Neurocognitive Endophenotypes for Bipolar Disorder Identified in Multiplex Multigenerational Families

David C. Glahn, PhD; Laura Almasy, PhD; Marcela Barguil, MA; Elizabeth Hare, PhD; Juan Manuel Peralta, MA; Jack W. Kent Jr, PhD; Albana Dassori, MD; Javier Contreras, MD; Adriana Pacheco, PhD; Nuria Lanzagorta, MA; Humberto Nicolini, MD; Henriette Raventós, MD; Michael A. Escamilla, MD

**Context:** Although genetic influences on bipolar disorder are well established, localization of genes that predispose to the illness has proven difficult. Given that genes predisposing to bipolar disorder may be transmitted without expression of the categorical clinical phenotype, a strategy for identifying risk genes is to identify and map quantitative intermediate phenotypes or endophenotypes.

**Objective:** To adjudicate neurocognitive endophenotypes for bipolar disorder.

**Design:** All participants underwent diagnostic interviews and comprehensive neurocognitive evaluations. Neurocognitive measures found to be heritable were entered into analyses designed to determine which test results are impaired in affected individuals, are sensitive to the genetic liability for the illness, and are genetically correlated with affection status.

**Setting:** Central valley of Costa Rica; Mexico City, Mexico; and San Antonio, Texas.

**Participants:** Seven hundred nine Latino individuals participated in the study. Of these, 660 were members of extended pedigrees with at least 2 siblings diagnosed as having bipolar disorder (n=230). The remaining subjects were community control subjects drawn from each site who did not have a personal or family history of bipolar disorder or schizophrenia.

**Main Outcome Measure:** Neurocognitive test performance.

**Results:** Two of the 22 neurocognitive variables were not significantly heritable and were excluded from subsequent analyses. Patients with bipolar disorder were impaired on 6 cognitive measures compared with nonrelated healthy controls. Nonbipolar first-degree relatives were impaired on 5 of these, and the following 3 tests were genetically correlated with affection status: Digit Symbol Coding Task, Object Delayed Response Task, and immediate facial memory.

**Conclusion:** This large-scale extended pedigree study of cognitive functioning in bipolar disorder identifies measures of processing speed, working memory, and declarative (facial) memory as candidate endophenotypes for bipolar disorder.

Arch Gen Psychiatry. 2010;67(2):168-177

Despite considerable evidence that the risk for bipolar disorder is inherited, the molecular genetic basis for this illness remains elusive. A recent genome-wide association analysis with more than 5000 cases and 6000 control subjects implicated 2 risk genes for the illness: ANK3 and CACNA1C. Because these genes regulate voltage-gated sodium and calcium channels, respectively, the findings suggest that ion channel dysfunction may play an important role in the pathophysiologic process of bipolar disorder. However, as noted by the authors, these genes have relatively small risk ratios, explaining little of the genetic contribution to the illness. Given the high heritability and familial relative risk of bipolar disorder, there is little doubt that additional genes are involved in the etiology of the illness. The identification of these genes is of paramount importance because they hold the potential for spurring novel treatments for this common and debilitating illness.

Given that genes predisposing to bipolar disorder may be transmitted without expression of the clinical phenotype, a strategy for identifying risk genes is the use of quantitative intermediate phenotypes or endophenotypes for the illness. A recent National Institute of Mental Health work group called for the widespread implementation of endophenotypic markers in the search for genes predisposing to bi-

Author Affiliations are listed at the end of this article.
polar disorder. However, relatively few endophenotypes have been proposed for the illness, and fewer still have been validated in large-scale studies.

Findings that asymptomatic individuals with bipolar disorder have neuropsychological impairments, that these deficits appear to be stable over time, and that unaffected first-degree relatives of bipolar probands have similar, although less pronounced, impairments suggest that neurocognitive measures may be candidate endophenotypes for the illness. The goal of the present study was to test this hypothesis.

Quantitative neurocognitive intermediate phenotypes can be assessed in affected and unaffected individuals, providing much greater statistical power to localize and identify disease-related genes than affection status alone. However, establishing a particular measure as an endophenotype requires that the trait be heritable, sensitive to affectation status, and genetically correlated with the illness. Although neurocognitive endophenotypes are appropriate for linkage and association experimental designs, the establishment of a particular trait as an intermediate phenotype is most efficient with large extended pedigrees because these families provide a single sample in which all of the criteria for effective endophenotypes can be assessed simultaneously (e.g., heritability, sensitivity to the illness, and genetic correlation).

