Journal of Abnormal Psychology
2007, Vol. 116, No. 2, 290–305

0021-843X/07/$12.00 DOI: 10.1037/0021-843X.116.2.290

Cognitive Processing of Facial Affect: Connectionist Model of Deviations in Schizophrenia

Jeffrey R. Carter
Mme Vanier Children's Services and
University of Western Ontario

Richard W. J. Neufeld
University of Western Ontario

Connectionist-model simulations of competing hypotheses of cognition in schizophrenia were constructed and tested. Emphasis was placed on judgment of affect, a prominent area of disturbance in this disorder with potential implications for social impairment. Participants with paranoid or nonparanoid schizophrenia and control participants provided judgments of affect as expressed in photographic faces. Schizophrenia groups were less accurate than control groups, and the paranoid group had greater latencies than did other groups. Model predictions simultaneously addressed judgment content and latencies for each trial. Results provide a connectionist extension of an account of deficits in schizophrenia that originated at the computational (stochastic modeling) level of analysis. This account postulates extra stages of item encoding but no reduction in formally defined processing capacity. It also provides for abnormalities in both judgment patterns and duration and is consistent with biological accounts of schizophrenia deficits. The substantive findings are supported by strategic innovations in the construction and testing of connectionist models.

Keywords: schizophrenia, encoding deficit, connectionist modeling, facial affect, cognitive performance

Supplemental material: http:dx.doi.org/10.1037/0021-843x.116.2.290.supp

Several lines of evidence suggest that people with schizophrenia have difficulty processing information related to facial affect (Corrigan & Green, 1993). Researchers have long known that these individuals have difficulty interpreting complex, meaningful stimuli (e.g., Gessler, Cutting, Frith, & Weinman, 1989; Maher, 1966; Neufeld, 1976) and that they have particular difficulty processing affective information (e.g., Deering, 1963). Reduced sensitivity to nonverbal social cues, such as those indicating others' emotions, may contribute to inappropriate responses and impaired social functioning (Walker, McGuire, & Bettes, 1984). Previous cognitive research (Carter & Neufeld, 1999; Neufeld, Vollick, & Highgate, 1993) mathematically modeled a cognitive process, and certain model parameters, that tenably account for certain deficits in schizophrenia. The current investigation builds on those findings by demonstrating through a different modeling framework that deficits in processing facial affect may also be caused by the same mechanism: Additional processing while encoding information.

Jeffrey R. Carter, Mme Vanier Children's Services, London, Ontario, Canada, and Department of Psychology and Department of Psychiatry, University of Western Ontario, London, Ontario, Canada; Richard W. J. Neufeld, Department of Psychology and Department of Psychiatry, University of Western Ontario.

This article is based on a dissertation by Jeffrey R. Carter that was supervised by Richard W. J. Neufeld and submitted in partial fulfillment of the requirements for the doctoral degree in clinical psychology at the University of Western Ontario. This research was supported by a doctoral fellowship and an operating grant from the Social Sciences and Humanities Research Council of Canada, an Ontario graduate scholarship, a University of Western Ontario graduate research fellowship, graduate tuition and special university scholarship, an Ontario Mental Health Foundation senior research fellowship, and an operating grant from the Canadian Institutes of Health Research.

We thank Ken McRae for his helpful advice throughout the conduct of this research and John Kruschke, Elizabeth Hayden, and David Dozois for their observations on a previous version of this article.

Correspondence concerning this article should be addressed to Jeffrey R. Carter, Mme Vanier Children's Services, 871 Trafalgar Street, London, Ontario N5Z 1E6, Canada. E-mail: jcarter@vanier.com

Paranoid and Nonparanoid Schizophrenia

The present project applied separate modeling analyses to participants with paranoid and nonparanoid types of schizophrenia. This distinction has proven meaningful according to multivariate and other empirical studies of symptomatology (reviewed in Neufeld & Williamson, 1996; Nicholson & Neufeld, 1993) and has been identified with differing patterns of cognitive performance (e.g., Neufeld, 1991; Neufeld & Williamson, 1996). Izard (1959) appears to be the first to suggest that paranoid schizophrenia is associated with difficulty processing information from faces. Some investigators (e.g., Ennis & Whelton, 1994; Penn & Spaulding, 1997; Silva, Leong, Wine, & Saab, 1992) have suggested possible links between paranoid schizophrenia and deficits in social skills, facial identity recognition, and facial affect recognition. Other researchers (e.g., Borod et al., 1989; Mueser et al., 1996; Schneider, Gur, Gur, & Shtasel, 1995) have suggested that deficits in interpreting facial expressions are related to negative schizophrenia symptoms.

Accuracy and Latency Effects

Research on facial affect perception in schizophrenia has been extensive. Most, but not all, of the results indicate deficit in

accuracy (Flack, Cavallaro, Laird, & Miller, 1997; Levy, Orr, & Rosenzweig, 1960; Pilowsky & Bassett, 1980; and especially Davis & Gibson, 2000, and LaRusso, 1978). When latency has been measured, people with schizophrenia have tended to take longer than control participants to recognize facial affect (Burch, 1995; Frith, Stevens, Johnstone, Owens, & Crow, 1983; Mandal & Rai, 1987).

Possible Causes of the Processing Deficit

Research has uncovered a vast array of cognitive deficits and biological anomalies associated with schizophrenia, and one can safely assume that several brain areas are involved in any given complex task. The following examples focus specifically on the processing of facial affect information. Gur et al. (2002) found decreased activation of the limbic regions among people with schizophrenia on an emotional valence discrimination task. Hempel, Hempel, Schonknecht, Stippich, and Schroder (2003) found decreased activation of the amygdala and hippocampus among people with partially remitted schizophrenia on a facial affect labeling task, but Kosaka et al. (2002) found exaggerated amygdala responses among people medicated for schizophrenia on an emotional intensity judgment task. One of the challenges in this field is to bring order to such findings. The current study attempts to bring greater understanding of a symptom (poor social functioning) by linking a related cognitive deficit (decreased accuracy and increased latency when processing facial affect information) to known cognitive and neurobiological anomalies.

Several hypotheses at different levels of explanation have been investigated regarding cognitive deficits in schizophrenia. The following sections address each in turn. The first two are framed primarily in cognitive terms: Additional processing steps and reduced processing speed. The remaining hypotheses tend more to the biological: neurotransmitter dysfunction, neural atrophy, deficits in visual processing systems, and deviant cortical pruning.

Additional Processing Steps (Subprocesses)

One cognitive deficit in schizophrenia identified by stochastic models of cognitive-task performance is the presence of additional operations during encoding. Encoding is the process of transforming a presented stimulus into a format suitable to the task at hand (Neufeld, Carter, Boksman, Jetté, & Vollick, 2002; Neufeld et al., 2007, 1993). An in-depth exposition of subprocesses in stochastic modeling, including their involvement in serial- and parallel-processing architectures, is given by Townsend (1984; see also Townsend & Ashby, 1983, chapter 4; for the present applications, see Neufeld et al., 1993). Examples of encoding processes include extracting physical features, identifying the name of a presented item, or locating a stimulus on its component dimensions (e.g., "pleasure" and "arousal" for faces). Cognitive–behavioral correlates of the additional steps in encoding may include repetition of previously completed operations, preliminary orienting activity, and overinclusiveness (Neufeld et al., 2007). *Rate*, or capacity for dispatching constituent subprocesses (Townsend & Ashby, 1978), thus remains intact but is implemented less economically. A metaphor that might describe this deficit is that of one train taking longer to pass a crossing than another because it has more cars—even though both are traveling at the same speed.

Biological substrates of additional encoding have been indicated according to high-field-strength (4.0 Tesla) functional magnetic resonance imaging studies, as follows. Task trials rich in encoding demands (endowing consonants with lexical properties) have been accompanied by diminished right anterior cingulate cortex (ACC) activation among participants with first-episode, never-treated schizophrenia (Boksman et al., 2005). Moreover, effective connectivity (essentially time-series covariance distinguishing the above encoding activity; Friston et al., 1997) between the right ACC and left inferior temporal region, apparent among control participants, was absent among the schizophrenia participants. Instead, diffuse connectivity of the ACC to multiple brain regions (of note, prefrontal and parietal regions) was observed. Similarly, more diffuse ACC connectivity occurred among schizophrenia participants performing a memory-search task (ascertaining the presence of a target item among a memorized item set) during trial epochs estimated by stochastic performance models (Neufeld et al, 2002; 2007) to accentuate target-item encoding (Boksman et al., 2007).

Reduced Processing Speed

An alternative explanation for increased latencies would be a decreased processing speed. The train metaphor for this hypothesis would be that one train takes longer than another because it is moving more slowly. Previous stochastic modeling (Carter & Neufeld, 1999; Neufeld et al., 1993) found evidence against this alternative hypothesis as an explanation for cognitive deficits in schizophrenia. Further, previous research (Chechile, 1987) has suggested that interference from additional processing accounts for deficits in content better than does the mere passage of time. From this perspective, reduced processing speed seems a less likely contender than increased processing. Because processing speed has been identified with processing capacity in the stochastic modeling literature, and processing capacity has been impugned in schizophrenia cognition (Neufeld et al., 1993), evaluating processing speed within the present modeling framework appears warranted. The biological substrate is even less clear for the reduced processing speed hypothesis than for the preferred cognitive explanation of increased processing.

Neurotransmitter Dysfunction

Turning from the cognitive explanations to the more biological, one sees that dopamine abnormalities have often been associated with schizophrenia (e.g., Creese, Burt, & Snyder, 1976; Kapur et al., 1998; Remington, Kapur, & Kipursky, 1998; Seeman, Lee, Chau-Wong, & Wong, 1976). The most common association is that schizophrenia symptoms are related to excess dopamine-receptor activity (Bartha et al., 1997; Kapur et al., 1998; Seeman & Lee, 1975). Cohen and Servan-Schreiber (1992) offered a different explanation when they simulated the effects of dopamine anomalies with a connectionist model. They theorized that dopamine modulates the response properties of the postsynaptic cells, enhancing their sensitivity specifically to meaningful elements of the stimuli (the signal-to-noise ratio). Their model accounted for deficits among people with schizophrenia on three different cognitive tasks: Stroop, continuous performance (identical pairs), and lexical disambiguation.

Neural Atrophy

Another possibility is that schizophrenia deficits are related to reduced brain mass or function (Crow, 1990). That is, entire groups of neurons at specific physical regions may be lost or rendered ineffective (Benes, Davidson, & Bird, 1986; Benes, McSparren, Bird, SanGiovanni, & Vincent, 1991). Potential enlargement of the ventricles in schizophrenia, for example, has been associated with a reduction in the temporal lobes (e.g., Barta, Pearlson, Powers, Richards, & Tune, 1990). The information lodged in those regions would be eliminated.

The temporal lobes may be of specific interest to the current investigation because individual features of facial expressions may be represented there (Deakin, 1994). Several sources indicate that the superior temporal sulcus of the macaque monkey contains cells that are involved in face perception (Heywood & Cowey, 1992; Perrett, Hietanen, Oram, & Benson, 1992; Perrett & Mistlin, 1990). These cells appear insensitive to other stimuli (Hasselmo, Rolls, & Baylis, 1989; Perrett et al., 1988; Rolls, 1992; Tsao, Freiwald, Tootell, & Livingstone, 2006). Different cells within the superior temporal sulcus are selectively responsive to different features, feature subsets, and feature configurations, and they appear to be sensitive to both macaque and human faces.

