Schizophrenia Spectrum & Othee Psychotic Disorders (Case Study).

(DISCUSSION 3)

Prior to beginning work on this discussion, read the assigned chapters from the text. It is highly recommended that you review each of the brief Blumenfeld (2012) video clips demonstrating the administration of a mental status examination. These are listed in the recommended resources and may require that you download Quicktime in order to view them. Although not required, these videos show the administration of a mental status exam and may prove helpful in this discussion.

Access the Barnhill (2014) *DSM-5 Clinical Cases* e-book in the DSM-5 library, and select one of the case studies. The case study you select must be one in which the client could be assessed using one or more of the assessment instruments discussed in this week's reading.

For this discussion, you will take on the role of a psychology intern at a mental health facility working under the supervision of a licensed psychologist. In this role, you will conduct a psychological evaluation of a client referred to you for a second opinion using valid psychological tests and assessment procedures. The case study you select from the textbook will serve as the information provided to you from the professional who previously evaluated the client (e.g., the psychologist or psychiatrist).

In your initial post, begin with a paragraph briefly summarizing the main information about the case you selected. Evaluate and describe the ethical and professional interpretation of any assessment information presented in the case study. Devise an assessment battery for a psychological evaluation that minimally includes a clinical interview, mental status exam, intellectual assessment, observations of the client, and at least two assessment instruments specific to the diagnostic impressions (e.g., attention deficit/hyperactivity disorder, post-traumatic stress disorder, autism spectrum disorder, etc.). The assessment battery must include at least one approach to assessing your client which is different from the assessments previously administered. The assessment plan must be presented as a list of recommended psychological tests and assessment procedures with a brief sentence explaining the purpose of each test or procedure. Following the list of tests and assessment procedures you recommend for your client, compare the assessment instruments that fall within the same categories (e.g., intellectual or achievement), and debate the pros of cons of the instruments and procedures you selected versus the instruments and procedures reported by the referring professional.

(PLEASE USE REQUIRED SOURCES)

Back to table of contents Previous Chapter Next Chapter

Chapter 2. Schizophrenia Spectrum and Other Psychotic Disorders

https://doi-org.proxy-library.ashford.edu/10.1176/appi.books.9781585624836.jb02

Introduction

John W. Barnhill, M.D.

Schizophrenia is the prototypical psychotic disorder. Not only is it the most common psychosis, but schizophrenia tends to involve abnormalities in all five of the emphasized symptom domains: hallucinations, delusions, disorganized thinking (speech), grossly disorganized or abnormal motor behavior (including catatonia), and negative symptoms. Like the DSM-5 neurodevelopmental disorders, schizophrenia is viewed as a neuropsychiatric disorder with complex genetics and a clinical course that tends to begin during a predictable stage of development. Whereas the neurodevelopmental disorders tend to begin during childhood, symptoms of schizophrenia tend to reliably develop during late adolescence and early adulthood.

The schizophrenia diagnosis has undergone some minor revisions for DSM-5. First, because of their limited diagnostic stability, low reliability, and poor validity, schizophrenia subtypes have been eliminated. They had included such categories as disorganized, paranoid, and residual types of schizophrenia.

Long associated with schizophrenia, catatonia remains one of the potential diagnostic criteria for most of the psychotic diagnoses, including schizophrenia, but it can now be

designated as a specifier for other psychiatric and nonpsychiatric medical conditions, including depressive and bipolar disorders. "Other specified catatonia" can also be diagnosed when criteria are either uncertain or incomplete for either the catatonia or the comorbid psychiatric or nonpsychiatric medical condition.

The DSM-5 schizophrenia diagnosis requires persistence of two of five symptomatic criteria (delusions, hallucinations, disorganized speech, disorganized behavior or catatonia, and negative symptoms). One pertinent change is the elimination of a special status for particular types of delusions and hallucinations, any one of which would previously have been adequate to fulfill symptomatic criteria for schizophrenia. A second change is the requirement for one of the two symptomatic criteria to be a positive symptom, such as delusions, hallucinations, or disorganized thinking.

