Reinforcement Ambiguity and Novelty Do Not Account for Transitive Inference Deficits in Schizophrenia

Michael J. Coleman, Debra Titone, Olga Krastoshevky, Verena Krause, Zhuying Huang, Nancy R. Mendell, Howard Eichenbaum, and Deborah L. Levy

The capacity for transitive inference (TI), a form of relational memory organization, is impaired in schizophrenia patients. In order to disambiguate deficits in TI from the effects of ambiguous reinforcement history and novelty, 28 schizophrenia and 20 nonpsychiatric control subjects were tested on newly developed TI and non-TI tasks that were matched on these 2 variables. Schizophrenia patients performed significantly worse than controls on the TI task but were able to make equivalently difficult nontransitive judgments as well as controls. Neither novelty nor reinforcement ambiguity accounted for the selective deficit of the patients on the TI task. These findings implicate a disturbance in relational memory organization, likely subserved by hippocampal dysfunction, in the pathophysiology of schizophrenia.

Key words: relational memory/difficulty/hippocampus

Introduction

Of the many domains of cognitive functioning known to be impaired in schizophrenia, memory deficits are among the more pronounced. In particular, declarative (ie, explicit) memory, which requires conscious recall of events within a context, is compromised, whereas nondeclarative (eg, implicit) forms of memory, such as perceptual priming and procedural memory, remain relatively intact. Higher order memory dysfunction is observed even when impairments in other cognitive abilities (eg, attention, executive functioning, intelligence) and clinical factors (eg, positive symptoms, duration of illness, and medication effects) are statistically controlled.

The medial temporal lobe (MTL), particularly the hippocampus (HP), is a critical substrate for declarative/explicit memory. The HP is part of the larger neurocognitive network that subserves memory encoding, recognition, and retrieval of declarative information in humans and rodents.

Abnormalities in HP structure and function have been strongly linked to the pathophysiology of schizophrenia. Both postmortem and imaging studies document a wide range of structural abnormalities in HP. Schizophrenia patients also show morphological, functional, and/or biochemical abnormalities in other parts of MTL and in other brain regions that connect directly or indirectly with HP, including neocortex, thalamus, anterior cingulate, amygdala, and entorhinal and prefrontal cortices. Moreover, substantial evidence supports a disturbance in “functional connectivity” between HP and one or more of these regions. For example, HP shape deformities in schizophrenia are localized to regions of HP that send projections to prefrontal cortex. Defective modulation and underrecruitment of HP in schizophrenia are localized to the anterior portion of HP, a key site of projections between HP and prefrontal cortex. Further, schizophrenia patients show recruitment of prefrontal cortex during a variety of cognitive and sensory tasks, possibly to compensate for underrecruitment of HP, extraprefrontal regions (eg, middle temporal area), or other regions within prefrontal cortex (eg, Bonner-Jackson, Barch et al, Nagel et al, and Chen et al).

In addition to these structural and functional abnormalities, alterations of neural circuitry in the HP have been strongly implicated in schizophrenia. Specifically, N-methyl-d-aspartic acid receptor hypofunction, resulting in abnormal modulation of neural excitation within the HP, has been hypothesized to be a fundamental component of the pathophysiology underlying memory and other cognitive impairments in schizophrenia.

In light of the centrality of the HP in the pathophysiology of schizophrenia, behavioral probes of HP function that have been extensively studied in rodents are prime candidates for translational applications. The
HP is an essential part of the neural circuitry that sub-serves relational memory organization in animals and humans—relating different elements of experience to each other and making ad hoc rearrangements of them as needed. Relational memory is thought to depend on intact function of the HP, although a larger network that also includes prefrontal cortex, posterior parietal cortex, and midbrain regions underlies memory retrieval and relational reasoning.

Transitive inference (TI) is one form of relational memory organization that has been studied extensively in rodents using a task that has been adapted for use in human populations. Subjects are presented with a set of premises for pairs of items where the individual items in the premises overlap. One can assess the establishment of a relational memory representation that incorporates all the items by measuring the capacity for inferential judgments about items that are only indirectly related (ie, the judgment requires reference to an other, intermediate stimulus that is not present at the time of judgment). For example, if the premises are “Howard is smarter than Bob” and “Bob is smarter than Larry,” then the establishment of the relational hierarchy, Howard > Bob > Larry, is tested by assessing the capacity to make the TI that “Howard is smarter than Larry.”