In the present study, we conducted comprehensive neurocognitive assessments on large multigenerational pedigrees selected for sibling pairs concordant for bipolar disorder or schizoaffective disorder, bipolar subtype, and community controls. The focus on multiplex families reduces the potential of spurious or nongenetic forms of the illness and increases the potential that identified intermediate phenotypes reflect genetic factors that influence the risk of bipolar disorder. To determine whether a neurocognitive measure is a candidate endophenotype for bipolar disorder, we document its heritability, demonstrate that patients and their unaffected first-degree relatives are impaired on the measure, and establish genetic correlation between the measure and affection status.

### METHODS

#### PARTICIPANTS

Seven hundred nine Latino individuals from the central valley of Costa Rica (328 individuals [46.3%]); Mexico City, Mexico (139 [19.6%]); and San Antonio, Texas (242 [34.1%]), participated in the study. Of these, 660 were members of 45 families (mean [SD] family size, 14.67 [10.56] members; range, 3-37 members) with at least 2 siblings diagnosed as having bipolar disorder. Probands were recruited through systematic screening of outpatient and inpatient facilities. Inclusion and exclusion criteria were the same across sites and required a previous diagnosis of bipolar I disorder in probands and at least 1 sibling with bipolar disorder or schizoaffective disorder, bipolar type. Affected individuals were excluded if they did not provide written consent to contact family members or had a history of mental retardation, a neurological disorder, or severe head trauma. Once an affected sibling pair provided informed consent, attempts were made to recruit all first-, second-, and third-degree relatives. Inclusion and exclusion criteria were identical for family members, with the exception of requiring a personal history of bipolar disorder. The remaining subjects were ethnically matched community controls drawn in equal numbers from each site who did not have a personal or family history (first-degree relatives) of bipolar disorder or schizophrenia. Community controls were recruited through advertisements in local newspapers and via fliers placed in medical clinics and had the same inclusion and exclusion criteria as family members.

All subjects provided written informed consent on forms in the language of their choice (Spanish or English) approved by the review boards of all participating study sites.

### DIAGNOSTIC ASSESSMENT

All participants, regardless of diagnostic or family status, received the Diagnostic Interview for Genetic Studies and the Family Interview for Genetic Studies. Interviews were conducted by psychiatrists with established reliability (κ = 0.85). Final diagnoses were determined through a best estimation process in which 2 psychiatrists review available records, arrive at independent DSM-IV diagnoses, and reach a consensus. If consensus could not be reached, a third best estimator reviewed the case independently (this occurred once in this sample).

### NEUROCOGNITIVE ASSESSMENT

Each participant received the South Texas Assessment of Neurocognition, a 90-minute neuropsychological evaluation consisting of standard and computerized measures. Tests were selected on evidence of heritability, sensitivity to bipolar disorder, and minimization of the effects of language or the availability of parallel English and Spanish forms. Instructions for computerized neuropsychological tests were translated to Spanish by bilingual psychologists and translated back into English by professional translators. Other tests (eg, the California Verbal Learning Test [CVLT]) have published English and Spanish editions. All subjects received the same battery of tests in the language of their choice and in a fixed order and were allowed breaks as needed.

The South Texas Assessment of Neurocognition battery included 16 tests with 21 separate measures (Table 1). The computerized Digit Symbol Coding Task required subjects to indicate, via button press, if a centrally presented digit-symbol pair was identical to 1 of the 9 digit-symbol pairs in the reference list at the top of the screen. During the identical pairs version of the Continuous Performance Test, subjects viewed a series of numbers presented briefly on the screen, pressing the space bar whenever the same number appeared twice in a row. During the manual Stroop Test, subjects pressed arrow keys that faced in the same direction (congruent) or in the opposite direction (incongruent) as an arrow on the screen. The Spatial Delayed Response Task involved subjects remembering where a set of circles appeared and indicating, via button press, if a single circle was presented in the same location. During the Object Delayed Response Task, subjects remembered 3 abstract shapes (presented serially) and, after a delay, selected 1 of 4 shapes not previously presented. The Penn Face Memory Test involved the presentation of 20 target faces and then immediate and delayed recognition (indicated by button press) of these faces mixed with an equal number of foils. During the Penn Conditional Exclusion Test, subjects chose which stimulus of 4 did not belong, based on inferred rules, and were provided feedback. In addition, the variables derived from the South Texas Assessment of Neurocognition battery were en-
QUANTITATIVE GENETIC ANALYSES