Early Processing Deficits

An additional proposal for impaired performance on tasks with visual stimuli involves deficits in processing the raw physical features of the stimulus. Mechanisms that control early processing seem to be disrupted in schizophrenia (Beech, McManus, Baylis, Tipper, & Agar, 1991; Frith, 1979). Further, aberrant visual scan paths associated with certain symptoms may reflect early processing deficits (Phillips & David, 1997; but cf. Knight, 1993; Knight & Silverstein, 2001). According to this hypothesis, the processing system itself remains intact, but errors and delays occur because the material being processed has been garbled.

Deviant Cortical Pruning

Another prominent hypothesis regarding deficits in schizophrenia is the density of synapses, an important neurodevelopmental variable (Feinberg, 1990). Feinberg (1982/1983, p. 331) stated simply, "Too many, too few, or the wrong synapses are eliminated" (p. 331). The most common hypothesis is that people who have schizophrenia may have reduced synaptic density, possibly indicating either excessive pruning or inadequate growth of synapses. For example, people who mature later tend to have both lower synaptic density and higher incidence of schizophrenia (Saugstad, 1989). More specifically, reduced thickness in cerebral cortex among people with schizophrenia may be caused by decreased dendritic material (Selemon, Rajkoska, & Goldman-Rakic, 1995). On the basis of connectionist simulations, Hoffman and his associates (Hoffman & Dobscha, 1989, 1990; Hoffman & McGlashan, 1993, 1994, 1998, 2006; Hoffman et al., 1995) found that excessive cortical pruning could be associated with schizophrenia symptoms, including hallucinations and delusions.

The diffuse damage associated with excessive pruning could interfere with the spread of information needed to complete cognitive tasks (Devlin, Gonnerman, Andersen, & Seidenberg, 1998). Theoretically, information losses may be either random or systematic. From a neural Darwinist perspective (e.g., Edelman, 1986), the elimination of less important synaptic connections is part of normal neurodevelopment. According to this premise, some of the less consequential, but still necessary, connections are lost in schizophrenia because the elimination process is exaggerated. A variant would be that some of the more important connections are lost in schizophrenia because the elimination process is distorted. Alternatively, if losses were merely related to a physiological disease process, the importance of the lost connections would be irrelevant.

Summary

A deficit related to the encoding of information into a format suitable to the task at hand—specifically, adding extra processing steps—has been found to be related to other impairments in schizophrenia. The current study investigated whether the encoding deficit applies to deficits in facial affect recognition. Deficits in facial affect recognition have been well documented in schizophrenia, but their sources are unclear. Few studies have examined differences between people with different subtypes of schizophrenia, and fewer yet have considered response latency. The current study established a model of the performance of people without schizophrenia and then pit against each other several different hypotheses regarding possible sources of schizophrenia deficits. Possible mechanisms include not only the preferred hypothesis of additional processing but also reduced processing speed, neurotransmitter dysfunction, structural aberrations, deficits in visual processing, and reduced synaptic density.

Method

Participants

Schizophrenia outpatients from the London Psychiatric Hospital (now Regional Mental Health Care–London), in London, Ontario, Canada, were invited to enter the study, following approval from their primary clinical caregiver. Sixty-one individuals participated (30 with paranoid and 31 with other forms of schizophrenia). Clinical diagnosis according to the International Classification of Disease was corroborated through the Maine Scale for paranoid and nonparanoid schizophrenia (Magaro, Abrams, & Cantrell, 1981), using procedures of Lubow, Weiner, Schlossberg, and Baruch (1987; as in Carter & Neufeld, 1999). Control participants included 30 volunteers solicited from Human Resources Development Canada (HRDC), London, Ontario, Canada, and 31 psychology undergraduates from the University of Western Ontario. (See George & Neufeld, 1987a, 1987b, for examples of this combination of control participants.) Patients and HRDC volunteers were each paid $15 for participating, and each student received one introductory-course research credit. Age of participants was between 18 and 60 years. Each had a minimum education level of 6 years, as well as normal visual acuity (corrected or uncorrected). No patient had received electroconvulsive shock therapy, had been cumulatively hospitalized more than 36 months, or had a currently active diagnosis involving substance abuse. Exclusion criteria for control participants included any indication of brain damage or

dysfunction due to injury, epilepsy, or substance abuse, or any score greater than zero on either of the Maine Scale subscales.

Stimuli

Facial affect was presented in the form of photographs of human faces that were scanned into a bitmap format. A set of 20 images represented four distinct prototypical expressions—*happy, surprised, sad,* and *disgusted* (Ekman & Friesen, 1976). An additional six photographs (depicting *anger* and *fear*) served for practice trials. All bitmapped images had similar size, luminosity, and contrast.

Procedure and Apparatus

Participants read computer-presented instructions and completed six practice trials. Computer-presented feedback regarding accuracy accompanied practice trials. Any queries were answered verbally. Judgments were then obtained for the 190 unique stimulus pairs. Participants proceeded at their own pace (as in Carter & Neufeld, 1999). Order of presentation was randomized individually for each participant.

The 190 unique pairs of the 20 images were presented during the judgment trials, producing 10 trials each for each of the four same-emotion pairs (e.g., *happy–happy*) and 25 trials for each of the six unique combinations (e.g., *happy–sad*). Pairs of stimuli were presented on a Packard Bell (Sacramento, CA) SX16 computer with a 13-inch [33-cm] color monitor. Each image was 11 × 15.5 cm, and each pair of images was presented side by side, separated by 1 cm. The visual angle accommodated visual search strategies of both patients and control participants (Cegalis, Leen, & Solomon, 1977). Responses were collected with a three-button box: Participants pressed and held one button (to initiate the trial) and selected one of the other two to indicate their judgment choice (same or different). Processing latencies and movement duration were measured in milliseconds with a London Research and Development (London, Ontario, Canada) 8254 Timer Board. Processing time was taken as the interval between the stimulus presentation on the screen and the release of the cue button (Pachella, 1974), and movement time was taken as the latency between the release of the cue button and the pressing of the choice button.

Results

Results were generally consistent with those of previous research on facial affect recognition in schizophrenia. The major finding (elaborated upon below) was that the paranoid schizophrenia (PS) group had greater latencies than the other three groups, and the nonparanoid schizophrenia (NPS) and PS groups were less accurate than the control groups.

Participants

Table 1 provides a summary of participant characteristics. Group differences in socioeconomic status and age were analyzed further. Within-group correlations indicated that any relationships between demographic variables and the variables of interest were either artifacts of group differences or unique to specific groups. After the Bonferroni correction for multiple comparisons, the only statistically significant correlations were for age and overall latency for the community control ($r = .59$, $p < .002$) and NPS ($r = .51$, $p < .004$) groups. These two groups are not statistically significantly different in age. Socioeconomic differences existed in that the community control group had higher status than did other

Table 1
Sample Characteristics: Frequencies or Means (±SD)

Attribute	Group				Statistical test
	Paranoid	Nonparanoid	Community	Student	
Gender					$\chi^2(3) = 4.95$, *ns*
Men	26	24	20	20	
Women	4	7	10	11	
Age (years)[a]	45.1 (8.6)	42.7 (7.8)	40.23 (9.0)	19.6 (1.2)	$F(3, 118) = 77.9$, $p < .001$
Socioeconomic status[b]	52.9 (14.7)	55.45 (11.8)	39.8 (10.8)	49.7 (8.6)	$F(3, 118) = 10.42$, $p < .001$
Handedness[c]	0.66 (0.54)	0.51 (0.52)	0.66 (0.38)	0.41 (0.53)	$F(3, 118) = 1.79$, *ns*
Receptive vocabulary[d]	96.2 (19.2)	87.6 (19.1)	110.8 (16.9)	99.2 (13.5)	$F(3, 118) = 9.36$, $p < .001$
Cumulative hospitalization (days)	410.2 (329.7)	224.5 (254.6)			$F(1, 59) = 6.09$, $p < .05$
Antipsychotic medication (mg Chlorpromazine)[e]	316.6 (273.4)	330.4 (330.2)			$F(1, 59) < 1$, *ns*
Anti-Parkinsonian medication (mg Benztropine)[e]	1.0 (1.3)	1.7 (1.6)			$F(1, 59) = 3.17$, *ns*
Maine Scale score[f]					$F(1, 59) = 39.03$, $p < .001$
Paranoid	9.0 (1.9)	6.5 (1.3)			$F(1, 59) = 22.43$, $p < .001$
Nonparanoid	6.8 (1.1)	8.6 (1.8)			
Total	15.8 (2.7)	15.1 (2.9)			$F(1, 59) < 1$, *ns*

[a] Post hoc analysis (Scheffé procedure) revealed the student group to be younger than the other groups. Excluding this group, $F(2, 88) = 2.48$, *ns.*
[b] Hollingshead (1957) Two-Factor Index of Social Position; post hoc analysis revealed the community control group to be of greater status than the other groups. Excluding this group, $F(2, 89) = 1.78$, *ns.*
[c] Edinburgh Handedness Index (Oldfield, 1971).
[d] Peabody Picture Vocabulary Test—Revised (Dunn & Dunn, 1981); post hoc analysis revealed the community (Human Resources Development Canada) control group to have greater receptive vocabulary than did the patient groups but the student group to not differ significantly from any other group. Excluding the community control group, $F(2, 88) = 2.48$, *ns.*
[e] Davis (1976); Gillis (1998); Wyatt & Torgow (1976).
[f] Magaro et al. (1981).

groups.[1] Socioeconomic status was not statistically significantly correlated with the variables of interest (i.e., accuracy or latency) for any group. Most important, the two patient groups were comparable on all demographic variables, yet they exhibited significant differences in the variables of interest.

Turning to patient-specific variables, neither medication nor cumulative hospitalization was statistically significantly related to the variables of interest for either patient group (Spohn & Strauss, 1989). Although medication effects cannot be ruled out categorically, the two patient groups were comparable in terms of antipsychotic and anticholinergic dosage equivalencies.

Validity of Trials

A total of 20,676 individual trials remained for analysis after the removal of invalid responses (i.e., postresponse processing and disattention; Wenger & Townsend, 2000). *Postresponse processing* was deemed to have occurred when the time between stimulus presentation and the release of the cue button was less than the time between the release of the cue button and the depression of the judgment button. *Disattention* was identified as trials in which the time between stimulus presentation and finger lift was more than 10 s.

Analysis

The following three fixed-effect factors were included in a preliminary analysis of latency: groups (four levels), stimulus-defined same–different expressions (two levels), and correct–incorrect (conformity of same–different responses to the stimulus-defined expressions; two levels). Random-effects factors involved participants nested within groups (n = harmonic mean of 30.49) and expressions within stimulus categories of same–different (the selected combinations of expressions were regarded as a sample of all possible combinations of all possible expressions). Judgment content was quantified as the proportion of trials on which the individual's response agreed with the stimulus designation. This analysis was duplicated for latency, except that the correct–incorrect factor was eliminated. Throughout, Quasi-F ratios were computed when necessary (Myers & Well, 1991), and Greenhouse–Geisser epsilon was used to adjust degrees of freedom where within-participant tests violated assumptions of sphericity.