In a previous study, we compared the performance of schizophrenia patients and controls on an adaptation of the Eichenbaum rodent TI paradigm. Subjects were trained to select the correct visual pattern in each of 4 overlapping pairs of abstract visual patterns that had been reinforced hierarchically (see figure 1, also available in color as online Supplementary Material). They were then presented with previously learned pairs as well as novel combinations. One novel combination (AE) could be solved successfully only by manipulating a hierarchy that required a TI. Both groups performed equivalently on the non-TI (AE) probe pair, but schizophrenia patients were significantly less accurate than controls on the TI (BD) probe pair.

The selectivity of impaired performance in schizophrenia patients on the TI probe suggested that compromised relational memory organization is a significant component of the cognitive deficit observed in schizophrenia. An equally plausible interpretation, however, is that the performance deficit in the TI condition reflected greater task difficulty (ie, ambiguous reinforcement history) and/or novelty effects. Novelty per se is an unlikely explanation because both AE and BD were novel pairings, but the groups differed only on BD. BD is more difficult than AE, however, because of its ambiguous reinforcement history. The reinforcement histories for B and D varied, depending on whether B was paired with A (never reinforced) or C (always reinforced) and whether D was paired with C (always reinforced) or E (never reinforced). Thus, unlike AE, BD cannot be solved based on reinforcement history alone. In our previous work, the poorer performance of schizophrenia patients than controls in the BD condition could not be conclusively attributed to a relational memory deficit because reinforcement ambiguity/difficulty and novelty were confounded with the demand for TI.

In order to disambiguate deficits in TI from effects of ambiguous reinforcement history (which serves as a proxy for difficulty) and novelty, we extensively modified the original TI task. In this new paradigm, TI and non-TI tasks were symmetrically designed to include unambiguous and ambiguous reinforcement conditions as well as familiar and novel conditions. In TI and non-TI tasks that are equated for difficulty and novelty in controls, selectively impaired performance on the TI task in schizophrenia patients would provide strong support for a differential deficit implicating relational memory organization and exclude reinforcement history/difficulty and novelty as parsimonious explanations for group differences in performance.

Methods

Subjects

The subject groups included 28 individuals who met Diagnostic and Statistical Manual of Mental Disorders (Fourth Edition) (DSM-IV) criteria for schizophrenia or schizoaffective disorder and 20 nonschizophrenic controls (NCs). Demographic characteristics of the sample are presented in table 1. The groups did not differ in age or parental socioeconomic status. Patients had significantly fewer years of education (P = .01) than controls. The patients were chronically ill outpatients (mean duration of illness: 16.0 ± 8.7 y), were moderately symptomatic
and had elevated levels of thought disorder as measured by the Thought Disorder Index (TDI)\textsuperscript{95,96} (mean total TDI score: 20.2 ± 14.4). All schizophrenia patients were medicated at the time of testing: 72% on atypical antipsychotics, 14% on typical antipsychotics, and 14% on both atypical and typical antipsychotics; mean dose in chlorpromazine equivalent: 768.6 ± 1316.8 (529.2 ± 367.3 excluding one subject who was on an unusually high dose of neuroleptics).\textsuperscript{97,98} NCs did not meet DSM-IV criteria for any psychotic disorder (lifetime), bipolar disorder without psychotic features, or a schizophrenia-spectrum personality disorder. The principal diagnostic instrument for assessing Axis I disorders was the Structured Clinical Interview for DSM-IV.\textsuperscript{99} Schizotypal, schizoid, and paranoid personality disorders were assessed in controls using the Structured Interview for Schizotypal Symptoms.\textsuperscript{100} An experienced clinician administered the interviews, and an independent group of senior diagnosticians reviewed the interview materials and all available hospital records and assigned consensus Axis I and Axis II diagnoses based on best estimate methods.\textsuperscript{101} The interviews and the diagnostic evaluations were performed blind to group membership and to the results of the experimental procedures. Subjects who are included in this study are a subset of a much larger group that also included relatives of schizophrenic, schizoaffective, and bipolar patients as well as relatives of controls. The following exclusion criteria applied to all participants: (a) lack of fluency in English, (b) history of serious head trauma or organic brain disease, and (c) history of substance abuse or dependence during the past 2 years or previous chronic dependence. All participants had an estimated verbal IQ of 85 or greater based on the vocabulary subtest of the Wechsler Adult Intelligence Scale-Revised.\textsuperscript{102} All participants provided written informed consent as per Institutional Review Board guidelines and were paid for their participation.

### Procedure

The TI and non-TI tasks (figures 2 and 3, also available in color as online Supplementary Material) were symmetrically designed in that the stimuli for each task consisted of 4 premise pairs of abstract visual patterns, 2 of which were low difficulty (ie, unambiguous reinforcement history) and 2 of which were high difficulty (ie, ambiguous reinforcement history). Stimuli were presented on a computer screen, one pair at a time. Each task had a training phase and a test phase.