All analyses were conducted with the Sequential Oligogenic Linkage Analysis Routines (SOLAR) computer program, which uses maximum likelihood variance decomposition methods to determine the relative importance of genetic and environmental influences on a measure by modeling the covariance among family members as a function of genetic proximity (kinship). The polygenic model assumes that a trait is the result of multiple gene loci at unidentified positions in the genome. In the case of a dichotomous trait (e.g., affection status), these loci are thought to contribute to an underlying liability that results in the disorder if a threshold is reached. To ensure that neuropsychological traits conform to the assumptions of normality, an inverse normal transformation was applied.

HERITABILITY ANALYSES

Heritability ($h^2$) represents the portion of the phenotypic variance accounted for by the total additive genetic variance ($h^2 = \sigma^2_a/\sigma^2_g$). Genetic variance ($\sigma^2_a$) is the portion of phenotypic variance due to pedigree relationships rather than environmental factors or error. Indices with stronger covariance between genetically more similar individuals have higher heritability. Within the SOLAR program, this is assessed by contrasting the observed covariance matrices for a trait with the covariance matrix predicted by kinship. Heritability analyses were conducted twice, without demographic covariates and again with age, sex, age $\times$ sex interaction, age$^2$, age$^2 \times$ sex interaction, education level (in years), and location of assessment included as covariates in the model. Although the second $h^2$ estimates are critical for the current investigation, the raw estimates were included for comparability with published data. Only neuropsychological variables with significant heritability, corrected for multiple comparison at a 5% false discovery rate (FDR), were included in subsequent analyses.

Table 1. Neurocognitive Dependent Measures

<table>
<thead>
<tr>
<th>Neurocognitive Task</th>
<th>Dependent Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Letter fluency$^{19}$</td>
<td>No. of words, beginning with specific letter, generated in 60 s</td>
</tr>
<tr>
<td>Semantic fluency$^{19}$</td>
<td>No. of animal names generated in 60 s</td>
</tr>
<tr>
<td>Digit Symbol Coding Task$^{12}$</td>
<td>No. of correctly identified digit-symbol pairs in 90 s</td>
</tr>
<tr>
<td>Trail Making Test, part A$^{12}$</td>
<td>Time needed to connect letters in ascending order</td>
</tr>
<tr>
<td>CPT hits$^{12}$</td>
<td>No. of hits during 6-min identical pairs CPT</td>
</tr>
<tr>
<td>CPT catch trials$^{12}$</td>
<td>No. of false alarms during 6-min identical pairs CPT</td>
</tr>
<tr>
<td>Digit span forward$^{20}$</td>
<td>No. of correctly recalled digit strings in their original order of presentation</td>
</tr>
<tr>
<td>Digit span backward$^{20}$</td>
<td>Reaction time difference between incongruent and congruent trials of manual Stroop Test</td>
</tr>
<tr>
<td>Letter-number span$^{22}$</td>
<td>No. of correctly recalled number-letter strings in numerical and alphabetical order</td>
</tr>
<tr>
<td>Spatial Delayed Response Task$^{23}$</td>
<td>No. of correct responses on simple test</td>
</tr>
<tr>
<td>Object Delayed Response Task$^{24}$</td>
<td>No. of correct trials in object-delayed nonmatch to sample task</td>
</tr>
<tr>
<td>CVLT learning$^{15}$</td>
<td>No. of items recalled during 5 exposures of 16-word list of CVLT</td>
</tr>
<tr>
<td>CVLT semantic cluster$^{17}$</td>
<td>Proximal recall of semantically related list items during 5 repeated exposures of 16-word list</td>
</tr>
<tr>
<td>CVLT delayed recall</td>
<td>No. of list items recalled after 20-min delay</td>
</tr>
<tr>
<td>Digit symbol recall$^{13}$</td>
<td>No. of digits recalled when presented with corresponding symbols from the Digit Symbol Coding Task</td>
</tr>
<tr>
<td>Immediate facial memory$^{13}$</td>
<td>No. of faces recognized during initial condition of Penn Face Memory Test</td>
</tr>
<tr>
<td>Delayed facial memory$^{25}$</td>
<td>No. of faces recognized during 20-min delay condition of Penn Face Memory Test</td>
</tr>
<tr>
<td>Matrix reasoning$^{26}$</td>
<td>No. of correctly completed progressive matrices</td>
</tr>
<tr>
<td>PCET$^{27}$</td>
<td>No. of correctly matched shapes on PCET</td>
</tr>
<tr>
<td>Trail Making Test, part B$^{14}$</td>
<td>Time needed to connect alternating letters and numbers</td>
</tr>
</tbody>
</table>