With respect to judgment latency, results indicated that the two control groups formed one subset, separated from the PS group (Scheffé-based procedures; Myers & Well, 1991). Means (in milliseconds) for student control, HRDC, NPS, and PS groups were 1,748 (SD = 647), 2,243 (SD = 805), 2,356 (SD = 778), and 3,021 (SD = 936), respectively. With regard to accuracy, the two control groups again formed one subset, and the PS and NPS groups formed another. Proportions correct for the above respective groups were 0.92 (SD = 0.04), 0.91 (SD = 0.06), 0.83 (SD = 0.11), and 0.84 (SD = 0.09).

Aggregate Data

For modeling purposes, data needed to be assembled into aggregates representing "homogeneous participants" (Townsend, 1984). The results just described suggested that the control groups could be combined, but the PS and NPS groups needed to remain separate. Before combining these groups, however, we conducted various tests (e.g., on group differences, within-group correlations between extraneous and dependent variables, and the distributions of accuracy scores and of latencies) to rule out statistical artifacts (see Carter, Neufeld, & Benn, 1998; Neufeld & Gardner, 1990; Neufeld & McCarty, 1994). These tests identified no significant threats to model integrity from sample characteristics or data aggregation. For example, Kolmogorov–Smirnov tests were applied to the full array of values within each of the pairs of expressions for each group, separately for correct and incorrect responses. None of the tests on correct responses were significant, and only three approached significance ($p < .10$). Not one test on incorrect responses even approached significance. A similar analysis of the accuracy data also suggested that aggregation was appropriate. After data from human participants were assembled into four aggregates (two controls each containing a mixture of data from HRDC and student participants, and one representing each schizophrenia group), computer simulations could begin.

Simulation

Modeling Framework

Stochastic models have previously been used to describe cognitive deficits in schizophrenia (e.g., Carter & Neufeld, 1999; Neufeld et al., 2002, 2007). *Stochastic models* (see Carter, Neufeld, & Benn, 1998; Townsend & Ashby, 1983) identify essential features of a processing system that are then reflected in performance properties such as speed and accuracy. The earmark of stochastic modeling is its ability to accommodate uncertainty (Busemeyer & Townsend, 1993; Doob, 1953). Stochastic models allow for random events. They provide predictions regarding the distributions of responses (e.g., their means, standard deviations, and shapes) and the effects of specific parameters. By modeling variables such as the probability of an event within a certain time period, stochastic models provide insight regarding processes. Clear, quantitative decision rules regarding the acceptability of a model are among the strengths of stochastic models (e.g., Wagenmakers & Waldorp, 2006).

The current investigation features a connectionist model of cognitive processing. *Connectionist* (i.e., parallel distributed processing) *models* consist of mathematical representations of networks of simple processors, assumed to be functionally similar to neurons in the brain. Computer simulations of the workings of connectionist networks constitute a potentially productive level of quantitative modeling when interest lies in linking neurophysiological mechanisms to cognitive dysfunction related to symptoms (Cohen & Servan-Schreiber, 1992; Hoffman & Dobscha, 1989; Siegle, Steinhauer, Thase, Stenger, & Carter, 2002; and especially Spitzer, 2000). The links between specific neurophysiological mechanisms and network analogues tend to have high face validity, such as the elimination of weights corresponding to the elimination of synapses, and the elimination of units corresponding to the elimination of neurons (see, e.g., Cohen & Servan-Schreiber,

[1] We thank an anonymous reviewer for pointing out that the socioeconomic status of the students may not be comparable to that of patients in that the students' situation is presumably temporary.

1992; Devlin et al., 1998; Farah & McClelland, 1991; Hinton & Shallice, 1991; Hoffman, 1992; Hoffman et al., 1995). In contrast to stochastic models, which make predictions regarding probabilities, connectionist models focus on the activations of certain processing units. Both types of models may include conceptually similar parameters, such as number of subprocesses or rate of processing, but fundamental differences in two frameworks prevent a formal, direct comparison.

The basic strategy of the current investigation was to create a connectionist network that could simulate the typical performance of nonpatients and then modify that model by adjusting one or more parameters so that the same network could simulate the performance of patients. This approach is typical for stochastic modeling (e.g., Busemeyer & Stout, 2002; Yechiam, Veinott, Busemeyer, & Stout, 2007). It puts aside several aspects of the network that are often of critical importance to connectionist modelers, such as different learning rules or model structures. Instead, the current investigation focused on directly comparing several different hypotheses, each represented by specific parameter manipulations. Such direct competition between hypotheses has been relatively uncommon in connectionist modeling (Siegle, 1998; but cf. Hoffman et al., 1995).

One of the basic assumptions of the current study is that the same general model accounts for anyone completing the task, with individual differences reflected in varying parameter values within the model. This approach is consistent with the methodology that Batchelder (1998) and others (Riefer, Knapp, Batchelder, Bamber, & Manifold, 2002) refer to as *cognitive psychometrics*, in which the goal is to pinpoint the exact nature of cognitive deficits by interpreting differences in parameter values. For example, Wallsten, Pleskac, and Lejuez (2005) modeled risk-taking behavior with this approach and found that adjusting the basic structure of a model was no better than adjusting model parameters in terms of the correlations of model parameters with clinical variables. One objective of the current study was to identify which parameters to manipulate and to what degree so that the model accounted for not only the responses of the control participants, but also the responses of each of the schizophrenia groups. Note that nested models (i.e., in which one model is a special case of a more general model) allow for statistically disciplined comparisons of model fit to the empirical data, whereas models with different structures cannot be compared in this way.

Simulation Network

The framework for the network used in the current investigation was intentionally simple. More sophisticated frameworks were rejected to reduce the number of potential confounds to the findings of interest (i.e., the parameter values associated with specific deficits). As much as possible, the details of the network were determined by a small set of learning rules applied to the model, rather than being imposed by the modeler. The chosen framework consisted of three layers. The first layer was made up of input units, each representing some feature of facial affect expression. The second ("hidden") layer was made up of processing units. This layer received information from the input and output units and sent information forward to the third layer. The third layer consisted of a single output unit, which both received information from and sent information to the hidden layer. The feedback loop between the hidden and output layers creates a dynamic system allowing for an indeterminate length of processing. This feedback loop is consistent with other models that address latency on tasks involving competing responses (e.g., Yeung, Botvinick, & Cohen, 2004). The network continued calculating activations for each processing unit in successive iterations ("ticks") until all of the units settled into stable values.

The strategy for the simulation was to train the network with one subset of data ("training data") from the assembled control group until it provided an adequate fit to the remaining data from the assembled control group ("test data"), and then to adjust various network parameters until it provided an adequate fit to each of the schizophrenia groups. As a cross-validation procedure (Browne, 2000), this process was repeated on a second network of the same architecture with the data that had been test data used as training data, and vice versa. Further, the possibility that model improvement was an artifact of parameter manipulations was investigated by applying each manipulation to the control data ("mock testing").

Cost functions. The challenge of developing objective summaries of the accuracy of model predictions required the following methodological development. The guiding format was that of chi-square. Higher values of this statistic represent more error. Straightforward decision rules exist regarding how much error is acceptable. Because connectionist models do not provide estimates for variability in responding as traditional stochastic models do, certain modifications were required (see online appendix). The resulting statistic, provisional chi-square (p-χ^2), does not yield exact probability values but otherwise functions in the same manner as the usual chi-square statistic. The p-χ^2 statistic provides a strong test of the model in that it summarizes error for both accuracy and latency across all trials. These two components of responding are given equal weight in the p-χ^2 statistic.

Additional cost functions provided convergent evidence on the relative effectiveness of the model and its modifications. These functions included the sum of squared differences between model-predicted and observed response selections and latencies (*SS*) and pseudo-chi-square (pseudo-χ^2), a demonstrably useful parameter-estimating cost function resembling but not identical to chi-square (Neufeld & McCarty, 1994; Townsend, 1984; Townsend & Ashby, 1983, chapter 13). These two cost functions also both provide measures of the lack of fit at the level of individual trials for both accuracy and latency, but they do not lend themselves easily to adopting the decision rules associated with chi-square. Results from these analyses are reported only when the conclusions were not consistent with conclusions based on p-χ^2.

Architecture. The simulation comprised a 258-unit network, depicted in Figure 1. The 250 input units (two groups of 125, each unit representing a psychologically meaningful facial feature, and each group representing one face) sent activations to a set of seven hidden units. Hidden units sent activations to a single output unit, which sent its activation back to the hidden units. (See Roelfsema & van Ooyen, 2005, for an example of this architecture, and Botvinick & Plaut, 2006, for an example of a fully recurrent back-propagation network.)

Representations. Each of the photographic images was represented as a set of psychologically meaningful features. In a preliminary study, a group of 23 psychology undergraduate volunteers identified 125 specific features in the set of 20 images. Another

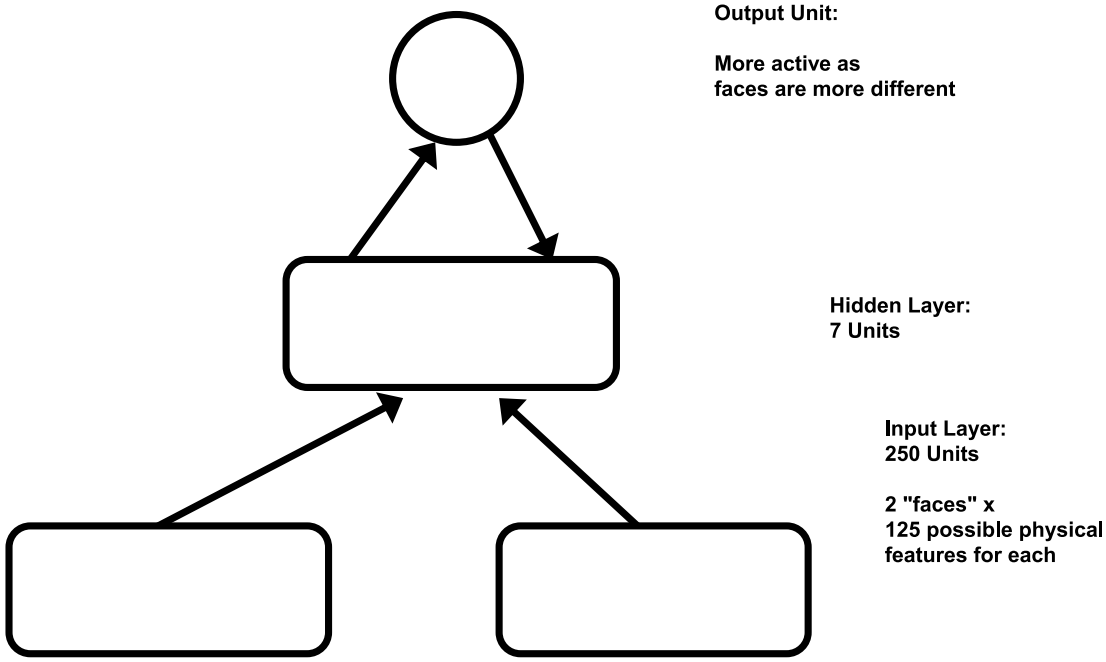

Figure 1. Back-propagation network for judgments of facial affect expressions.

group of 25 psychology undergraduate volunteers coded for the presence or absence of those features. See the study by Carter (2006, pp. 16–17, 29–40) for details.