#### Training Phase

During training, participants were instructed that one pattern in each pair would always hide a “smiley face” (ie, ☺) and that the task was to learn and remember which stimulus in each premise pair hid the smiley face. Subjects responded by pressing a button labeled “LEFT” (for the pattern on the left) or “RIGHT” (for the pattern on the right).

Sequential (TI) Series Training included 4 hierarchical sequential (ie, overlapping) premise pairs (A > B, B > C, C > D, D > E) (figure 2). The individual patterns were randomly assigned to positions (ie, A, B, C, D, E) within the series. Each of the low-difficulty, or “unambiguous,” premise pairs (AB, DE) contained an end-anchored stimulus that was unambiguously reinforced in that the smiley face was always under A and never under E. Neither of the 2 high-difficulty, or “ambiguous,” internal premise pairs (BC, CD) contained an unambiguously reinforced stimulus in that the smiley face was under C when C was paired with D but under B when C was paired with B.

### Table 1. Demographic and Clinical Characteristics of Schizophrenia Patients and Normal Controls\textsuperscript{a}

<table>
<thead>
<tr>
<th></th>
<th>Schizophrenia Patients (n = 28), Mean (SD)</th>
<th>Normal Controls (n = 20), Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (y)</td>
<td>41.9 (8.9)</td>
<td>41.5 (11.8)</td>
</tr>
<tr>
<td>Female/male</td>
<td>9/19</td>
<td>12/8</td>
</tr>
<tr>
<td>Socioeconomic status</td>
<td>71 (in social classes I–III)</td>
<td>90</td>
</tr>
<tr>
<td>Education (y)</td>
<td>13.3 (2.0)*</td>
<td>16.3 (2.4)</td>
</tr>
</tbody>
</table>

\textsuperscript{a}Includes only subjects who met learning criterion.

\textsuperscript{*}P < .01; all other group differences were not statistically significant.
Nonsequential (Non-TI) Series  Training for this series included 4 nonsequential (ie, nonoverlapping) premise pairs (F > G, H > I, J > K, L > M) (figure 3). Individual items are presented in figure 3. The 2 low-difficulty, or unambiguous, premise pairs had fully predictable reinforcement contingencies (100% reinforcement probability for F and J, 0% reinforcement for G and K). The 2 high-difficulty, or ambiguous, premise pairs had probabilistic reinforcement contingencies (75% reinforcement probability for H and L, 25% reinforcement for I and M). We chose the 75%/25% reinforcement ratio for the high-difficulty condition after demonstrating in 2 pilot studies that schizophrenia patients and NCs did not differ on 4 reinforcement schedules and that accuracy in the 75%/25% reinforcement condition was equivalent to accuracy on TI in NCs. (Specifically, we compared a 75%/25% reinforcement ratio group \([n = 18]\) with a 66%/34% reinforcement ratio group \([n = 20]\) in NCs. The 75%/25% reinforcement ratio group was better matched across the sequential/nonsequential, learned/novel, and unambiguous/ambiguous conditions than the 66%/34% reinforcement ratio group. We also compared 13 schizophrenia outpatients and 13 age-matched NCs who were demographically similar to the present sample on 4 reinforcement schedules [100%/0%, 83%/17%, 75%/25%, 66%/34%]. Schizophrenia patients and NCs did not differ in sensitivity to reinforcement history on any of the 4 reinforcement schedules as measured by accuracy on learned and novel pairs in this nonsequential series.)

Training for both the nonsequential and sequential series occurred independently in staged designs totaling 192 trials per series, ie, 48 trials for each premise pair. In each series, the first training block had 48 trials and was front loaded, ie, subjects received twice as many trials of 2 pairs of adjacent stimuli relative to the other 2 pairs of adjacent stimuli in that series. The second training block of each series also had 48 trials and was back loaded; the number of trials of adjacent stimuli was reversed relative to the first training block. The third training block had 96 trials and was balanced; subjects received an equal number of exposures to all pairs of stimuli. The staged training method was chosen because balanced training blocks in a sequential series resulted in hierarchical responding in fewer than 50% of controls, whereas staged training produced a better-than-chance likelihood that controls treated the stimuli hierarchically. Within each training block, order of presentation of the relevant pairs was randomized, and the left/right position of each stimulus within a pair was counterbalanced. In order to minimize the potential confound that might result from unpredictable manipulation of response contingencies in the nonsequential series on later combined novel and learned premise pair testing, we elected to train all subjects first on the nonsequential premise pairs.

