negative symptoms within schizophrenia: Implications for assessment. *Schizophrenia Bulletin, 32,* 238–245. http://dx.doi.org/10.1093/schbul/sbj013

Brinley, J. F. (1965). Cognitive sets, speed and accuracy of performance in the elderly. In A. T. Welford & J. E. Birren (Eds.), *Behavior, aging and the nervous system* (pp. 114–149). Springfield, IL: Thomas.

Brown, S. D., & Heathcote, A. (2008). The simplest complete model of choice response time: Linear Ballistic Accumulation. *Cognitive Psychology, 57,* 153–178.

Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: Understanding AIC and BIC in model selection. *Sociological Methods & Research, 33,* 261–304. http://dx.doi.org/10.1177/0049124104268644

Butler, P. D., Silverstein, S. M., & Dakin, S. C. (2008). Visual perception and its impairment in schizophrenia. *Biological Psychiatry, 64,* 40–47. http://dx.doi.org/10.1016/j.biopsych.2008.03.023

Carpenter, R. H. S., & Williams, M. L. L. (1995). Neural computation of log likelihood in control of saccadic eye movements. *Nature, 377,* 59–62. http://dx.doi.org/10.1038/377059a0

Cerella, J. (1994). Generalized slowing in Brinley plots. *Journal of Gerontology, 49,* P65–P71. http://dx.doi.org/10.1093/geronj/49.2.P65

Chen, Y., Bidwell, L. C., & Holzman, P. S. (2005). Visual motion integration in schizophrenia patients, their first-degree relatives, and patients with bipolar disorder. *Schizophrenia Research, 74*(2–3), 271–281. http://dx.doi.org/10.1016/j.schres.2004.04.002

Donkin, C., Brown, S. D., & Heathcote, A. (2009). The overconstraint of response time models: Rethinking the scaling problem. *Psychonomic Bulletin & Review, 16,* 1129–1135. http://dx.doi.org/10.3758/PBR.16.6.1129

Donkin, C., Brown, S., & Heathcote, A. (2011). Drawing conclusions from choice response time models: A tutorial using the linear ballistic accumulator. *Journal of Mathematical Psychology, 55,* 140–151. http://dx.doi.org/10.1016/j.jmp.2010.10.001

Eidels, A., Donkin, C., Brown, S. D., & Heathcote, A. (2010). Converging measures of workload capacity. *Psychonomic Bulletin & Review, 17,* 763–771. http://dx.doi.org/10.3758/PBR.17.6.763

Fassbender, C., Scangos, K., Lesh, T. A., & Carter, C. S. (2014). RT distributional analysis of cognitive-control-related brain activity in first-episode schizophrenia. *Cognitive, Affective, & Behavioral Neuroscience, 14,* 175–188. http://dx.doi.org/10.3758/s13415-014-0252-4

Forstmann, B. U., Dutilh, G., Brown, S., Neumann, J., von Cramon, D. Y., Ridderinkhof, K. R., & Wagenmakers, E.-J. (2008). Striatum and pre-SMA facilitate decision-making under time pressure. *Proceedings of the National Academy of Sciences of the United States of America, 105,* 17538–17542. http://dx.doi.org/10.1073/pnas.0805903105

Forstmann, B. U., Schafer, A., Anwander, A., Neumann, J., Brown, S. D., Wagenmakers, E.-J., . . . Turner, R. (2010). Cortico-striatal connections predict control over speed and accuracy in perceptual decision making. *Proceedings of the National Academy of Sciences of the United States of America, 107,* 15916–15920. http://dx.doi.org/10.1073/pnas.1004932107

Frangou, S., & Kington, J. (2004). Schizophrenia. *Medicine, 32,* 21–25. http://dx.doi.org/10.1383/medc.32.7.21.36676

Green, M. F., Kern, R. S., Braff, D. L., & Mintz, J. (2000). Neurocognitive deficits and functional outcome in schizophrenia: Are we measuring the "right stuff"? *Schizophrenia Bulletin, 26,* 119–136. http://dx.doi.org/10.1093/oxfordjournals.schbul.a033430

Green, M. F., Nuechterlein, K. H., Gold, J. M., Barch, D. M., Cohen, J., Essock, S., . . . Marder, S. R. (2004). Approaching a consensus cognitive battery for clinical trials in schizophrenia: The NIMH-MATRICS conference to select cognitive domains and test criteria. *Biological Psychiatry, 56,* 301–307. http://dx.doi.org/10.1016/j.biopsych.2004.06.023

Heathcote, A., & Brown, S. (2004). Reply to Speckman and Rouder: A theoretical basis for QML. *Psychonomic Bulletin & Review, 11,* 577–578. http://dx.doi.org/10.3758/BF03196614