Abbreviations: CPT, Continuous Performance Test; CVLT, California Verbal Learning Test; PCET, Penn Conditional Exclusion Test.

SENSITIVITY TO LIABILITY FOR BIPOLAR DISORDER

Heritable neuropsychological variables were examined to determine which measures differentiate individuals with bipolar disorder from unrelated, unaffected individuals. Within the SOLAR program, these analyses were conducted by including a variable that identified individuals in specific liability groups (e.g., bipolar vs unrelated) and testing whether the mean of these groups differed within the context of the known pedigree structure and covariates ($\chi^2$ test). Evidence that a particular neurocognitive measure was sensitive to bipolar disorder was provided if the grouping variable explained a significant proportion of trait variance at 5% FDR. Effect size estimates (standardized mean differences) were derived using the following equation:

$$\text{Effect Size} = \left(\frac{4 \times \chi^2}{(n-1) \times \chi^2} \right)^{-1/2},$$

where $n$ is the number of individuals included in the analyses. Given the nonindependence of participants, this effect size estimate is conservative.

Neurocognitive measures sensitive to bipolar disorder were examined in analogous analyses in unaffected (nonbipolar) first-degree relatives to confirm that the candidate endophenotype reflected underlying genetic vulnerability. An individual’s liability for bipolar disorder was defined as the shortest genetic distance to an affected individual.

BIVARIATE ANALYSES

Although differences in neurocognitive measures between unaffected first-degree relatives and controls that rule out effects secondary to disease in the affected individuals are indicative of a genetic vulnerability, they could also be due to environmental factors shared by individuals with bipolar disorder and their unaffected relatives. To determine whether neurocognitive performance and liability for bipolar disorder have common genetic or environmental influences, mixed discrete/continuous trait bivariate analyses were conducted. These analyses decompose phenotypic correlations into genetic ($p_g$) and environmental ($p_e$) terms.
and environmental ($p_e$) correlations between the 2 traits. If the genetic correlation is significantly different from 0, then the traits are considered to be influenced by the same genetic factors (eg, $\geq 1$ genes influence both traits, or pleiotropy). If the environmental correlation is significantly different from 0, then the traits are considered to be influenced by the same environmental factors (eg, education level). The significance (5% FDR) of these correlations was tested by comparing the natural log (ln) likelihood for 2 restricted models (with $p_g$ or $p_e$ constrained to equal 0.0) against the ln likelihood for the model in which these parameters were estimated.

Finally, we performed multivariate analyses on those neurocognitive traits genetically correlated with bipolar disorder to determine whether these traits represent independent risk factors.

### RESULTS

#### SAMPLE CHARACTERISTICS

Of the 660 subjects from extended pedigrees, 230 had a best-estimate consensus DSM-IV diagnosis of bipolar disorder (type I, 161 subjects; type II, 51; and not otherwise specified, 6) or schizoaffective disorder, bipolar subtype (12 subjects), and are considered part of a broad bipolar phenotype (Table 2). Six individuals were diagnosed as having schizophrenia or schizoaffective depressive subtype and were excluded from all analyses. Among family members without major psychosis, 243 were unaffected (nonbipolar spectrum) first degree, 86 were unaffected second-degree, and 42 were unaffected third-degree relatives of affected individuals. One hundred eight subjects were not biologically related to affected individuals and were used to form an unrelated control sample; 59 were subjects who had married into selected families and 49 were community controls. Liability groups differed in average age ($F_{3,704} = 34.57$ [P < .001]) and in sex distributions ($\chi^2 = 7.15$ [P = .05]) but not in education level ($F_{3,704} = 0.89$ [P = .44]) (Table 2).