The input units for the network were divided into two sets of 125. Each "face" was equally likely to be represented on the "left" (i.e., the first 125 units) or "right" (i.e., the second 125 units). Each input unit represented a specific feature. The activation was set according to the proportion of endorsements in the preliminary study of that feature for the image at hand. Interstimulus dissimilarity (represented in the model by the activation of the output unit) reflected the proportion of individuals in the experimental sample who responded "different" to the pair of photographic images. The activation of the output unit was the model's representation of response content.

One of the major challenges for this project was to represent response latencies. In this model, time to process a set of stimuli is the product of the number of processing steps and the time required to complete each step. The number of processing steps is treated as the number of successive iterations of unit-activation adjustment before settling, represented as K. The time (in real-time units) required for each processing step is summarized by a parameter, h (after Busemeyer & Townsend, 1993). Thus, the model's representation of the time required to complete processing is hK. The value of h was calculated with the following formula:

$$h = \frac{\sum_{i=1}^{N} (K_i RT_i)}{\sum_{i=1}^{N} K_i^2},$$

where N is the number of stimulus judgments, K_i is the number of activation-change iterations for the model to settle to the target activation for stimulus i, and RT_i is the real time required for human participants to process stimulus i. (See Carter, 2006, pp. 1–2, for the derivation of this formula.)

Training-rule and parameter estimation. The network was trained with the back-propagation-through-time algorithm (see online appendix). Each epoch involved the entire set of 190 unique pairs of faces, each pair presented once. The only parameter of theoretical interest for which a closed form solution was available was h. Other parameters were estimated with a direct-search algorithm. Many starting points within the range of possible values for that parameter were selected. From each starting point, the value of the parameter was adjusted by small amounts until the value of the cost function reached a local minimum. Although this procedure was computationally intense, it made a minimum of assumptions regarding the shape of the prediction-error surface. To avoid confounds with specific cost functions, we estimated parameters by using each of the cost functions that measured error at the level of the individual trial (i.e., p-χ^2, *SS*, and pseudo-χ^2).

Processing of a stimulus pattern. The activations of the input units were clamped at the values associated with the stimulus for that trial throughout all training and testing trials. During initial training, the hidden and output unit activations were set according to a random (uniform) distribution with a mean of 0.2 and a standard deviation of 0.03 at the beginning of each training trial. When p-χ^2 was introduced as the error criterion (see online appendix), these values were consistently set to 0.2 because setting them to random values would destabilize this statistic. Within each iteration, activations were updated sequentially (i.e., hidden units, then the output unit).

In this model, processing for a particular trial stopped when the activations for all of the individual processing units changed by less than some arbitrary amount from one tick to the next. This "just-noticeable difference" was represented by the parameter ι.

Table 2
Summary of Cognitive Deficits, Their Representations, and Results

Cognitive dysfunction	Neurological correlate	Connectionist analogue	Result	Qualitative inferences
Additional subprocesses	Diminished activation and diffuse neuroconnectivity of encoding-significant brain regions	Decrease ι	$\iota_{Controls} = 0.001$ $\iota_{PS} = 0.0002$ $\iota_{NPS} = 0.000452$	Stable parameter values with statistically significant improvements when using group-specific parameter values
Decreased processing speed	Unknown	Increase *h*	$h_{Controls} = 0.014$ $h_{PS} = 0.021$ $h_{NPS} = 0.0163$	Unstable parameter values and substantially inflated model predictions for latency variances for the paranoid schizophrenia group
Modified signal-to-noise ratio	Neurotransmitter dysfunction	Adjust *gain*	$gain_{Default} = 1.00$ $gain_{Controls} = 1.15$ $gain_{PS} = 1.32$ or 1.46 $gain_{NPS} = 1.28$ or 1.46	Unstable parameter values across network instances and lack-of-fit indices
Complete loss of specific information	Reduced number of (functional) neurons	Eliminate a proportion of units	Eliminating units did not consistently improve model fit	Results were unstable, ranging from 0.02 to 0.78 across instances of the network and lack-of-fit indices
Early processing deficits	Dysfunction in the visual system	Add noise to input units	Adding noise did not consistently improve model fit	Results were unstable across instances of the network and lack-of-fit indices, with a mean of 0.09 and standard deviation of 0 being fairly common
Diffuse information loss or distortion	Cortical pruning	Eliminate a proportion of weights	Eliminating weights failed to improve model fit	Alternative manipulations yielded unstable results

Note. PS = paranoid schizophrenia group; NPS = nonparanoid schizophrenia group.

When none of the processing units in the model had a change greater than ι, processing for that trial stopped. For practical purposes, especially during training, processing was also stopped at an arbitrary maximum of 250 iterations if the network failed to settle. The network was considered to have settled when the change in activation for all units from one iteration to the next was less than the value of the parameter representing a just-noticeable difference, ι.[2]

Training regimen. The initial weights were drawn from a random (uniform) distribution with a mean of 0 and a standard deviation of 0.17. Weights were updated at the end of each training epoch. Preliminary testing indicated that the model needed to achieve some level of accuracy before the variables associated with latency were introduced. To this end, the network was initially trained on training-sample data with the latency parameters set to default values so that the network functioned as a simple feed-forward network. (See McRae & Hetherington, 1993, for an example of a pretrained network.) This initial training continued to an arbitrary reduction in total back-propagation error (see online appendix). After the parameters associated with latency were added, training continued until the network achieved the modal (most likely) value of the p-χ^2 statistic (which occurs when p-$\chi^2_{df} = df - 2$). Note that other options, such as $R^2_{predictions,\ observations}$ (e.g., Zorzi, 2002) or back-propagation error, require some arbitrary cutoff.

Simulation Results

The simulation findings were robust. Similar results were found for different cost functions and on cross-validation. Differences between the two instances of the network were minor. Hence, findings from the second instance are reported only when they are not redundant to those of the first. Similarly, the results of mock testing (i.e., estimation of optimum values with the use of control data for parameters theoretically associated with schizophrenia dysfunction) are reported only when adjusting the parameters improves the fit from their default values. The additional cost functions *SS* and pseudo-χ^2 produced parameter values and associated changes in model fit that were essentially consistent with those of p-χ^2. They are reported only when they are informative.

The preferred modification of adjusting the parameter associated with additional processing steps was most consistently supported. Some other modifications, however, also yielded positive results. Table 2 provides a summary of the results.

Initial Model Results

The simulation effectively predicted control group performance. For the first instance of the network, p-χ^2_{379} was 378.73 by training epoch 1 623, and for the second instance of the network,

[2] An alternative processing rule that more closely resembled decision-field theory (Busemeyer & Townsend, 1993), in that processing stopped after the activation of the output unit crossed a certain threshold, failed to replicate the observed data even after several architectures were attempted and extensive training was conducted. With the alternative rule, the model tended to overestimate the proportion of correct trials and underestimate error latencies.

p-χ^2_{379} was 378.41 by training epoch 1 653. The default parameter values (including *h* as calculated from the control data) did not adequately fit the data from the PS (p-χ^2_{380} = 952.19, *p-p* < .001) or NPS (p-χ^2_{380} = 491.38, *p-p* < .001) groups.

Model Representations of Hypothesized Deficits

Reduced processing speed. The model representation for processing speed is *h*, the duration (in real-time units) of the time required to complete each processing step. Higher values of *h* reflect more time (and reduced processing speed). Values for *h* were calculated according to the formula shown above rather than estimated. The usual decision rules from stochastic modeling were applied to determine whether a given data set required a unique value for *h*.

The optimal value of the time constant, *h*, was calculated to be 0.014. For the PS group data, calculating *h* to be 0.021 significantly reduced p-χ^2_{379} to 354.07 (*ns*, Δp-χ^2_1 = 598.12, *p-p* < .0001). This reduction, however, was an artifact of substantially inflated variance of model predictions for latency. The observed standard deviation for latency was 0.63046 s, compared with 1.327 s for the model predictions. Corresponding standard deviations of observed dissimilarity values and their model predictions were 0.3106 and 0.2263. Using the control value for *h*, we found that p-χ^2_{380} was 952.19 (*p-p* < .001); observed and modeled standard deviations were more similar for both judgment values and their latencies: 0.3106, 0.2263 and 0.6346, 0.7584.

For the NPS group, recalculating *h* to 0.016 improved model fit (p-χ^2_{379} = 348.25, *ns*, Δp-χ^2_1 = 143.13, *p-p* < .001). In contrast to the PS's comparison, there was no indication of model-variance inflation. Observed and modeled standard deviations for judgment latency and dissimilarity were 1.4567 (observed) versus 1.4495 and 0.5642 (observed) versus 0.2265. An unexpected result was that importing the value of *h* calculated for the PS group decreased p-χ^2 for the NPS group. Closer inspection revealed that model variances were enlarged relative to those of the NPS's own recomputed *h*: a standard deviation of 1.8244 (as opposed to 1.4495 using the *h* value from the NPS group, and as compared with the observed standard deviation of 1.4567).

On balance, an increase in *h* improved model fit for the NPS. There was some indication that the PS and NPS participants share similar values, but those values were associated with model-variance inflation for the PS group. Subsequent analyses used the control value for *h*. In general, model fit tended to be highly sensitive to small changes in *h*.

Additional processing. The model representation for the number of processing steps was the number of iterations before the model settled, *K*. This parameter was manipulated indirectly by adjusting the threshold for a just-noticeable difference, ι. Smaller values of ι imposed a stricter criterion (requiring more processing and resulting in higher values for *K*), whereas larger values of ι reflected a more lenient criterion (requiring less processing and resulting in lower values for *K*).

The optimum value for ι estimated for the control group was 0.001. For the PS group, reducing ι from 0.001 to 0.0002 significantly improved model prediction (p-χ^2_{379} = 432.64, *p-p* < .03, Δp-χ^2_1 = 545.27, *p-p* < .001). Similar results occurred for the second instance of the network, except that an estimated value of 0.0003 lowered p-χ^2_{379} to 348.92 (*ns*). For the NPS group, reducing ι to 0.00045 significantly improved fit (p-χ^2_{379} = 321.25, *ns*, Δp-χ^2_1 = 120.13, *p-p* < .001). Improvements in fit with changes in ι were demonstrated across cost functions, and they were more pronounced than was found for changes in *h* (see detailed comparisons in Carter, 2006, p. 63). Figure 2 compares observed data and model predictions for the preferred hypothesis of additional processing steps.

As predicted by the preferred hypothesis of additional encoding processes, the value for ι was reduced for the schizophrenia groups as compared with the control groups. Further, this manipulation demonstrated specificity for the parameter values between the two patient groups. Inserting the PS value of 0.0002 resulted in a significantly poor fit for the NPS group (Δp-χ^2_1 = 71.06, *p-p* < .001). That is, the two patient groups need to have different values for ι for the model to account for their respective data. This finding provides additional evidence for the additional processing hypothesis because the observed data are different for the two patient groups, and any model that accounts for both must therefore have different parameter values.