Heathcote, A., Brown, S., & Mewhort, D. J. K. (2002). Quantile maximum likelihood estimation of response time distributions. *Psychonomic Bulletin & Review, 9,* 394–401. http://dx.doi.org/10.3758/BF03196299

Heathcote, A., & Hayes, B. (2012). Diffusion versus linear ballistic accumulation: Different models for response time with different conclusions about psychological mechanisms? *Canadian Journal of Experimental Psychology, 66,* 125–136. http://dx.doi.org/10.1037/a0028189

Heathcote, A., & Love, J. (2012). Linear deterministic accumulator models of simple choice. *Frontiers in Psychology, 3,* 292. http://dx.doi.org/10.3389/fpsyg.2012.00292

Hintze, B., Bebenek, A., Kuhn-Dymecka, A., Wronska, A., & Wciorka, J. (2006). Working memory dysfunction in patients suffering from schizophrenia and their first-degree relatives. *Archives of Psychiatry and Psychotherapy, 8,* 17–29.

Ho, T. C., Yang, G., Wu, J., Paulus, M. P., Cassey, P., Brown, S. D., Connolly, C. G., . . . Yang, T. T. (2014). Functional connectivity of negative emotional processing in adolescent depression. *Journal of Affective Disorders, 155,* 65–74.

Jogems-Kosterman, B. J. M., Zitman, F. G., Van Hoof, J. J., & Hulstijn, W. (2001). Psychomotor slowing and planning deficits in schizophrenia. *Schizophrenia Research, 48*(2–3), 317–333. http://dx.doi.org/10.1016/S0920-9964(00)00097-9

Joyce, E., & Huddy, V. (2004). Defining the cognitive impairment in schizophrenia. *Psychological Medicine, 34,* 1151–1155. http://dx.doi.org/10.1017/S0033291704003472

Kappenman, E. S., Kaiser, S. T., Robinson, B. M., Morris, S. E., Hahn, B., Beck, V. M., . . . Luck, S. J. (2012). Response activation impairments in schizophrenia: Evidence from the lateralized readiness potential. *Psychophysiology, 49,* 73–84. http://dx.doi.org/10.1111/j.1469-8986.2011.01288.x

Karras, A. (1967). The effect of stimulus-response complexity on the reaction time of schizophrenics. *Psychonomic Science, 7,* 75–76. http://dx.doi.org/10.3758/BF03331085

Keefe, R. S., & Harvey, P. D. (2012). Cognitive impairment in schizophrenia. *Handbook of Experimental Pharmacology, 213,* 11–37. http://dx.doi.org/10.1007/978-3-642-25758-2_2

Kieffaber, P. D., Kappenman, E. S., Bodkins, M., Shekhar, A., O'Donnell, B. F., & Hetrick, W. P. (2006). Switch and maintenance of task set in schizophrenia. *Schizophrenia Research, 84*(2–3), 345–358. http://dx.doi.org/10.1016/j.schres.2006.01.022

Kim, C. Y., Lee, G., Choi, H., & Goh, J. (2009). Repeated measures of reaction times among patients with schizophrenia. *Clinical Psychopharmacology and Neuroscience, 7,* 20–22.

Kirby, N. H. (1976). Sequential effects in two-choice reaction time: Automatic facilitation or subjective expectancy? *Journal of Experimental Psychology: Human Perception and Performance, 2,* 567–577. http://dx.doi.org/10.1037/0096-1523.2.4.567

Krieger, S., Lis, S., & Gallhofer, B. (2001). Cognitive subprocesses and schizophrenia. A. Reaction-time decomposition. *Acta Psychiatrica Scandinavica, 104,* 18–27. http://dx.doi.org/10.1034/j.1600-0447.2001.104s408018.x

Krieger, S., Lis, S., Cetin, T., Gallhofer, B., & Meyer-Lindenberg, A. (2005). Executive function and cognitive subprocesses in first-episode, drug-naive schizophrenia: An analysis of N-back performance. *The American Journal of Psychiatry, 162,* 1206–1208. http://dx.doi.org/10.1176/appi.ajp.162.6.1206

Luce, R. D. (1986). *Response times: Their role in inferring elementary mental organization.* New York, NY: Oxford University Press.

Lyon, N., & Gerlach, J. (1988). Perseverative structuring of responses by schizophrenic and affective disorder patients. *Journal of Psychiatric Research, 22,* 261–277. http://dx.doi.org/10.1016/0022-3956(88)90035-0

Lyon, N., Mejsholm, B., & Lyon, M. (1986). Stereotyped responding by schizophrenic outpatients: Cross-cultural confirmation of perseverative switching on a two-choice task. *Journal of Psychiatric Research, 20,* 137–150. http://dx.doi.org/10.1016/0022-3956(86)90013-0

Mansfield, E. L., Karayanidis, F., Jamadar, S., Heathcote, A., & Forstmann, B. U. (2011). Adjustments of response threshold during task switching: A model-based functional magnetic resonance imaging study. *The Journal of Neuroscience, 31,* 14688–14692. http://dx.doi.org/10.1523/JNEUROSCI.2390-11.2011

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorials in Quantitative Methods for Psychology, 4,* 61–64.