Criteria for schizoaffective disorder have been significantly tightened. As was the case in DSM-IV, a diagnosis of schizoaffective disorder requires that the patient meet criteria for schizophrenia and have symptoms of either major depressive or bipolar disorder concurrent with having active symptoms of schizophrenia. Also, as was the case previously, there must have been a 2-week period of delusions or hallucinations without prominent mood symptoms. The significant change is that in DSM-5 symptoms that meet criteria for a major mood disorder must be present for the majority of the total duration of the active and residual phases of the overall illness. Therefore, the DSM-5 schizoaffective diagnosis requires more attention to the longitudinal course than was previously the case. Furthermore, the diagnostic requirement that major mood symptoms be present during most of the course of the psychotic disorder (including both the acute and the residual phases) will likely lead to a significant reduction in the number of people who meet criteria for schizoaffective disorder.

Delusional disorder remains focused on the presence of delusions in the absence of other active symptoms of schizophrenia, depressive or bipolar disorders, and pertinent substance use. Bizarre delusions are now included as symptomatic criteria for delusional disorder, whereas delusions that are considered to be part of body dysmorphic disorder and obsessive-compulsive disorder should not lead to a delusional disorder diagnosis but rather to a primary diagnosis of either body dysmorphic disorder or obsessive-compulsive disorder, along with the "absent insight/delusional beliefs" specifier.

Brief psychotic disorder and schizophreniform disorder remain essentially unchanged in DSM-5. They remain distinguished from schizophrenia primarily on the basis of the duration of symptoms.

Not specifically discussed in this text are diagnoses that involve atypical or incomplete presentations or involve situations such as the emergency room setting where information is often incomplete. These include "other specified schizophrenia spectrum and other psychotic disorder," "unspecified catatonia," and "unspecified schizophrenia spectrum and other psychotic disorder."

These "other" diagnoses reflect the reality that humans' thoughts, feelings, and behaviors lie on a continuum, as do their disorders, and the "other" option is a diagnostic option through much of DSM-5. This diagnostic gray zone is particularly poignant in regard to schizophrenia spectrum illness. For many people who end up with a chronic illness such as schizophrenia or schizoaffective disorder, there exists a period of time in which they begin to show symptoms but are not yet diagnosed. It had been proposed that this issue be addressed in DSM-5 by creating a new diagnosis, *attenuated psychosis syndrome.* Psychiatrists are not yet able to robustly predict which patients are most likely to go on to develop full-blown psychotic symptoms, but accurate prediction is important enough that the syndrome is mentioned in two places in DSM-5. First, attenuated psychosis syndrome can be used as a specifier within this chapter of DSM-5, where it would be listed as "other specified schizophrenia spectrum and other disorders (attenuated psychosis syndrome)." The condition is also discussed in more detail among the "Conditions for Further Study."

Suggested Readings

Bromet EJ, Kotov R, Fochtmann LJ, et al: Diagnostic shifts during the decade following first admission for psychosis. Am J Psychiatry 168(11):1186–1194, 2011 PubMed ID: 21676994

Lieberman JA, Murray RM: Comprehensive Care of Schizophrenia: A Textbook of Clinical Management, 2nd Edition. New York, Oxford University Press, 2012

Tamminga CA, Sirovatka PJ, Regier DA, van Os J (eds): Deconstructing Psychosis: Refining the Research Agenda for DSM-V. Arlington, VA, American Psychiatric Association, 2010

Case 2.1 Emotionally Disturbed

Carol A. Tamminga, M.D.

Felicia Allen was a 32-year-old woman brought to the emergency room (ER) by police after she apparently tried to steal a bus. Because she appeared to be an "emotionally disturbed person," a psychiatry consultation was requested.

According to the police report, Ms. Allen threatened the driver with a knife, took control of the almost empty city bus, and crashed it. A more complete story was elicited from a friend of Ms. Allen's who had been on the bus but who had not been arrested. According to her, they had boarded the bus on their way to a nearby shopping mall. Ms. Allen became frustrated when the driver refused her dollar bills. She looked in her purse, but instead of finding exact change, she pulled out a kitchen knife that she carried for protection. The driver fled, so she got into the empty seat and drove the bus across the street into a nearby parked car.

On examination, Ms. Allen was a handcuffed, heavyset young woman with a bandage on her forehead. She fidgeted and rocked back and forth in her chair. She appeared to be mumbling to herself. When asked what she was saying, the patient made momentary eye contact and just repeated, "Sorry, sorry." She did not respond to other questions.