Neurotransmitter dysfunction. In the current model, dopamine function is represented by the network's *gain* parameter. A decrease in *gain* flattens the connectionist model's classical activation function (see Cohen & Servan-Schreiber, 1992), effectively creating the noisier signals that these authors associated with decreased dopamine activity. Conversely, an increase in *gain* would sharpen the activation function, representing increased dopamine activity.

One significant difference between the current model and that of Cohen and Servan-Schreiber (1992) is that they adjusted the *gain* parameter for only certain units, identified as *contextualizing units.* Contextual cues were not an integral aspect of the current task. The effectiveness of manipulating the *gain* parameter for the entire network was confirmed empirically (see online appendix).

Results were in the opposite direction than would be predicted by a decreased dopamine hypothesis. For the control group, mock testing of the *gain* parameter to an estimated value of 1.15 (vs. the default value of 1.00) significantly improved model fit (p-χ^2_{378} = 317.44, *ns*, Δp-χ^2_1 = 51.29, *p-p* < .001). For the PS group, an estimated value for *gain* of 1.32 significantly improved fit (p-χ^2_{379} = 735.82, *p-p* < .001, Δp-χ^2_1 = 216.37, *p-p* < .001) and was significantly superior to the optimal value for the control group (Δp-χ^2_1 = 67.32, *p-p* < .001). For the NPS group, increasing *gain* to an estimated value of 1.28 significantly improved predictive accuracy (Δp-χ^2_1 = 72.08, *p-p* < .001). Substituting the PS value of 1.32 for the NPS data retained the predictive accuracy achieved with the estimated value of 1.28, but substituting the controls' estimated value of 1.15 for the NPS data lowered the model fit (Δp-χ^2_1 = 5.25, *p-p* < .025).

Estimated values for *gain* were not stable. Across cost functions, estimates using both *SS* and pseudo-χ^2 were less than one for both the PS and NPS groups. This result tends to support a decreased dopamine hypothesis, contrary to the results found with the use of the p-χ^2 criterion. Further, the estimated value of *gain* for the NPS data was not stable across instances of the network. For the second instance of the network, the optimal value for the NPS data was the same as for the PS data (i.e., 1.46).

The current findings did not support the neurotransmitter dysfunction hypothesis within this model for a facial affect judgment task. Mock testing indicated that adjusting *gain* improved fit even

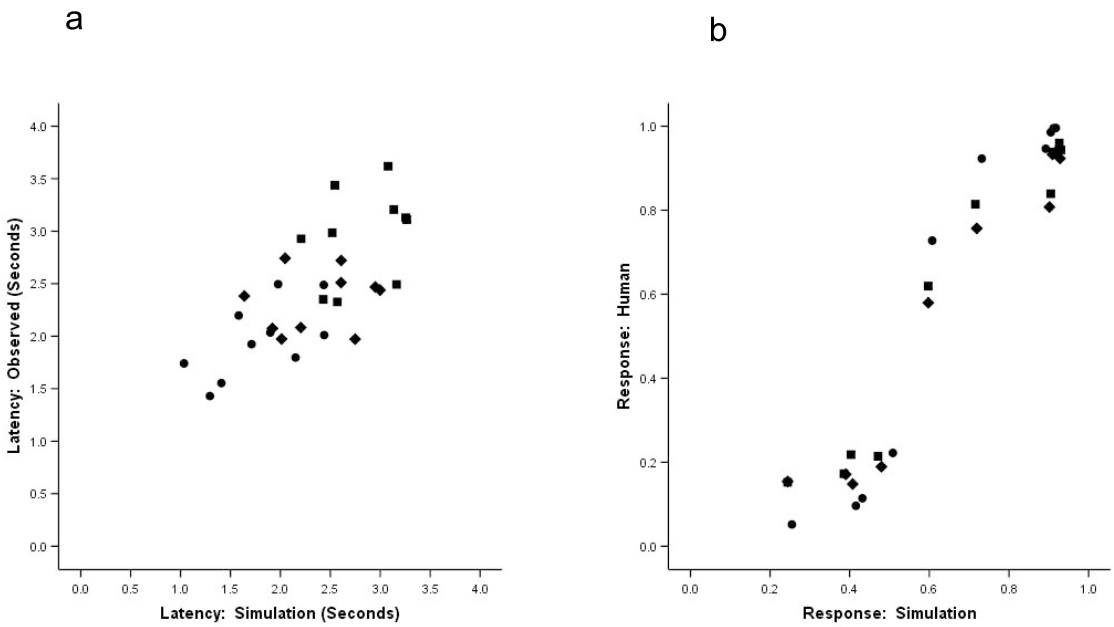

Figure 2. Observed latencies (a) and judgment responses (b) plotted against model predictions. Each datum represents the mean for one combination of facial expression (e.g., *happy–happy*), for the first instance of the network (the configuration being essentially identical for the second instance). Elevation in latencies from control to nonparanoid schizophrenia to paranoid schizophrenia groups expresses increased response time occurring to additional iterations (corresponding to added subprocesses), effected through reduction in ι. Circles = control participants; diamonds = nonparanoid schizophrenia participants; squares = paranoid schizophrenia participants.

for the control group. Results were not consistent across different cost functions or instances of the network.

Neural atrophy. To simulate the (functional) elimination of entire groups of neurons, one can set the activation of a unit to zero (e.g., Hinton & Shallice, 1991). A rudimentary neural Darwinist perspective predicts that less important neurons are systematically eliminated, and its variant predicts that more important neurons are eliminated. A plain physiological disease mechanism would eliminate neurons without regard for their importance. In this model, reduced importance is represented by relatively small weights (either afferent or efferent) or a low contribution to the response (mathematically, the partial derivative of error with respect to the activation value of the unit). Loss may be biased toward units whose connections are predominantly excitatory or inhibitory (modeled by the actual weights and derivatives) or by their relative strength (modeled by the squared weights and derivatives). Loss may also be random with respect to the importance of units. All of these possibilities were tested against each other.

This manipulation did not account for the observed data. For the PS data, eliminating network units failed to improve model fit. For the NPS data, eliminating network units improved model fit, but only for the first instance of the network. Further, the extensiveness of the damage and the exact units to be eliminated were not clear. Removing 0.21 of the units with the lowest sum of squared efferent weights improved model fit (p-$p < .001$), but so did removing 0.02 of the units with the lowest sum of afferent weights (p-$p < .001$).

Early processing deficits. Deficits in visual processing were simulated by adding noise to the activations of the input units (as in Farah & McClelland, 1991). Because the most appropriate shape for the noise distribution was unknown, two viable alternatives (uniform and Gaussian) were implemented. Two parameters are associated with this manipulation. The first parameter, the standard deviation of the noise distribution, reflects dispersion. Greater standard deviations correspond to more severe distortions. The second parameter, the mean of the distribution, reflects bias.

Early processing deficits do not appear to be sufficient on their own to account for the observed deficits. For the PS data, model fit was improved for the second instance of the network when noise was added to the input units. In each case, however, the fit was improved even further when noise was also added to other unit layers. (See Carter & Neufeld, 2003, for details.) Turning to the NPS data, one sees that the results were similar to those for the PS, except for one instance in which adding noise to the input units was sufficient. For the second instance of the network, adding noise from a Gaussian distribution with a mean of −0.091 and a standard deviation of 0.089 increased fit (p-χ^2 (378) = 327.03, Δp-χ^2_2 = 87.07, p-$p < .001$). In this case, adding noise to other unit layers in addition to the input layer failed to improve model fit (Δp-χ^2 (*df* for change) < 1.0).

The current data did not support the early processing deficit hypothesis on this task. Results were not consistent across instances of the network, and the results were not specific in that adding noise to other units further improved fit.

Deviant cortical pruning. As in previous research (e.g., Devlin et al., 1998), the loss of specific neural connections was simulated by eliminating the weights between processing units. The distinction between systematic and random elimination has important

theoretical implications (i.e., between a neurodevelopmental or neural Darwinist perspective and that of a purely physiological disease process). The operationalization of the (un)importance of connections is unresolved in the connectionist literature. One method is to consider that weights (the connections between processing units) with smaller absolute values represent the less effective and less significant connections (e.g., Hoffman et al., 1995). Alternatively, Li (1994) suggested that the weights that carry less information should be considered less important. This value is operationalized mathematically as the partial derivative of the change in error with respect to the change in weight. The situation is further complicated because both the weight and the partial derivative can potentially be either positive or negative.

All possible options were directly compared. In turn, the weights with the highest (representing the most excitatory connections) and lowest (representing the most inhibitory connections) values were eliminated. Similarly, the weights with the highest excitatory and inhibitory effects on empirical-response reproduction were eliminated. These analyses were repeated using the squared values, a function of the absolute magnitude of the weights and derivatives. The highest squared values represented the greatest influence, and the lowest squared values represented those with the least influence. Finally, weights were also eliminated at random. The effectiveness of deactivating individual weights was confirmed for the current modeling context (see online appendix).

Cortical pruning did not appear to account for the data. Eliminating network weights failed to improve model fit for either the PS or the NPS groups.

Other mechanisms may also be appropriate to simulate cortical pruning. To account for the possibility that the distortion of information may be greater for either excitatory or inhibitory activity (Grossberg, 2002), one can use an alternative method of simulating synaptic dysfunction: adding noise to the activations of individual units (e.g., Hoffman, 1992). The details of this intervention were the same as for adding noise to the input units, except that noise was added to the input and hidden units. Alternatively, David (1994) has developed a theory of schizophrenia symptomatology based on the neurocognitive concept of modularity, in which dysfunction results from insufficient cortical pruning. This possibility can be simulated by multiplying each weight by some number drawn from a random distribution, allowing for the possibility that connections may be either weakened or excessively strengthened. Both of these alternative representations of the cortical pruning hypothesis were tested extensively (see Carter & Neufeld, 2003), but they were rejected for two reasons. First, adding noise and distorting weights improved fit on mock testing. Second, results were inconsistent across instances of the network, cost functions, and details of the timing of noise during any given judgment trial.

The current results did not support the cortical pruning hypothesis for a facial affect judgment task. A wide range of possible options in terms of which connections should be eliminated were tested by three different representations of the hypothesis, but no consistent findings emerged.

Discussion

This study evaluated connectionist simulations of several competing hypotheses concerning cognitive deficits regarding processing facially expressed affect in schizophrenia. These particular deficits may relate directly to social impairment (Penn, Corrigan, Bentall, Racenstein, & Newman, 1997). The hypotheses were tested in direct competition with each other through stringent tests of predictive accuracy. The model predicted both judgment accuracy and latency. Consistency of modeling results was evaluated across schizophrenia subgroups and with different cost functions. Generalization testing (Busemeyer & Wang, 2000) and cross-validation (Browne, 2000) were given high priority in selecting a preferred model. This study is an example in clinical cognitive science (Stein & Young, 1992) of the convergence of all three levels of analysis (implementational, algorithmic, and computational; Marr, 1982). Although the current study focused on the algorithmic level (connectionist modeling), it is flanked by previous findings at the computational (stochastic modeling) and implementational (neurobiological) levels (Boksman et al., year?; Boksman et al., 2007).

The modeled hypothesis that best survived this crucible of empirical testing was that additional processing subprocesses account for deviations in both judgment content and their durations for people with either paranoid or other subtypes of schizophrenia on a facial affect judgment task. This hypothesis originates from computational-level (stochastic modeling) analysis. The mechanism of increased subprocesses has been identified as the key difference between schizophrenia and control participants across multiple cognitive paradigms (Neufeld, 1989, 1990; Neufeld et al., 2002, 1993; Neufeld & Williamson, 1996).