Posttraining Tests. Immediately following the completion of training for each series, participants were first tested on the 4 learned premise pairs from that series. Participants were instructed to choose one pattern in each pair based on their memory of where the smiley face had appeared during training. Order of presentation was random and counterbalanced for left/right orientation.

Fig. 3. Nonsequential Series (Nontransitive Inference Task).

Non-Sequential Series (Non-Transitive Inference Task)

Learned Premise Pairs

<table>
<thead>
<tr>
<th>F &gt; G</th>
<th>H &gt; I</th>
<th>J &gt; K</th>
<th>L &gt; M</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
<td><img src="image3.png" alt="Image" /></td>
<td><img src="image4.png" alt="Image" /></td>
</tr>
</tbody>
</table>

Novel Probe Pairs

Low Difficulty

<table>
<thead>
<tr>
<th>F &gt; K</th>
<th>J &gt; G</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image5.png" alt="Image" /></td>
<td><img src="image6.png" alt="Image" /></td>
</tr>
</tbody>
</table>

High Difficulty

<table>
<thead>
<tr>
<th>H &gt; M</th>
<th>L &gt; I</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image7.png" alt="Image" /></td>
<td><img src="image8.png" alt="Image" /></td>
</tr>
</tbody>
</table>
of the individual stimuli. The tests consisted of 16 trials each of the unambiguous and ambiguous sequential premise pairs and 16 trials each of the unambiguous and ambiguous nonsequential premise pairs or 32 trials for each series.

Combined Novel and Learned Premise Pair Testing. Participants were then tested on both the learned premise pairs and on pairs involving novel combinations, ie, pairs that were not previously learned. In the sequential series, novel combinations included AC and CE (sequential novel unambiguous) and BD (sequential novel ambiguous); AE was not included in this modified paradigm because it is separated by 3 positions. In the nonsequential series, novel combinations included FK and GJ (nonsequential novel unambiguous) and HM and IL (nonsequential novel ambiguous). Participants were instructed that they would see combinations of patterns not previously seen and to choose the pattern in each pair based on what they had learned during training. Each of the 8 conditions, ie, sequential and nonsequential learned premise pairs and novel combinations, was presented 40 times across 4 blocks for a total of 320 test trials. Within each test block, order of presentation of the learned premise and novel pairs was randomized, and the left/right position of each stimulus within a pair was counterbalanced. Subjects received no feedback during either test. After the testing was completed, subjects were asked a single debriefing question to assess conscious awareness of the hierarchy. (How did you make your decisions?)

Subjects’ responses to this question were classified in terms of whether or not they articulated an overt awareness of the hierarchy for the sequential series.

Statistical Analysis
In order to separate the effects of group, task, novelty, and difficulty/reinforcement ambiguity, we performed a multivariate analysis of variance, using diagnosis (NCs, schizophrenia patients) as the between-group factor and task, novelty, and difficulty/reinforcement ambiguity as within-group factors. For each within-group factor, mean difference scores of the proportion of correct responses were used to evaluate group interactions (a between-group difference in difference scores indicates an interaction). Within-group factors were assessed for task (sequential-nonsequential conditions), novelty (learned-novel conditions), and difficulty/reinforcement ambiguity (unambiguous-ambiguous conditions) Two-tailed $P$ values are reported. Planned comparisons were carried out using $t$ tests and Wilcoxon signed rank tests as appropriate. Fisher exact tests were used to determine whether there were significant group differences in the proportions of above and below chance responding in each condition.

Results
Learning
Subjects were included in the analyses only if their performance on the original premise pairs during the first posttraining test was significantly better than chance. Comparable proportions of subjects in each group (32% of 41 schizophrenia patients and 31% of 29 NCs) did not meet the learning criterion. Accuracy scores of subjects who met the learning criterion (28 schizophrenia patients and 20 NCs) are presented in the top section of figure 4 (available as table 2 in the online Supplementary Material). The criterion for better-than-chance performance, according to binomial probabilities, was $\geq22$ correct responses in 32 trials (69%). (The base probability of a correct response is 50% on each of the 32 trials for a given type of premise pair. Based on the binomial theorem, the probability of correct guessing on $\geq 22/32$ trials is 0.03 [1 tailed]. That is, the probability that “good guessing” would result in $\geq 22/32$ correct responses by chance in any subject is $\leq 3\%$.) Significantly better-than-chance performance was required for the unambiguous sequential, ambiguous sequential, and unambiguous nonsequential premise pairs. (Because of the staged training design, the usual learning curves for the 3 training blocks are not informative. Despite the fact that all subjects included in the study learned to criterion, on the final balanced training block NCs achieved significantly higher accuracy levels than schizophrenia patients on the unambiguous nonsequential premise pairs [$t_{31.9} = 3.48$, $P = .002$; NCs: 99% ± 1.9%; schizophrenia patients: 94% ± 7.8%], the unambiguous sequential premise pairs [$t_{37.3} = 2.41$, $P = .02$; NCs: 98% ± 3.2%; schizophrenia patients: 94% ± 6.6%], and the ambiguous sequential premise pairs [$t_{37.3} = 4.30$, $P = .001$; NCs: 96% ± 3.5%; schizophrenia patients: 85% ± 4.7%]. The behavioral data for all 3 training blocks are available as table 3 in the online Supplementary Material.) Above-chance performance on the ambiguous nonsequential premise pairs was not required because the reinforcement contingencies had been manipulated to increase the difficulty level.