Mowrer, O. H. (1941). Preparatory set (expectancy): Further evidence of its "central" locus. *Journal of Experimental Psychology, 28,* 116–133. http://dx.doi.org/10.1037/h0061076

Nuechterlein, K. H. (1977). Reaction time and attention in schizophrenia: A critical evaluation of the data and theories. *Schizophrenia Bulletin, 3,* 373–428. http://dx.doi.org/10.1093/schbul/3.3.373

Nuechterlein, K. H., & Dawson, M. E. (1984). Information processing and attentional functioning in the developmental course of schizophrenic disorders. *Schizophrenia Bulletin, 10,* 160–203. http://dx.doi.org/10.1093/schbul/10.2.160

Nuechterlein, K. H., Subotnik, K. L., Green, M. F., Ventura, J., Asarnow, R. F., Gitlin, M. J., . . . Mintz, J. (2011). Neurocognitive predictors of work outcome in recent-onset schizophrenia. *Schizophrenia Bulletin, 37,* S33–S40. http://dx.doi.org/10.1093/schbul/sbr084

Peralta, V., & Cuesta, M. J. (1994). Psychometric properties of the Positive and Negative Syndrome Scale (PANSS) in schizophrenia. *Psychiatry Research, 53,* 31–40. http://dx.doi.org/10.1016/0165-1781(94)90093-0

Rae, B., Heathcote, A., Donkin, C., Averell, L., & Brown, S. (2014). The hare and the tortoise: Emphasizing speed can change the evidence used to make decisions. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40,* 1226–1243. http://dx.doi.org/10.1037/a0036801

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation, 20,* 873–922. http://dx.doi.org/10.1162/neco.2008.12-06-420

Ratcliff, R., & Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychological Review, 111,* 333–367. http://dx.doi.org/10.1037/0033-295X.111.2.333

Ratcliff, R., Spieler, D., & McKoon, G. (2000). Explicitly modeling the effects of aging on response time. *Psychonomic Bulletin & Review, 7,* 1–25. http://dx.doi.org/10.3758/BF03210723

Ratcliff, R., & Van Dongen, H. P. A. (2011). Diffusion model for one-choice reaction-time tasks and the cognitive effects of sleep deprivation. *Proceedings of the National Academy of Sciences of the United States of America, 108,* 11285–11290. http://dx.doi.org/10.1073/pnas.1100483108

R Core Team. (2014). R: A language and environment for statistical computing [Computer software]. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

Rentrop, M., Rodewald, K., Roth, A., Simon, J., Walther, S., Fiedler, P., . . . Kaiser, S. (2010). Intra-individual variability in high-functioning patients with schizophrenia. *Psychiatry Research, 178,* 27–32. http://dx.doi.org/10.1016/j.psychres.2010.04.009

Salthouse, T. A., & Somberg, B. L. (1982). Isolating the age deficit in speeded performance. *Journal of Gerontology, 37,* 59–63. http://dx.doi.org/10.1093/geronj/37.1.59

Schatz, J. (1998). Cognitive processing efficiency in schizophrenia: Generalized vs domain specific deficits. *Schizophrenia Research, 30,* 41–49. http://dx.doi.org/10.1016/S0920-9964(97)00125-4

Schmiedek, F., Oberauer, K., Wilhelm, O., Süβ, H.-M., & Wittmann, W. W. (2007). Individual differences in components of reaction time distributions and their relations to working memory and intelligence. *Journal of Experimental Psychology: General, 136,* 414–429. http://dx.doi.org/10.1037/0096-3445.136.3.414

Schneider, F. C., Royer, A., Grosselin, A., Pellet, J., Barral, F. G., Laurent, B., . . . Lang, F. (2011). Modulation of the default mode network is task-dependant in chronic schizophrenia patients. *Schizophrenia Research, 125,* 110–117. http://dx.doi.org/10.1016/j.schres.2010.11.013

Silverstein, S. M. (2008). Measuring specific, rather than generalized, cognitive deficits and maximizing between-group effect size in studies of cognition and cognitive change. *Schizophrenia Bulletin, 34,* 645–655. http://dx.doi.org/10.1093/schbul/sbn032

Smith, P. L., & Ratcliff, R. (2009). An integrated theory of attention and decision making in visual signal detection. *Psychological Review, 116,* 283–317. http://dx.doi.org/10.1037/a0015156