A particular advantage of the additional processing model is the parsimonious explanation for both response content and response latency across groups. Not only does the extra processing take time, it may also introduce retroactive interference that undermines the accuracy of judgments (Carter & Neufeld, 1999; Chechile, 1987). In this sense, fitting predictions to schizophrenia-performance data is analogous to *model overfitting*, whereby model predictions have been crafted to accommodate both unstable and systematic fluctuations in empirical observations (Myung, Forster, & Browne, 2000). An important aspect of this study was the successful modeling of different results from the PS and NPS groups. The observed differences were parsimoniously explained by the same basic mechanism, but the two groups required different parameter values. We found that the value for the PS data is more deviant from the control value than is the value for the NPS data, results that are consistent with previous findings.

A competing hypothesis bearing directly on response latencies is reduced processing speed (represented by an increased value of the time constant, h). In the current study, additional processing outperformed reduced processing speed by improving model fit everywhere that reduced processing speed led to improvement—and additionally by improving fit in tests in which reduced processing speed failed. This finding suggests that the time for the processing of each subprocess completion is intact among schizophrenia participants, which is consistent with previous findings (reviewed in Neufeld et al., 2002).

Note that despite meeting the present homogeneous-group criterion, latency for the student control participants was less than that of the HRDC controls. Thus, the admixture of the two sets of participants in each of the model-training groups produced lower latency values than would have been the case without student participation. It might be conjectured, therefore, that despite other

advantages (e.g., increased power for model rejection), the success obtained by the parameterized additional-subprocess hypothesis (reduced ι) was an artifact of such reduced latencies. This possibility, however, seems unlikely, as follows.

Because predictions involving ι were constrained to accommodate both judgments and their latencies across the 180 judgment trials, any dissociation between these two response properties would strain empirical fit (see, e.g., Ashby, Maddox, & Lee, 1994). In contrast, change in the time constant *h* affects latency predictions only, thereby potentially accommodating more readily any disproportionate lowering of observed values. Reduction in ι nevertheless led to a more satisfactory empirical fit throughout.

Moreover, if a change in ι were being driven primarily by latency differences, judgment response predictions would, it could be argued, suffer by comparison. Inspection of Figure 2 fails to confirm this possibility (indeed, it indicates the opposite; these observations being ratified with subsidiary analysis of associated cost functions, applied separately to latency and judgment responses). Finally, the above results were duplicated when students were not involved as control participants (see online appendix).

One advantage of a formal model is that the mathematical formulation exposes the limits of knowledge. The deficit related to additional encoding subprocesses has been theoretically related to thought-content disorder, stress-negotiation deficit, and other clinical variables, and it has been found across a wide range of cognitive tasks and with both inpatient and outpatient samples (see, e.g., Carter & Neufeld, 1999; Neufeld et al., 2002, 1993). This deficit has now been replicated in this study for a facial affect judgment task. As Hoffman and McGlashan (2006) pointed out, modeling the entirety of schizophrenia is clearly an impossible task. Other connectionist modeling presents compelling accounts of prominent schizophrenia symptomatology, such as speech-processing alterations associated with auditory hallucinations (e.g., Hoffman & McGlashan, 2001, 2006; McGlashan & Hoffman, 2000), as well as deficits on a range of specific cognitive tasks (e.g., Cohen & Servan-Schreiber, 1992; Schweikert, 1978). Successful pretesting of the network manipulations (see online appendix) demonstrated that the current modeling framework could detect cases in which deficits other than additional processing were present.

Methodological innovations required within the current investigative context may potentially be exported to other connectionist modeling endeavors. The time constant, *h*, provides for continuous-time mapping of simulation data onto empirical responses. The formula for *h* was derived through a closed-form solution, eliminating the uncertainty and additional computation associated with computer-search estimates. This parameter can be applied to any number of similar modeling situations. The parameter *h* allows, as well, for the separation of parameters associated with the rate of processing from the number of processing steps. The importance of this separation has been underscored in formal modeling, specifically of cognitive performance in schizophrenia (Neufeld, in press).

Several modeling procedures were used to test the reliability and validity of the conclusions drawn from these studies. Principled model competition was undertaken, in that each competitor was motivated by previous theory and empirical findings. Cross-validation of hypothesized dysfunction assessed the reliability of inferences across control samples and cost functions. In addition, mock testing of modeled dysfunction screened out spurious increases in model predictive accuracy.

The current study focused on identifying the mechanism underlying deficits in the accuracy and latency in facial affect judgments among people with schizophrenia, as if those deficits were static. This approach allowed for detailed and rigorous descriptions of the deficit phenomenon and its immediate sources. It nevertheless puts aside potentially important developmental and learning considerations, as follows. Schizophrenia may be the product of a normal developmental process with a malfunction (e.g., Saugstad, 1989; Waddington, 1993; Weinberger, 1987) or a subtle variation (e.g., Benes et al., 1986). Connectionist frameworks other than back propagation may be more neurologically plausible and better suited to investigate these more distant causes of deficit in schizophrenia.[3]

Of (computational) necessity, the current model also ignored learning that occurred during the task itself. Other algorithms (e.g., Vickers & Lee, 2000) may be able to account for serial position effects. One related finding is worthy of further investigation. For all groups, latency tended to decrease over trials (r ranged from $-.45$ to $-.54$, all ps $< .001$). For the control groups and the PS group, accuracy increased over trials (r ranged from .17 to .20, all ps $< .02$). For the NPS group, however, the trend was for accuracy to decrease over trials ($r = -.12$, $p < .10$). Accounting for such serial-position effects would require potentially intractable expansion of the present modeling (cf., Yechiam et al., 2007). Note that the above across-trial variation in no way undermines the present modeling of within-trial dynamics, because the model's empirical fit withstands challenges from this additional source of variance.

Current computer simulations provide a mathematical representation of processing deficits in schizophrenia. These deficits have previously been framed in closed-form expressions as stochastic models. This account represents a brain metaphor (McClelland & Rumelhart, 1988) for additional subprocesses in cognitive processing. Emphasis has been on facial affect judgments, in part because of their potential significance for social functioning. Standardized arrays of posed, still expressions (required for tractability of modeling) may not be as ecologically valid as moving displays of spontaneous expressions in context (Carroll & Russell, 1996, 1997). The underlying processing mechanisms and their associated parameters, however, could well be general. Indeed, they were derived from formal mathematical models of different tasks altogether. In terms of parsimony, the simultaneous prediction of both response content and response latency is a notable asset of the presented formulation.

[3] We thank an anonymous reviewer for this observation.

References

Ashby, F. G., Maddox, W. T., & Lee, W. W. (1994). On the dangers of averaging across subjects when using multidimensional scaling or the similarity-choice model. *Psychological Science, 5,* 144–151.

Barta, P. E., Pearlson, G. D., Powers, R. E., Richards, S. S., & Tune, L. E. (1990). Auditory hallucinations and smaller superior temporal gyral volume in schizophrenia. *American Journal of Psychiatry, 147,* 1457–1462.

Bartha, R. B., Williamson, P. C., Drost, D. J., Malla, A., Carr, T. J.,

Cortese, L., et al. (1997). Measurement of glutamate and glutamine in the medial prefrontal cortex of never-treated schizophrenic patients and healthy controls by proton Magnetic Resonance Spectroscopy. *Archives of General Psychiatry, 54,* 959–965.

Batchelder, W. H. (1998). Multinomial processing tree models and psychological assessment. *Psychological Assessment, 10,* 331–334.

Beech, A. R., McManus, D., Baylis, G., Tipper, S. P., & Agar, K. (1991). Individual differences in cognitive processes: Towards an explanation of schizophrenic symptomatology. *British Journal of Psychology, 82,* 417–426.

Benes, F. M., Davidson, J., & Bird, E. D. (1986). Quantitative cytoarchitectural studies of the cerebral cortex of schizophrenics. *Archives of General Psychiatry, 43,* 31–35.

Benes, F. M., McSparren, J., Bird, E. D., SanGiovanni, J. P., & Vincent, S. L. (1991). Deficits in small interneurons in prefrontal and cingulate cortices of schizophrenic and schizoaffective patients. *Archives of General Psychiatry, 48,* 996–1001.

Boksman, K., Miller, J., Williamson, P., Drost, D., Manchanda, R., Densmore, M., et al. (2007). *High-field fMRI-monitored effective neuroconnectivity of formally modeled encoding deficit in schizophrenia.* Manuscript in preparation.

Boksman, K., Théberge, J., Williamson, P., Drost, D., Malla, A., Densmore, M., et al. (2005). A 4.0 Tesla fMRI study of brain connectivity during word fluency in first episode schizophrenia. *Schizophrenia Research, 75,* 247–263.

Borod, J. C., Alpert, M., Brozgold, A., Martin, C., Welkowitz, J., Diller, L., et al. (1989). A preliminary comparison of flat affect schizophrenics and brain-damaged patients on measures of affective processing. *Journal of Communication Disorders, 22,* 93–104.

Botvinick, M. M., & Plaut, D. C. (2006). Short-term memory for serial order: A recurrent neural network model. *Psychological Review, 133,* 201–233.

Browne, M. W. (2000). Cross-validation methods. *Journal of Mathematical Psychology, 44,* 108–132.

Burch, J. W. (1995). Typicality range deficit in schizophrenic's recognition of emotion in faces. *Journal of Clinical Psychology, 51,* 140–152.

Busemeyer, J. R., & Stout, J. C. (2002). A contribution of cognitive decision models to clinical assessment: Decomposing performance on the Bechara Gambling Task. *Psychological Assessment, 14,* 253–262.

Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review, 100,* 432–459.

Busemeyer, J. R., & Wang, Y. (2000). Model comparisons and model selections based on generalization test methodology. *Journal of Mathematical Psychology, 44,* 171–189.

Carroll, J. M., & Russell, J. A. (1996). Do facial expressions signal specific emotions? Judging emotion from the face in context. *Journal of Personality and Social Psychology, 70,* 205–218.

Carroll, J. M., & Russell, J. A. (1997). Facial expressions in Hollywood's portrayal of emotion. *Journal of Personality and Social Psychology, 72,* 164–176.

Carter, J. R. (2006). *Technical report on facial expression analysis in schizophrenia.* Retrieved May 1, 2006, from http://www.vanier.com/professionals.innovations.shtm

Carter, J. R., & Neufeld, R. W. J. (1999). Cognitive processing of multidimensional stimuli in schizophrenia: Formal modeling of judgment speed and content. *Journal of Abnormal Psychology, 108,* 633–654.

Carter, J. R., & Neufeld, R. W. J. (2003). *Department of Psychology Research Bulletin No. 753.* London, Ontario, Canada: University of Western Ontario.

Carter, J. R., Neufeld, R. W. J., & Benn, K. D. (1998). Application of process models in assessment psychology: Potential assets and challenges. *Psychological Assessment, 10,* 279–298.

Cegalis, J. A., Leen, D., & Solomon, E. J. (1977). Attention in schizophrenia: An analysis of selectivity in the functional visual field. *Journal of Abnormal Psychology, 86,* 470–482.