Disambiguating Novelty, Difficulty/Reinforcement Ambiguity, and Task
Across all subjects, main effects were obtained for novelty ($F_{1,47} = 14.77$, $P = .0004$) and for difficulty ($F_{1,47} = 14.27$, $P = .0004$). Subjects performed better on the previously learned premise pairs than on novel premise pairs and on unambiguous premise pairs than on ambiguous premise pairs. These main effects in the combined cohort were primarily a result of novelty ($F_{1,27} = 11.45$, $P = .002$) and difficulty ($F_{1,27} = 12.35$, $P = .002$) effects in schizophrenia patients, although there was also a trend toward a significant novelty effect ($F_{1,19} = 3.90$, $P = .06$) in NCs. The main effect of difficulty was not statistically significant.
in the NCs ($F_{1,19} = 2.65, P = .12$). Group-by-novelty ($t_{44.3} = 1.63, P = .10$) and group-by-difficulty ($t_{46} = 1.32, P = .19$) interactions were not statistically significant. There was no significant effect of task ($F_{1,47} = 0.01, P = .93$).

Post Hoc Analyses

Novelty. We further parsed the effects of novelty by examining its relation to task and difficulty in each group. Performance on unambiguous sequential and nonsequential pairs differed as a function of novelty. Schizophrenia patients performed worse on novel unambiguous pairs than on previously learned unambiguous pairs in both the sequential ($S = 68.0, P = .0004$) and nonsequential ($S = 185.0, P = .0001$) conditions. Significant novelty effects were observed in NCs only on the unambiguous nonsequential ($S = 104.0, P = .0001$) pairs (figure 5). In contrast, performance on ambiguous sequential and ambiguous nonsequential pairs did not differ as a function of novelty (all $P$ values > .28).

Difficulty/Reinforcement Ambiguity. We also examined difficulty in relation to task and novelty in each group. The effects of difficulty were observed for previously learned pairs (figure 6). Schizophrenia patients performed significantly worse on previously learned ambiguous pairs than on previously learned unambiguous pairs in both the sequential ($S = 84.0, P = .0009$) and nonsequential conditions ($S = 100.5, P = .0001$). Difficulty
effects were evident in NCs only on the previously learned nonsequential pairs ($S = 23.5, P = .02$). In contrast, in both groups, subjects performed equivalently on novel ambiguous pairs and novel unambiguous pairs regardless of task (all $P$ values $> .09$).

**Sequential Novel Ambiguous Pairs (TI).** The schizophrenia patient group had significantly worse accuracy ($t_{33.4} = 2.11, P = .02$, estimated effect size: 0.6) in the TI (BD) condition than NC subjects (figure 4 and table 2 of Supplementary Material). Moreover, a significantly ($P < .03$) larger proportion of schizophrenia patients (6/28, 21%) performed below chance than NCs (0/20) on these pairs, demonstrating a clear TI deficit in this subgroup of schizophrenia subjects.

**Nonsequential Novel Ambiguous Pairs (Non-TI).** The schizophrenia patient and NC groups performed equivalently on nonsequential novel ambiguous pairs ($t_{46} = 0.72, P = .47$, estimated effect size: 0.2), indicating that schizophrenia patients can make difficult nontransitive judgments. Consistent with this finding, the proportions of schizophrenia patients (5/28, 17.9%) and NCs (2/20, 10%) who performed below chance on these pairs did not differ ($P = .68$).

**Sequential Novel Ambiguous Pairs Vs Nonsequential Novel Ambiguous Pairs.** The novel ambiguous sequential and nonsequential pairs were matched for difficulty in that NCs performed them with equivalent accuracy (see also figure 4) ($t_{19} = -1.46, P = .16$, estimated effect size: 0.4). Performance on these tasks was not significantly correlated in NCs ($r = .24, n = 20, P = .3$) or schizophrenia patients ($r = .19, n = 28, P = .3$), however, indicating that the 2 tasks tap different abilities.