Individuals with the broad bipolar phenotype had significantly higher rates of anxiety disorders ($\chi^2 = 32.34$ [P < .001]) and past alcohol abuse/dependence ($\chi^2 = 63.52$ [P < .001]) than did their nonbipolar family members and unrelated participants. Among these individuals, 129 (36.1%) were prescribed psychotropic medications (some of them, multiple medications) at the time of assessment: 66 (28.7%) were prescribed antidepressants, 16 (7.0%) were prescribed lithium, 41 (17.8%) were prescribed mood stabilizers, 30 (13.0%) were prescribed antipsychotics, 57 (24.8%) were prescribed sedatives, 37 (16.1%) were prescribed atypical antipsychotics, 23 (10.0%) were prescribed typical antipsychotics, and 11 (4.8%) were taking stimulants.

#### HERITABILITY

Estimated heritability for 22 neurocognitive variables are presented in Table 3. Although demographic covariates were fixed for all analyses, covariates that reached significance for a specific neurocognitive measure are included in the far right column. Two measures failed to reach significant levels of heritability: CVLT semantic clustering and CVLT delayed recall. Performance on these measures was significantly correlated with education and differed by location. When these covariates were omitted from analyses, these indices were significantly heritable.

#### LOCATION OF ASSESSMENT

Location of assessment significantly influenced performance on many neurocognitive measures (Table 3). Examining differences between individual sites indicated that the effect sizes for these differences were generally small. Indeed, the average effect size for neurocognitive performance differences between San Antonio and Costa Rica was 0.32; between San Antonio and Mexico City, 0.20; and between Mexico City and Costa Rica, 0.17. In contrast, the effect size for formal education was 0.61. Although performance on neurocognitive tests is influenced by myriad factors, educational attainment, which varies significantly across countries, has been shown to have a more profound effect on test performance than other factors. This pattern of results is consistent with our prior work in these locations.

#### SENSITIVITY TO LIABILITY FOR BIPOLAR DISORDER

Results from analyses determining the sensitivity of neurocognitive measures to liability for bipolar disorder are given in Table 4. Individuals with the broad bipolar phenotype were statistically impaired on 6 of the 20 heritable cognitive measures compared with unrelated, unaffected subjects after controlling for FDR. To examine diagnostic specificity, analyses were repeated after constraining the affected group to patients with bipolar I disorder. These patients were impaired on all of the measures identified in the broad phenotype group and also on semantic fluency.
Given that individuals with bipolar disorder have increased rates of anxiety and alcohol use disorders (Table 2), it is unclear whether neurocognitive impairments are due to these co-occurring illnesses or to bipolar disorder per se. Hence, analyses were repeated with lifetime history of anxiety disorders and alcoholism in-
BIVARIATE ANALYSES

Genetic and environmental correlations performed on neurocognitive measures impaired in bipolar disorder are given in Table 5 and in the Figure. Of these measures, the Digit Symbol Coding Task, Object Delayed Response Task, and immediate facial memory were significantly and negatively correlated with affection status, indicating that poorer cognitive performance on these measures is related to an increasing genetic risk of bipolar disorder. None of the environmental correlations were significant, given the FDR. Similar correlations were observed when affection status was defined as bipolar I disorder (\( \rho_g \) [SE] for Digit Symbol Coding Task, \(-0.51 \pm 0.21\), \( P = .04\); \( \rho_g \) [SE] for Object Delayed Response Task, \(-0.53 \pm 0.17\), \( P = .009\); and \( \rho_g \) [SE] for immediate facial memory, \(-0.44 \pm 0.19\), \( P = .03\)).

In the full sample, performance on the Digit Symbol Coding Task was significantly genetically correlated with that of the Object Delayed Response Task (\( \rho_g \) [SE], \(0.78 \pm 0.09\); \( P = .003\)) and immediate facial memory (\(0.59 \pm 0.16\); \( P = .001\)). Similarly, the Object Delayed Response Task performance was significantly correlated with that of immediate facial memory (\(0.62 \pm 0.12\); \( P < .001\)).