Chechile, R. A. (1987). Trace susceptibility theory. *Journal of Experimental Psychology, 116,* 203–222.

Cohen, J. D., & Servan-Schreiber, D. (1992). Context, cortex, and dopamine: A connectionist approach to behavior and biology in schizophrenia. *Psychological Review, 99,* 45–77.

Corrigan, P. W., & Green, M. F. (1993). Schizophrenic patients' sensitivity to social cues: The role of abstraction. *American Journal of Psychiatry, 150,* 589–594.

Creese, I., Burt, D. R., & Snyder, S. H. (1976, April 30). Dopamine receptor binding predicts clinical and pharmacological potencies of antischizophrenic drugs. *Science, 192,* 481–483.

Crow, T. J. (1990). Temporal lobe asymmetries as the key to the etiology of schizophrenia. *Schizophrenia Bulletin, 16,* 433–443.

David, A. S. (1994). Dysmodularity: A neurological model for schizophrenia. *Schizophrenia Bulletin, 20,* 249–255.

Davis, J. (1976). Comparative doses and costs of antipsychotic medication. *Archives of General Psychiatry, 33,* 858–861.

Davis, P. J., & Gibson, M. G. (2000). Recognition of posed and genuine facial expressions of emotion in paranoid and nonparanoid schizophrenia. *Journal of Abnormal Psychology, 109,* 445–450.

Deakin, J. F. W. (1994). Neuropsychological implications of brain changes in schizophrenia: An overview. *Psychopathology, 27,* 251–254.

Deering, G. (1963). Affective stimuli and disturbance of thought processes. *Journal of Consulting Psychology, 27,* 338–343.

Devlin, J. T., Gonnerman, L. M., Andersen, E. S., & Seidenberg, M. S. (1998). Category-specific semantic deficits in focal and widespread brain damage: A computational account. *Journal of Cognitive Neuroscience, 10,* 77–94.

Doob, J. L. (1953). *Stochastic processes.* New York: Wiley.

Dunn, L. M., & Dunn, L. M. (1981). *Peabody Picture Vocabulary Test—Revised.* Circle Pines, MN: American Guidance Service.

Edelman, G. (1986). *Neural Darwinism: The theory of neuronal group selection.* New York: Basic Books.

Ekman, P., & Friesen, W. V. (1976). *Pictures of facial affect.* Palo Alto, CA: Consulting Psychologists Press.

Ennis, J. H., & Whelton, C. (1994). The relationship between face recognition, facial affect recognition and social skills in schizophrenia. *Canadian Journal of Psychiatry, 39,* 58.

Farah, M. J., & McClelland, J. L. (1991). A computational model of semantic memory impairment: Modality specificity and emergent category specificity. *Journal of Experimental Psychology: General, 120,* 339–357.

Feinberg, I. (1982/1983). Schizophrenia: Caused by a fault in programmed synaptic elimination during adolescence? *Journal of Psychiatric Research, 17,* 319–334.

Feinberg, I. (1990). Cortical pruning and the development of schizophrenia. *Schizophrenia Bulletin, 16,* 567–568.

Flack, W. F., Jr., Cavallaro, L. A., Laird, J. D., & Miller, D. R. (1997). Accurate encoding and decoding of emotional facial expressions in schizophrenia. *Psychiatry, 60,* 197–210.

Friston, K. J., Buechel, C., Fink, G. R., Morris, J., Rolls, E., & Dolan, R. J. (1997). Psychophysiological and modulatory interactions in neuroimaging. *NeuroImage, 6,* 218–229.

Frith, C. D. (1979). Consciousness, information processing and schizophrenia. *British Journal of Psychiatry, 134,* 225–235.

Frith, C. D., Stevens, M., Johnstone, E. C., Owens, D. G. C., & Crow, T. J. (1983). Integration of schematic faces and other complex objects in schizophrenia. *Journal of Nervous and Mental Disease, 171,* 34–39.

George, L., & Neufeld, R. W. J. (1987a). Attentional resources and hemispheric functional asymmetry in schizophrenia. *British Journal of Clinical Psychology, 26,* 35–45.

George, L., & Neufeld, R. W. J. (1987b). Magical ideation and schizophrenia. *Journal of Consulting and Clinical Psychology, 55,* 778–779.

Gessler, S., Cutting, J., Frith, C. D., & Weinman, J. (1989). Schizophrenic inability to judge facial emotion: A controlled study. *British Journal of Clinical Psychology, 28,* 18–29.

Gillis, C. G. (1998). (Ed.) *Compendium of pharmaceuticals and specialties* (33rd ed.). Toronto, Ontario, Canada: Canadian Pharmacists Association.

Grossberg, S. (2002, September). *The imbalanced brain: From normal behavior to schizophrenia.* Paper presented at the meeting of the Society for Research in Psychopathology, San Francisco, CA.

Gur, R. E., McGrath, C., Chan, R. M., Schroeder, L., Turner, T., Turestsky, B. I., et al. (2002). An fMRI study of facial emotional processing in patients with schizophrenia. *American Journal of Psychiatry, 159,* 1992–1999.

Hasselmo, M. E., Rolls, E. T., & Baylis, G. C. (1989). The role of expression and identity in the face-selective responses of neurons in the temporal visual cortex of the monkey. *Behavioural Brain Research, 32,* 203–218.

Hempel, A., Hempel, E., Schonknecht, P., Stippich, C., & Schroder, J. (2003). Impairment in basal limbic function in schizophrenia during affect recognition. *Psychiatry Research, 122,* 115–124.

Heywood, C. A., & Cowey, A. (1992). The role of the "face-cell" area in the discrimination and recognition of faces by monkeys. *Philosophical Transactions of the Royal Society of London, Series B, 335,* 31–38.

Hinton, G. E., & Shallice, T. (1991). Lesioning an attractor network: Investigations of acquired dyslexia. *Psychological Review, 98,* 74–95.

Hoffman, R. E. (1992). Attractor networks and psychotic disorders [CD-ROM]. *Psychiatric Annals, 27,* 119–124.

Hoffman, R. E., & Dobscha, S. K. (1989). Cortical pruning and the development of schizophrenia. *Schizophrenia Bulletin, 15,* 477–490.

Hoffman, R. E., & Dobscha, S. K. (1990). Reduced cortico-cortical exchange of information as a possible etiology of schizophrenia: Reply to Feinberg. *Schizophrenia Bulletin, 16,* 569–570.

Hoffman, R. E., & McGlashan, T. H. (1993). Parallel distributed processing and the emergence of schizophrenic symptoms. *Schizophrenia Bulletin, 19,* 119–140.

Hoffman, R. E., & McGlashan, T. H. (1994). Corticocortical connectivity, autonomous networks, and schizophrenia. *Schizophrenia Bulletin, 20,* 257–261.

Hoffman, R. E., & McGlashan, T. H. (1998). Reduced corticocortical connectivity can induce speech perception pathology and hallucinated "voices." *Schizophrenia Research, 30,* 137–141.

Hoffman, R. E., & McGlashan, T. H. (2001). Neural networks of schizophrenia. *Neuroscientist, 7,* 441–454.

Hoffman, R. E., & McGlashan, T. H. (2006). Using a speech perception neural network simulation to study normal neurodevelopment and auditory hallucinations in schizophrenia. In R. W. J. Neufeld (Ed.), *Advances in clinical cognitive science: Formal modeling and assessment of processes and symptoms.* Washington, DC: American Psychological Association.

Hoffman, R. E., Rapaport, J., Ameli, R., McGlashan, T. H., Harcherik, D., & Servan-Schreiber, D. (1995). A neural network simulation of hallucinated "voices" and associated speech perception impairments in schizophrenic patients. *Journal of Cognitive Neuroscience, 7,* 479–496.

Hollingshead, B. A. (1957). *Two-factor index of social position.* Unpublished manuscript (available from Richard W. J. Neufeld).

Izard, C. E. (1959). Paranoid schizophrenic and normal subjects' perceptions of photographs of human faces. *Journal of Consulting Psychology, 23,* 119–124.

Kapur, S., Zipursky, R. B., Remington, G., Jones, C., DaSilva, J., Wilson, A. A., & Houle, S. (1998). 5-HT_2 receptor occupancy of olanzapine in schizophrenia: A PET investigation. *American Journal of Psychiatry, 155,* 921–928.

Knight, R. A. (1993). Comparing cognitive models of schizophrenics' input dysfunction. In R. L. Cromwell & C. R. Snyder (Eds.), *Schizophrenia: Origins, processes, treatment, and outcome* (pp. 151–175). London: Oxford University Press.

Knight, R. A., & Silverstein, S. M. (2001). A process-oriented approach for averting confounds resulting from general performance deficiencies in schizophrenia. *Journal of Abnormal Psychology, 110,* 15–30.

Kosaka, H., Omori, M., Murata, T., Iidaka, T., Yamada, H., Okada, T., et al. (2002). Differential amygdala response during facial recognition in patients with schizophrenia: An fMRI study. *Schizophrenia Research, 57,* 87–95.

LaRusso, L. (1978). Sensitivity of paranoid patients to nonverbal cues. *Journal of Abnormal Psychology, 87,* 463–471.

Levy, L. H., Orr, T. B., & Rosenzweig, S. (1960). Judgments of emotion from facial expressions by college students, mental retardates, and mental hospital patients. *Journal of Personality, 28,* 342–349.

Li, S.-C. (1994). *A diagnostic approach to model evaluation: Parameter sensitivity and interdependence analysis in models of cognition.* Unpublished doctoral dissertation. Norman: University of Oklahoma.

Lubow, R. E., Weiner, I., Schlossberg, A., & Baruch, I. (1987). Latent inhibition and schizophrenia. *Bulletin of the Psychonomic Society, 25,* 464–467.

Magaro, P., Abrams, L., & Cantrell, P. (1981). The Maine Scale of paranoid and nonparanoid schizophrenia: Reliability and validity. *Journal of Consulting and Clinical Psychology, 49,* 438–447.

Maher, B. (1966). *Principles of psychopathology: An experimental approach.* New York: McGraw-Hill.

Mandal, M. K., & Rai, A. (1987). Responses to facial emotion and psychopathology. *Psychiatry Research, 20,* 317–323.

Marr, D. (1982). *Vision.* San Francisco: Freeman.

McClelland, J. L., & Rumelhart, D. E. (1988). *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol. 3). Cambridge, MA: Bradford Books.

McGlashan, T. H., & Hoffman, R. E. (2000). Schizophrenia as a disorder of developmentally reduced synaptic connectivity. *Archives of General Psychiatry, 57,* 637–648.

McRae, K., & Hetherington, P. A. (1993). Catastrophic interference is eliminated in pretrained networks. In *Proceedings of the Fifteenth Annual Conference of the Cognitive Science Society* (pp. 723–728). Hillsdale, NJ: Erlbaum.

Mueser, K. T., Doonan, R., Penn, D. L., Blanchard, J. J., Bellack, A. S., Nishith, P., & DeLeon, J. (1996). Emotion recognition and social competence in chronic schizophrenia. *Journal of Abnormal Psychology, 105,* 271–275.

Myers, J. L., & Well, A. D. (1991). *Research design and statistical analysis.* New York: HarperCollins.