**Group-by-Task Interaction: Clarifying the Distinction Between Difficulty/Ambiguous Reinforcement and Capacity for TI.** Ambiguous reinforcement history was a feature of both the sequential and nonsequential novel conditions. The key distinction between the 2 conditions was that nontransitive judgments were required to perform accurately in the nonsequential novel condition, whereas transitive judgments were required to perform accurately in the sequential novel condition. Schizophrenia patients and NCs performed with equivalent accuracy in one ambiguously reinforced condition (nonsequential novel ambiguous), but schizophrenia patients performed significantly less accurately than NCs in the other, equivalently difficult, ambiguously reinforced condition (sequential novel ambiguous). This series of findings suggests that a disturbance in the ability to make a TI, rather than difficulty/ambiguous reinforcement per se, accounts for the selective deficit in the patients.

One way to further separate capacity for TI from effects of difficulty on the sequential novel ambiguous trials is to condition on above-chance performance on the nonsequential novel ambiguous trials. Stratifying in this way identifies subjects who were able to handle ambiguous reinforcement when no TI was involved and removes the effect of a subgroup of subjects for whom ambiguous reinforcement, independent of TI, was sufficient to impair performance. In order to evaluate an interaction with group, the difference in accuracy scores in the 2 novel ambiguous conditions (nonsequential-sequential) was used to compare performance as a function of task in schizophrenia patients ($n = 23$) and NCs ($n = 18$) who performed at above-chance levels on the nonsequential novel ambiguous trials. The analysis shows a strong trend for a group-by-task interaction ($t_{29.6} = 1.63, P = .055$; schizophrenia patients: $9.7\% \pm 29.5\%$; NCs: $-1.2\% \pm 11.2\%$; estimated effect size: 0.5), providing additional support for a differential deficit in TI in schizophrenia patients. Figure 7 shows the mean percent accuracy scores on nonsequential and sequential novel...
ambiguous pairs for these subgroups of schizophrenia and NC subjects.

Correlations Between Sequential Novel Ambiguous (TI) Performance and Clinical Demographic Variables

The only demographic or clinical variable that significantly correlated with TI performance in schizophrenia patients was thought disorder as measured by the TDI. The total TDI score was negatively and moderately correlated with the percentage of correct responses on sequential novel ambiguous pairs ($r = -.64$, $n = 28$, $P < .02$). The association between lower amounts of thought disorder and better TI capacity is supported by the correlation between total TDI score and the difference in accuracy scores in the 2 novel ambiguous conditions (sequential-nonsequential) in the subgroup of schizophrenia patients who could perform the nonsequential condition ($r = -.54$, $n = 23$, $P < .008$). Total TDI score was also negatively and moderately correlated with performance on sequential learned ambiguous pairs ($r = -.73$, $P < .05$), indicating that ability to master as well as to manipulate difficult hierarchical relationships is enhanced in the context of less disordered thinking. Notably, this relationship was present independent of novelty and reinforcement ambiguity. Performance on unambiguous sequential pairs was not significantly correlated with total TDI score independent of novelty ($P > .14$). The correlations between total TDI score and performance on nonsequential pairs were all nonsignificant (all $P > .15$).

Accuracy on sequential learned unambiguous pairs was inversely related to a particular category of thought disorder, combinatory thinking ($r = -.47$, $P < .01$). Combinatory thinking reflects the propensity to find relationships between unrelated things. Thus, patients who were less prone to create arbitrary relationships were more competent in handling straightforward relationships within a hierarchy. Combinatory thinking was not related to accuracy outside the context of a hierarchy (nonsequential conditions: all $P > .15$) or when manipulating within the context of a hierarchy (ie, BD; $P > .94$). Deviant verbalizations, a category of thought disorder that reflects idiosyncratic use of language, were not significantly correlated with performance on any of the sequential or nonsequential task conditions (all $P > .3$). The total TDI score was not significantly correlated with any clinical or demographic variables in schizophrenia patients ($r = -.2$, $P < .1$), and none of the other demographic and clinical variables were significantly correlated with TI performance (all $P > .46$).

A significantly larger proportion of NCs (7/14, 50%) than schizophrenia patients (4/26, 15%) ($P < .03$) expressed overt awareness of the hierarchy. Conscious awareness of the hierarchy was not significantly associated with TI performance in schizophrenia patients ($r = .09$, $n = 26$, $P = .7$) or in NCs ($r = .43$, $n = 14$, $P = .12$) (debriefing information available on a subset of the subjects).