Table 5. Bivariate Analysis Between Affection Status and Neurocognitive Measures

<table>
<thead>
<tr>
<th>Trait</th>
<th>Genetic Correlation</th>
<th>Environmental Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \rho_g ) (SE)</td>
<td>( P ) Value</td>
</tr>
<tr>
<td>Semantic fluency</td>
<td>0.22 (0.27)</td>
<td>.40</td>
</tr>
<tr>
<td>Digit Symbol Coding Task</td>
<td>-0.70 (0.18)</td>
<td>.002</td>
</tr>
<tr>
<td>Letter-number span</td>
<td>-0.16 (0.20)</td>
<td>.44</td>
</tr>
<tr>
<td>Object Delayed Response Task</td>
<td>-0.65 (0.16)</td>
<td>.002</td>
</tr>
<tr>
<td>CVLT learning</td>
<td>-0.47 (0.30)</td>
<td>.16</td>
</tr>
<tr>
<td>Immediate facial memory</td>
<td>-0.57 (0.18)</td>
<td>.004</td>
</tr>
<tr>
<td>Delayed facial memory</td>
<td>-0.40 (0.19)</td>
<td>.07</td>
</tr>
</tbody>
</table>

Abbreviation: CVLT, California Verbal Learning Test.

a Significant correlations, after controlling the false discovery rate, are given in boldface type.

MEASURES OF PROCESSING SPEED, WORKING MEMORY, AND DECLARATIVE (FACIAL) MEMORY ARE CANDIDATE ENDOPHENOTYPES FOR BIPOLAR DISORDER.

Measures of processing speed, working memory, and declarative (facial) memory are candidate endophenotypes for bipolar disorder. Each of these measures was heritable, impaired in individuals with the illness and their nonbipolar relatives, and genetically correlated with affection status. Initial evidence for the specificity of these neurocognitive measures for liability to bipolar disorder was provided because the index of generalized cognitive functioning was not impaired in the illness. Although a number of investigators have demonstrated that healthy first-degree relatives of bipolar probands have memory or executive functioning impairments, this is, to our knowledge, the first large-scale family-based study to provide evidence that the neurocognitive deficits found in bipolar disorder are linked to genetic liability for the...
illness. Although there are various definitions for intermediate phenotypes or endophenotypes, there is universal agreement that these measures must be associated with genetic liability for the illness under investigation. Thus, an implicit assumption about endophenotypes is that the same genes that convey risk for disease also influence the endophenotype. More formally, there is an assumption of pleiotropy. Pleiotropy can be demonstrated through estimation of genetic correlation.31 Unfortunately, in nontwin designs, genetic correlation requires relatively large samples of related individuals (range, >300 to 600 persons) and, hence, is rarely applied in psychiatric genetics. However, without formally demonstrating pleiotropy, one of the primary criteria of an endophenotype is left untested. Each of the candidate endophenotypes identified in this study meet the necessary and sufficient criteria for a viable intermediate phenotype, providing testable hypotheses about brain systems implicated in the pathophysiologic process of bipolar disorder.

Performance on the Digit Symbol Coding Task is an index of processing speed or of the time needed to execute simple cognitive operations. Although processing speed has not been localized to a single brain region, cognitive neuroscience emphasizes the integration of information across spatially distinct brain regions.29,35 suggesting that cognitive slowing as indexed by the processing rate is related to neuronal efficiency.30 Several investigators have reported that euthymic individuals with bipolar disorder have moderate to severe digit-symbol coding impairments.6,37,38 Although processing speed measures such as the Digit Symbol Coding Task may be negatively influenced by psychotropic medications, a recent meta-analysis in schizophrenia suggested that medication status had little effect on patient performance.30 Furthermore, at least 2 previous studies have reported poor Digit Symbol Coding Task performance in unaffected family members of probands with bipolar disorder.39,40 Our results further these findings by showing a significant negative genetic correlation between bipolar disorder and Digit Symbol Coding Task performance.

The Object Delayed Response Task24 includes the maintenance and manipulation of memory for complex visual objects and was modeled after measures used in nonhuman primates.42 Functional magnetic resonance imaging experiments indicate that task performance engages a complex network of brain areas associated with working memory (eg, prefrontal, temporal, and parietal regions).43 Patients with bipolar disorder showed relative hyperactivation in the right prefrontal and anterior cingulate regions and relative hypoactivation in medial temporal and visual processing regions compared with healthy subjects.43 Although relatively few investigators have used delayed response tasks in bipolar disorder, working memory deficits—particularly when manipulation of information is critical—have been consistently reported in patients who are in remission.39 In contrast, there is little prior evidence of working memory dysfunction in unaffected relatives of patients with bipolar disorder.8 The present findings suggest that the application of working memory specifically designed for use in affective disorders could improve the sensitivity of these measures to liability for bipolar disorder.44 Nonetheless, the present findings should be confirmed in an independent sample.