Smyrnis, N., Karantinos, T., Malogiannis, I., Theleritis, C., Mantas, A., Stefanis, N. C., . . . Evdokimidis, I. (2009). Larger variability of saccadic RT in schizophrenia patients. *Psychiatry Research, 168,* 129–136. http://dx.doi.org/10.1016/j.psychres.2008.04.015

Soetens, E. (1998). Localizing sequential effects in serial choice reaction time with the information reduction procedure. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 547–568. http://dx.doi.org/10.1037/0096-1523.24.2.547

Soetens, E., Boer, L. C., & Hueting, J. E. (1985). Expectancy or automatic facilitation? Separating sequential effects in two-choice reaction time. *Journal of Experimental Psychology: Human Perception and Performance, 11,* 598–616. http://dx.doi.org/10.1037/0096-1523.11.5.598

Sutton, S., Hakerem, G., Zubin, J., & Portnoy, M. (1961). The effect of shift of sensory modality on serial reaction-time: A comparison of schizophrenics and normals. *The American Journal of Psychology, 74,* 224–232. http://dx.doi.org/10.2307/1419407

van Ravenzwaaij, D., Dutilh, G., & Wagenmakers, E.-J. (2012). A diffusion model decomposition of the effects of alcohol on perceptual decision making. *Psychopharmacology, 219,* 1017–1025. http://dx.doi.org/10.1007/s00213-011-2435-9

Wagenmakers, E.-J., & Farrell, S. (2004). AIC model selection using Akaike weights. *Psychonomic Bulletin & Review, 11,* 192–196. http://dx.doi.org/10.3758/BF03206482

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale—Third Edition.* San Antonio, TX: Psychological Corporation.

Wechsler, D. (2001). *Wechsler Test of Adult Reading.* San Antonio, TX: Psychological Corporation.

White, C., Ratcliff, R., Vasey, M., & McKoon, G. (2009). Dysphoria and memory for emotional material: A diffusion-model analysis. *Cognition and Emotion, 23,* 181–205. http://dx.doi.org/10.1080/02699930801976770

White, C. N., Ratcliff, R., Vasey, M. W., & McKoon, G. (2010a). Anxiety enhances threat processing without competition among multiple inputs: A diffusion model analysis. *Emotion, 10,* 662–677. http://dx.doi.org/10.1037/a0019474

White, C. N., Ratcliff, R., Vasey, M. W., & McKoon, G. (2010b). Using diffusion models to understand clinical disorders. *Journal of Mathematical Psychology, 54,* 39–52. http://dx.doi.org/10.1016/j.jmp.2010.01.004

Yogev, H., Hadar, U., Gutman, Y., & Sirota, P. (2003). Perseveration and over-switching in schizophrenia. *Schizophrenia Research, 61,* 315–321. http://dx.doi.org/10.1016/S0920-9964(02)00323-7

Zahn, T. P., Jacobsen, L. K., Gordon, C. T., McKenna, K., Frazier, J. A., & Rapoport, J. L. (1998). Attention deficits in childhood-onset schizophrenia: Reaction time studies. *Journal of Abnormal Psychology, 107,* 97–108. http://dx.doi.org/10.1037/0021-843X.107.1.97

Zubin, J. (1975). Problem of attention in schizophrenia. In M. Kietzman, S. Sutton, & J. Zubin (Eds.), *Experimental approaches to psychopathology* (pp. 139–165). New York, NY: Academic Press.

Members of Underrepresented Groups:
Reviewers for Journal Manuscripts Wanted

If you are interested in reviewing manuscripts for APA journals, the APA Publications and Communications Board would like to invite your participation. Manuscript reviewers are vital to the publications process. As a reviewer, you would gain valuable experience in publishing. The P&C Board is particularly interested in encouraging members of underrepresented groups to participate more in this process.

If you are interested in reviewing manuscripts, please write APA Journals at Reviewers@apa.org. Please note the following important points:

- To be selected as a reviewer, you must have published articles in peer-reviewed journals. The experience of publishing provides a reviewer with the basis for preparing a thorough, objective review.

- To be selected, it is critical to be a regular reader of the five to six empirical journals that are most central to the area or journal for which you would like to review. Current knowledge of recently published research provides a reviewer with the knowledge base to evaluate a new submission within the context of existing research.

- To select the appropriate reviewers for each manuscript, the editor needs detailed information. Please include with your letter your vita. In the letter, please identify which APA journal(s) you are interested in, and describe your area of expertise. Be as specific as possible. For example, "social psychology" is not sufficient—you would need to specify "social cognition" or "attitude change" as well.

- Reviewing a manuscript takes time (1–4 hours per manuscript reviewed). If you are selected to review a manuscript, be prepared to invest the necessary time to evaluate the manuscript thoroughly.

APA now has an online video course that provides guidance in reviewing manuscripts. To learn more about the course and to access the video, visit http://www.apa.org/pubs/authors/review-manuscript-ce-video.aspx.