Discussion

In this study, schizophrenia patients showed an impaired ability to make TIs. This finding could not be attributed to difficulty (ie, ambiguous reinforcement history) or to novelty. Novelty did not significantly worsen performance on pairs that had been ambiguously reinforced, including the key TI condition requiring manipulation of a hierarchy. Further, reinforcement ambiguity did not significantly impair performance on novel pairs, including the TI condition. These findings confirm our previous report using paradigms that provided improved experimental control over these 2 potentially confounding factors. The TI deficit implicates relational memory organization, which has clear links to hippocampal integrity as one component of the cognitive dysfunction in schizophrenia. The results are consistent with data showing that schizophrenics are also impaired on another task that is HP dependent, transverse patterning.

Both novelty and reinforcement history did significantly affect performance but not when TI was involved. Novelty worsened performance only on unambiguously reinforced items. Titone et al reported a similar finding for the AE condition relative to previously learned unambiguously reinforced sequential pairs. Consistent with the finding here, this effect was observed in patients but not in controls. Similarly, ambiguous reinforcement history worsened performance only on previously learned pairs, a finding also consistent with the results of Titone et al. Both subject groups were equivalently vulnerable to these effects on nonsequential trials. Only schizophrenia patients were susceptible to the effects of novelty and
ambiguous reinforcement history on sequential trials. The most likely explanation is that schizophrenia patients were less able than controls to make use of hierarchical relationships to enhance reinforcement history in making correct judgments. Thus, schizophrenia patients showed a decrement in performance both when manipulation of a hierarchy was required for correct judgments (the TI condition, BD) as well as when knowledge of a hierarchy could be used to complement ambiguous reinforcement history (previously learned ambiguous sequential pairs).

We used reinforcement ambiguity as a surrogate for task difficulty. The inference that difficulty per se is not a parsimonious explanation for TI deficits is also supported by 2 neuroimaging studies of nonpsychiatric controls. In a variation of the same paradigm used here and in a separate TI task, NCs showed selective activation of right anterior HP during transitive judgments but not during nontransitive judgments. Difficulty was not associated with HP activation in either study. The results of these imaging studies are consistent with the behavioral findings presented here showing that it is possible to distinguish difficulty and novelty effects from impaired TI. Moreover, the imaging studies provide support at the brain functional level that both paradigms are targeted behavioral probes that tap a key aspect of HP function, namely, the capacity for TI.

A generalized deficit is not a parsimonious explanation for the selective TI deficit in schizophrenia. In ambiguously reinforced tasks of equivalent difficulty, schizophrenia patients were able to make nontransitive judgments about novel pairs as well as NCs. Indeed, schizophrenia patients continued to perform worse than NCs on the TI condition even after conditioning on ability to make correct nontransitive judgments. Thus, difficult reinforcement contingencies per se are not a necessary condition for a selective deficit in TI. Impaired ability to respond to difficult reinforcement contingencies and/or generalized deficit may be sufficient to account for nonselective deficits. In a subgroup of schizophrenia patients on both novel transitive and novel nontransitive tasks, however. Notably, none of the NC subjects who performed below chance on ambiguously reinforced nontransitive novel pairs performed below chance on ambiguously reinforced transitive pairs.

The schizophrenia patient group was not uniformly impaired on the TI task. A relatively small subgroup (21%) of patients performed at lower accuracy than the worst performing control subject (see figure 8). Mean TI accuracy in this deviant subgroup (26.3%) was 72% lower than the mean of the NCs (95.4%), underscoring the profound nature of the impairment in these subjects. Indeed, all schizophrenia subjects in the deviant subgroup had accuracy scores more than 2 SD below the NC mean. Conversely, the estimated mean accuracy in the rest of the schizophrenia patient group was slightly better than that of NCs (96.9% vs 95.4%). Similar heterogeneity was described by Hanlon et al; the variance in performance among schizophrenia patients on a transverse patterning task was over 8 times that of the NCs.

The proportion of schizophrenia patients who showed a severe TI deficit is comparable to the proportion of schizophrenia patients who show marked HP volume reductions. In 2 structural magnetic resonance imaging studies, relatively small subgroups of schizophrenia patients (21% and 26%) had HP volumes that did not overlap with the range of HP volumes found in the control groups. In these 2 studies, 79% and 74% of the HP volumes of schizophrenia patients overlapped with those of NCs, respectively. Sim et al reported that the estimated mean total HP volume of the entire schizophrenia patient group was 10% smaller than that of the NCs. However, the estimated mean volume loss in the subgroup of schizophrenia patients with HP volumes outside the range observed in controls was actually 25%, 2.5 times greater than the estimate based on the entire patient sample. In the presence of such heterogeneity within the schizophrenia patient group, between-group comparisons of schizophrenia patients and NCs underestimate the magnitude of the change present in the deviant subgroup. Taken together, both the behavioral findings reported here and the structural indices of HP integrity by other investigators demonstrate marked changes in comparable proportions of schizophrenia patients. Whether HP volume reduction is related to behavioral deficits on tasks that depend on an intact HP cannot be addressed from these data.