Declarative memory impairments are among the most commonly reported cognitive deficits in bipolar disorder. Although most investigators report verbal memory impairments, visuospatial and facial memory deficits have also been observed in euthymic patients with bipolar disorder.5,8 Indeed, in an early investigation of monozygotic twin-pairs discordant for bipolar disorder, Gourovitch and colleagues45 reported facial memory impairment in unaffected co-twins of bipolar probands. In the present study, verbal declarative memory was assessed with the CVLT.17,18 Although CVLT performance was heritable and impaired in individuals with bipolar disorder, nonbipolar first-degree relatives were not statistically different from healthy subjects. Our findings are similar to those reported in a recent meta-analysis of cognitive performance in unaffected first-degree relatives of bipolar probands.5 Although the present study nominates a facial declarative memory task as a candidate endophenotype for bipolar disorder, verbal declarative memory measures cannot be excluded and may represent important risk factors for the illness. More generally, although findings from the present study are encouraging, they are limited by the families studied and the neurocognitive tests used. Studies in different samples and with overlapping test batteries are warranted.

Significant heritability estimates indicate that genetic factors contribute substantially to the variation in the phenotype. However, heritability estimates do not provide information concerning the underlying genetic architecture of a phenotype and are subject to a number of assumptions.46 Indeed, phenotypes with high heritability estimates are not necessarily influenced by fewer genes or by genes with larger effects. For example, normal variation in adult height is highly heritable (h²=0.89-0.93),47 but current estimates suggest that up to 44 independent loci are associated with normal stature.38,49 In contrast, phenotypes with somewhat lower heritability estimates may have less complex genetic architectures. For example, the estimated heritability of the neuregulin 1 transcript was 0.50, but linkage analysis indicated a single locus (logarithm of odds =15.8) on chromosome 8.49,50 Thus, although an intermediate phenotype should be heritable, the strength of the heritability estimate may be less critical. Heritability estimates for bipolar disorder are typically higher than those reported herein for neurocognitive traits. However, it remains to be seen whether the genetic architectures of these neurocognitive endophenotypes are less complex than for the illness itself. Furthermore, simulation studies indicate that quantitative endophenotypes should be significantly more powerful than qualitative diagnoses for identifying genes.52

Endophenotypes can inform psychiatric nosology. Although bipolar disorder and schizophrenia are considered discrete psychiatric illnesses with unique etiologies, growing evidence suggests that the illnesses may have some common genetic roots.53 Each of the neurocognitive endophenotypes nominated for bipolar disorder have also been found in schizophrenia. Indeed, the Digit Symbol Coding Task applied herein was the most sensitive cognitive
measure of genetic liability for schizophrenia in our extended pedigree study.13 Similarly, the facial declarative memory measure has previously been found to be sensitive to risk of schizophrenia.54 Although our Object Delayed Response Task24 has not been used in schizophrenia family studies, similar measures have been proposed as intermediate phenotypes for the illness.54 Although the Spatial Delayed Response Task applied in the present study was previously shown to be sensitive to the genetic liability for schizophrenia,13,23 individuals with the broad bipolar phenotype were not impaired on this measure, implying some level of diagnostic specificity. However, we have shown in 2 separate samples that patients with bipolar disorder and a lifetime history of psychotic symptoms are impaired on this test, whereas patients without formal psychosis are not,44,55 suggesting that the manifestation of psychosis may be linked to spatial working memory deficits rather than to a specific diagnostic category. It is possible that by using a broad bipolar phenotype, in which few individuals manifest psychosis, we obscured deficits on this test. Indeed, when affection status was limited to bipolar I disorder alone, impairments on this test were more pronounced, though still below the FDR requirement. These findings raise questions about the specificity of endophenotypes for bipolar disorder alone rather than for psychotic illness more generally. However, it is unlikely that these questions will be fully addressed until the genetic factors that influence these neurocognitive endophenotypes and bipolar disorder are identified and found to confer risk for schizophrenia.