Myung, I. J., Forster, M. R., & Browne, M. W. (2000). Special issue on model selection. *Journal of Mathematical Psychology, 44,* 1–2.

Neufeld, R. W. J. (1976). Simultaneous processing of multiple stimulus dimensions among paranoid and nonparanoid schizophrenics. *Multivariate Behavioral Research, 4,* 425–442.

Neufeld, R. W. J. (1989, July). *Stochastic modeling of memory search among paranoid schizophrenics.* Paper presented at the Society for Mathematical Psychology, Irvine, CA.

Neufeld, R. W. J. (1990, August). *Stimulus encoding in schizophrenia, and the capacity deficit hypothesis.* Paper presented at the Annual Meeting of the Society for Mathematical Psychology, Toronto, Ontario, Canada.

Neufeld, R. W. J. (1991). Memory in paranoid schizophrenia. In P. A. Magaro (Ed.), *The cognitive bases of mental disorders: Annual review of psychopathology* (Vol. 1, pp. 231–261). Newbury Park: Sage.

Neufeld, R. W. J. (in press). Composition and uses of formal clinical cognitive science. In W. Spaulding & J. Poland (Eds.), *Modeling complex systems: Motivation, cognition and social processes: Nebraska Symposium on Motivation.* Lincoln: University of Nebraska Press.

Neufeld, R. W. J., Carter, J. R., Boksman, K., Jetté, J., & Vollick, D. (2002). Application of stochastic modeling to group and individual differences in cognitive functioning. *Psychological Assessment, 14,* 279–298.

Neufeld, R. W. J., Carter, J. R., Vollick, D., Boksman, K., Levy, L., George, L., & Jetté, J. (2007). A mathematical account of group and individual differences in memory-search facilitative stimulus encoding, with application to schizophrenia. In R. W. J. Neufeld (Ed.), *Advances in clinical cognitive science: Formal modeling and assessment of processes and symptoms* (pp. 147–177). Washington, DC: American Psychological Association.

Neufeld, R. W. J., & Gardner, R. C. (1990). Data aggregation in evaluating psychological constructs: Multivariate and logical-deductive considerations. *Journal of Mathematical Psychology, 34,* 276–296.

Neufeld, R. W. J., & McCarty, T. (1994). A formal analysis of stressor and stress-proneness effects on basic information processing. *British Journal of Mathematical and Statistical Psychology, 47,* 193–226.

Neufeld, R. W. J., Vollick, D., & Highgate, S. (1993). Stochastic modeling of stimulus encoding and memory search in paranoid schizophrenia: Clinical and theoretical implications. In R. L. Cromwell & R. C. Snyder (Eds.), *Schizophrenia: Origins, processes, treatment, and outcome: The Second Kansas Series in Clinical Psychology* (pp. 176–196). Oxford, England: Oxford University Press.

Neufeld, R. W. J., & Williamson, P. (1996). Neuropsychological correlates of positive symptoms: Delusions and hallucinations. In C. Pantelis, H. E. Nelson, & T. R. E. Barnes (Eds.), *Schizophrenia: A neuropsychological perspective* (pp. 205–235). London: Wiley.

Nicholson, I. R., & Neufeld, R. W. J. (1993). Classification of the schizophrenias according to symptomatology: A two-factor model. *Journal of Abnormal Psychology, 102,* 259–270.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychology, 9,* 97–113.

Pachella, R. (1974). The interpretation of reaction time in information processing research. In B. H. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition* (pp. 41–82). Hillsdale, NJ: Erlbaum.

Penn, D. L., Corrigan, P. W., Bentall, R. P., Racenstein, J. M., & Newman, L. (1997). Social cognition in schizophrenia. *American Journal of Psychiatry, 153,* 607–617.

Penn, D. L., & Spaulding, P. W. (1997). Introduction: Factors underlying social functioning in schizophrenia: Information processing and social perception. *Psychiatry, 60,* 279–280.

Perrett, D. I., Hietanen, J. K., Oram, M. W., & Benson, P. J. (1992). Organization and function of cells responsive to faces in the temporal cortex. *Philosophical Transactions of the Royal Society of London, Series B, 335,* 23–30.

Perrett, D. I., & Mistlin, A. J. (1990). Perception of facial characteristics by monkeys. In W. C. Stebbins & M. A. Berkley (Eds.), *Comparative perception: Vol. II: Complex signals* (pp. 187–215). New York: Wiley.

Perrett, D. I., Mistlin, A. J., Chitty, A. J., Harries, M. H., Newcombe, F., & de Haan, E. (1988). Neuronal mechanisms of face perception and their pathology. In C. Kennard & F. C. Rose (Eds.), *Physiological aspects of clinical neuro-ophthalmology* (pp. 137–154). London: Chapman & Hall.

Phillips, M. L., & David, A. S. (1997). Visual scan paths are abnormal in deluded schizophrenics. *Neuropsychologia, 35,* 99–105.

Pilowsky, I., & Bassett, D. (1980). Schizophrenia and the response to facial emotions. *Comprehensive Psychiatry, 21,* 236–244.

Remington, G., Kapur, S., & Zipursky, R. (1998). The relationship between risperidone plasma levels and dopamine D_2 occupancy: A positron emission tomography study [CD-ROM]. *Journal of Clinical Psychopharmacology, 18,* 82–83. (Review available at http://www.futur.com/resframes.htm)

Riefer, D. M., Knapp, B., Batchelder, W. H., Bamber, D., & Manifold, V. (2002). Cognitive psychometrics: Assessing storage and retrieval deficits in special populations. *Psychological Assessment, 14,* 184–201.

Roelfsema, P. R., & van Ooyen, A. (2005). Attention-gated reinforcement learning of internal representations for classification. *Neural Computation, 17,* 2176–2214.

Rolls, E. T. (1992). Neurophysiological mechanisms underlying face processing within and beyond the temporal cortical visual area. *Philosophical Transactions of the Royal Society of London, Series B, 336,* 11–21.

Saugstad, L. F. (1989). Age at puberty and mental illness: Towards a neurodevelopmental aetiology of Kraepelin's endogenous psychoses. *British Journal of Psychiatry, 155,* 536–544.

Schneider, F., Gur, R. C., Gur, R. E., & Shtasel, D. L. (1995). Emotional processing in schizophrenia: Neurobehavioral probes in relation to psychopathology. *Schizophrenia Research, 17,* 67–75.

Schweikert, R. A. (1978). A critical path generalization of the additive factor method: Analysis of a Stroop task. *Journal of Mathematical Psychology, 18,* 105–139.

Seeman, P., & Lee, T. (1975, June 20). Antipsychotic drugs: Direct correlation between clinical potency and presynaptic action on dopamine neurons. *Science, 188,* 1217–1219.

Seeman, P., Lee, T., Chau-Wong, M., & Wong, K. (1976, June 24). Antipsychotic drug doses and neuroleptic/dopamine receptors. *Nature, 261,* 717–719.

Selemon, L. D., Rajkowska, G., & Goldman-Rakic, P. S. (1995). Abnormally high neuronal density in the schizophrenia cortex. *Archives of General Psychiatry, 52,* 805–818.

Siegle, G. J. (1998). *Connectionist models of cognitive, affective, brain, and behavioral disorders.* Retrieved August 1, 1999, from http://www.sci.sdsu.edu/CAL/connectionist-models/

Siegle, G. J., Steinhauer, S. R., Thase, M. E., Stenger, V. A., & Carter, C. S. (2002). Can't shake that feeling: Event-related fMRI assessment of sustained amygdala activity in response to emotional information in depressed individuals. *Biological Psychiatry, 51,* 693–707.

Silva, J. A., Leong, G. B., Wine, D. B., & Saab, S. (1992). Evolving misidentification syndromes and facial recognition deficits. *Canadian Journal of Psychiatry, 37,* 574–576.

Spitzer, M. (2000). *The mind within the net: Models of learning, thinking, and acting.* Cambridge, MA: MIT Press.

Spohn, H. E., & Strauss, M. E. (1989). Relation of neuroleptic and anticholinergic medication to cognitive functions in schizophrenia. *Journal of Abnormal Psychology, 98,* 367–380.

Stein, D. J., & Young, J. E. (Eds.). (1992). *Cognitive science and clinical disorders.* San Diego, CA: Academic Press.

Townsend, J. T. (1984). Uncovering mental processes with factorial experiments. *Journal of Mathematical Psychology, 28,* 363–400.

Townsend, J. T., & Ashby, F. G. (1978). Methods of modeling capacity in simple processing systems. In J. Castellan & F. Restle (Eds.), *Cognitive theory* (Vol. 3). Hillsdale, NJ: Erlbaum.

Townsend, J. T., & Ashby, F. G. (1983). *Stochastic modeling of elementary psychological processes.* Cambridge, England: Cambridge University Press.

Tsao, D. Y., Freiwald, W. A., Tootell, R. B. H., & Livingstone, M. S. (2006, February 3). A cortical region consisting entirely of face-selective cells. *Science, 311,* 670–674.

Vickers, D., & Lee, M. D. (2000). Dynamic models of simple judgments: II. Properties of a self-organizing PAGAN (Parallel, Adaptive, Generalized Accumulator Network) model for multi-choice tasks. *Nonlinear Dynamics, Psychology, and Life Sciences, 4,* 1–31.

Waddington, J. T. (1993). Neurodynamics of abnormalities in cerebral metabolism and structure in schizophrenia. *Schizophrenia Bulletin, 19,* 55–69.

Wagenmakers, E., & Waldorp, L. (2006). Special issue on model selection. *Journal of Mathematical Psychology, 50,* 99–214.

Walker, E., McGuire, M., & Bettes, B. (1984). Recognition and identifi-

cation of facial stimuli by schizophrenics and patients with affective disorders. *British Journal of Clinical Psychology, 23,* 37–44.

Wallsten, T. S., Pleskac, T. J., & Lejuez, C. W. (2005). Modeling behavior in a clinically diagnostic sequential risk-taking task. *Psychological Review, 112,* 862–880.

Weinberger, D. R. (1987). Implications of normal brain development for the pathogenesis of schizophrenia. *Archives of General Psychiatry, 44,* 660–669.

Wenger, M. J., & Townsend, J. T. (2000). Basic response time tools for studying general processing capacity in attention, perception, and cognition. *Journal of General Psychology, 127,* 67–99.

Wyatt, R. J., & Torgow, J. S. (1976). A comparison of equivalent clinical potencies of neuroleptics as used to treat schizophrenia and affective disorders. *Journal of Psychiatric Research, 13,* 91–98.

Yechiam, E., Veinott, E. S., Busemeyer, J. R., & Stout, J. C. (2007). Cognitive models for evaluating basic decision processes in clinical populations. In R. W. J. Neufeld (Ed.), *Advances in clinical cognitive science: Formal modeling of processes and symptoms* (pp. 81–111). Washington, DC: American Psychological Association.

Yeung, N., Botvinick, M. M., & Cohen, J. D. (2004). The neural basis of error detection: Conflict monitoring and error-related negativity. *Psychological Review, 111,* 931–959.

Zorzi, M. (2002, August). *Network simulations of number processing and neural correlates.* Paper presented at the Second Annual Summer Interdisciplinary Conference, Squamish, British Columbia, Canada.

Received January 12, 2005
Revision received December 12, 2006
Accepted December 22, 2006 ■