It is not surprising that a disturbance in TI would be associated with schizophrenia. TI is a form of deductive reasoning. Delusional thinking and autistic logic, common symptoms of schizophrenia, involve failures of deductive reasoning. These symptoms are often state-related features of the disorder, although they can also persist when the acute symptoms of psychosis have subsided. The direction of the connection between relational
memory and these symptoms is not clear, however. For example, the hypervigilance associated with paranoia may be more likely to accompany intact capacity for TI, at least in nonacute states, whereas it would not be surprising if the delusional state itself interfered with TI.

We found an association between thought disorder, a core symptom of schizophrenia, and TI capacity. In addition, combinatory thinking was inversely related to ability to learn straightforward relationships within a hierarchy. In a previous study, we found that impaired relational interpretations on a conceptual combination task were associated with increased amounts of total thought disorder on the TDI. This finding suggested that a predisposition to make inappropriate relational interpretations may contribute to deficient inhibition of contextually irrelevant semantic interpretations. See also studies of Goldberg and Weinberger and Kerns and Berenbaum. The finding that the ability to learn simple relationships within a hierarchy would be impaired by a tendency to infer relationships between unrelated items is consistent with other data linking thought disorder to the HP. Several studies that used the BPRS to measure formal thought disorder (FTD) reported significant negative associations between the amount of FTD and decreased volume in the HP and related regions.

The issue of whether conscious awareness of hierarchy is necessary for correct TI judgments is unresolved. Consistent with previous findings in schizophrenia patients and NCs, conscious awareness was not significantly associated with ability to make a TI. In other studies, however, conscious hierarchical awareness was associated with TI. The possibility that a covert level of awareness is required cannot be ruled out but was not evaluated by asking subjects to order the pairs as part of the debriefing.

Although our data support the interpretation that the TI deficit in schizophrenia is related to hierarchical ordering and inferential processes, we cannot conclusively rule out 2 other possible interpretations. The first is that reinforcement history accounts for response selection during TI tasks in rodents and in humans. Our results are not consistent with this interpretation, but an additional test of this possibility required a 5-premise pair set. Pilot testing, however, showed that a 5-premise pair set was too difficult for controls and thus was unfeasible. The second alternative explanation is that encoded associations to overlapping stimuli in the sequential series may have led to generalizations based on “integrative encoding” at the time stimuli were learned rather than through inferences made at a later time. Both the novel sequential and nonsequential conditions do involve associative novelty, i.e., the detection of new arrangements of familiar stimuli. The novel nonsequential pairs and the novel unambiguous sequential pairs could be solved by associations to individual stimuli without reference to any stimulus that was not present at the time of judgment. Although novel ambiguous sequential judgments also required associations to individual stimuli, these judgments entailed an additional component—a comparison of these individual stimuli to another element that was not present at the time of judgment and thus was only indirectly related to the presented stimuli. In this latter comparison, inferential judgments are based on context-dependent relationships within a superordinate hierarchy. Which of these explanations has primacy cannot be determined from this study. Notably, both tasks involving associative novelty as well as those requiring a TI have been shown to selectively engage HP in rats and humans. Thus, regardless of whether the mechanism underlying performance in the ambiguous sequential condition is reinforcement history, integrative encoding, or TI, the group difference in performance between schizophrenia patients and controls implicates a disturbance in HP function.

Supplementary Material

Supplementary figures 1–3 and tables 2 and 3 are available at http://schizophreniabulletin.oxfordjournals.org

Acknowledgments

The authors thank Dr. Stephan Heckers for providing the raw data on hippocampal volume. The authors are grateful to Anne Gibbs for subject recruitment and to study participants for the dedication of their time and effort.

References


17. Whatmough C, Chertkow H, rCBF to the hippocampal complex covaries with superior semantic memory retrieval. Behav Brain Res. 2007;181:262–269.


47. Glantz LA, Lewis DA. Decreased dendritic spine density on prefrontal cortical pyramidal neurons in schizophrenia. *Arch Gen Psychiatry*. 2000;57:65–73.


120. Driscoll I, Sutherland RJ, Prusky GT, Rudy JW. Damage to the hippocampal formation does not disrupt representational flexibility as measured by a novelty transfer test. *Behav Neurosci.* 2004;118:1427–1432.