A number of the neurocognitive measures impaired in individuals with bipolar disorder were not impaired in non-bipolar first-degree relatives. Performance on these measures may be sensitive to the environmental factors necessary for the diathesis of the illness, and thus impairment in affected individuals is not associated with the genetic risk of bipolar disorder. Alternately, these tests may be influenced by affective or psychotic symptoms, psychotropic medication, sleep disturbances, or other sequelae of the illness. Indeed, current symptoms and psychotropic medications used to treat bipolar disorder have been shown to impair performance on specific neurocognitive tests.6,56 Although 56.1% of the patients included in the study were receiving medication at the time of assessment, inclusion of a single covariate coding use of psychotropic medication did not significantly alter the observed results. Although this analysis was coarse and did not account for potentially important factors (eg, lifetime history of medication use, specific class of medication), each of the proposed endophenotypes was also impaired in nonsymptomatic and medication-naïve, unaffected first-degree relatives. Hence, it is unlikely that the candidate endophenotypes identified herein are explained by patient mood symptoms or medication use.

Although the pattern of results did not significantly differ when affection status was constrained to bipolar I disorder, it is difficult to make conclusions about diagnostic specificity from this sample. Seventy percent of the affected individuals had bipolar I disorder, and those individuals with bipolar II disorder, bipolar disorder not otherwise specified, or schizoaffective disorder were related to those with bipolar I disorder. To draw inferences about the utility of these cognitive endophenotypes for the genetic investigation of non–bipolar type I illnesses, families with these illnesses who do not have bipolar I disorder should be studied. Similar conclusions can be drawn about psychiatric comorbidities such as alcohol abuse and anxiety disorders.

A potential advantage of neurocognitive endophenotypes is that they may be less genetically complex than psychiatric diagnoses.12,57 Although this potential has yet to be empirically demonstrated,58 however, because multiple neurocognitive endophenotypes can be studied simultaneously, they may afford an additional advantage over diagnostic classification, regardless of their genetic complexity. By focusing on the gene or genes common to all studied intermediate phenotypes, the number of candidate genes nominated for subsequent validation or functional genomic studies could be significantly reduced. However, such a strategy is valid only when significant genetic correlation between endophenotypes can be established (eg, pleiotropy). In the present study, the candidate endophenotypes were genetically correlated with each other and with bipolar disorder. This implies that these endophenotypes may represent a single genetic pathway for the illness and thus may facilitate this strategy prioritizing genes for additional study.

As demonstrated here and elsewhere,13,54 computerized neurocognitive measures can be efficiently and reliably administered to large numbers of individuals, a requirement for effective endophenotypes.59 Indeed, each of the tests sensitive to risk of bipolar disorder was developed specifically for large-scale studies of psychopathology. These tests are being applied in a multisite family-based linkage study of bipolar disorder and in a molecular genetics study of randomly ascertained individuals in large extended pedigrees. Together, these studies should provide clues to the genetic architecture of the neurocognitive endophenotypes identified in the present study and should facilitate the localization and identification of specific genes that contribute to these endophenotypes. These genes, in turn, should be examined for association with bipolar disorder. This iterative process should provide a window into the causal neurobiological pathways involved in the illness and the potential to develop biomarkers for psychiatric disorders.

Submitted for Publication: December 15, 2008; final revision received June 16, 2009; accepted June 19, 2009.

Author Affiliations: Olin Neuropsychiatry Research Center, Institute of Living, Hartford, Connecticut (Dr Glahn); Department of Psychiatry, Yale University School of Medicine, New Haven, Connecticut (Dr Glahn); Departments of Psychiatry (Dr Glahn, Hare, Dassori, Contreras, and Escamilla) and Ms Barguil) and Cellular and Structural Biology (Dr Escamilla), The University of Texas Health Science Center at San Antonio, and Department of Genetics, Southwest Foundation for Biomedical Research (Drs Almasy and Kent and Mr Peralta), San Antonio, Centro de Investigación en Biología Molecular y Celular, Universidad de Costa Rica, San José, Costa Rica (Ms Barguil and Drs Contreras, Pacheco, and Raventós); and Grupo de Estudios Médicos y Familiares Carracci, Mexico City, Mexico (Ms Lanzagorta and Dr Nicolini).
References