More information was elicited from a psychiatrist who had come to the ER soon after the accident. He said that Ms. Allen and her friend were longtime residents at the state psychiatric hospital where he worked. They had just begun to take passes every week as part of an effort toward social remediation; it had been Ms. Allen's first bus ride without a staff member.

According to the psychiatrist, Ms. Allen had received a diagnosis of "childhood-onset, treatment-resistant paranoid schizophrenia." She had started hearing voices by age 5 years. Big, strong, intrusive, and psychotic, she had been hospitalized almost constantly since age 11. Her auditory hallucinations generally consisted of a critical voice commenting on her behavior. Her thinking was concrete, but when relaxed she could be self-reflective. She was motivated to please and recurrently said her biggest goal was to "have my own room in my own house with my own friends." The psychiatrist said that he was not sure what had caused her to pull out the knife. She had not been hallucinating lately and had been feeling

possible that she was just impatient and irritated. The psychiatrist also believed that she had spent almost no period of life developing normally and so had very little experience with the real world.

Ms. Allen had been taking clozapine for 1 year, with good resolution of her auditory hallucinations. She had gained 35 pounds during that time, but she had less trouble getting out of bed in the morning, was hoping that she could eventually get a job and live more independently, and had insisted on continuing to take the clozapine. The bus trip to the shopping mall was intended to be a step in that direction.

Diagnosis

- Schizophrenia, multiple episodes, currently in active phase

Discussion

Stealing a city bus is not reasonable, and it reflects Ms. Allen's inability to deal effectively with the world. Her thinking is concrete. She behaves bizarrely. She mumbles and talks to herself, suggesting auditory hallucinations. She lives in a state mental hospital, suggesting severe, persistent mental illness.

DSM-5 schizophrenia requires at least two of five symptoms: delusions, hallucinations, disorganized speech, disorganized or abnormal behavior, and negative symptoms. Functioning must be impaired, and continuous signs of the illness must persist for at least 6 months. Even without any more information about Ms. Allen's history, the diagnosis of schizophrenia is clear.

Ms. Allen's psychosis began when she was a child. Early-onset symptoms are often unrecognized because children tend to view their psychotic experience as "normal." Identifying the symptom (e.g., hearing voices that are not there) and associating this with a milestone (e.g., going to a certain grade or school) can help the adult patient retrospectively identify symptom onset. Although the symptoms and treatments are similar for both, childhood-onset schizophrenia is often more severe than adult-onset schizophrenia. Early psychotic symptoms are highly disruptive to normal childhood development. Florid psychotic symptoms are impairing in and of themselves, but they also deprive the young person of the social learning and cognitive development that take place during critical childhood years.

Ms. Allen's behavior on the bus likely reflects not only the psychosis and cognitive dysfunction that are part of schizophrenia but also her diminished experience in real-life social settings. In addition to treating her psychotic symptoms with clozapine, her psychiatric team appears to be trying to remediate her losses by connecting her to a "friend" and organizing the shopping trip. They are also quite active and involved, as reflected by the psychiatrist's almost immediate presence in the ER after the bus incident.

Schizophrenia is a heterogeneous disorder, affecting multiple domains. It is likely that there are multiple schizophrenias, differentiated by as yet unknown markers. Because of insufficient evidence about validity, DSM-5 has done away with categories such as schizophrenia, paranoid type. Instead, DSM-5 outlines several ways in which the diagnosis can be subtyped. One way is by overall activity and chronicity of symptoms (e.g., single vs. multiple episodes; in acute episode, in partial remission, in full remission). Another way to categorize is by assessing the severity of each of the five core schizophrenia symptoms, using a 0–4 scale.

For example, Ms. Allen was able to try to travel with a "friend," and her hospital-based psychiatrist did arrive in the ER very quickly. These might reflect an engaged, active treatment program, but when combined with her apologetic attitude and her stated efforts toward independence, they likely indicate a relative lack of negative symptoms such as anhedonia, reduced social networks, and alogia. Such activity-driven behavior is unusual in patients with schizophrenia and suggests that she is not depressed. It is hard to judge Ms. Allen's cognitive capacity without testing. Her obvious concrete thinking is represented by a failure to understand the process of paying for her bus ride or abstracting behavioral clues. Whether she has the additional characteristics of a schizophrenia-like working memory disorder or attentional dysfunction is hard to tell from this vignette, but she should be tested.

In addition to assessing the extent of positive symptoms, it is crucial for the field of psychiatry to better understand and categorize the negative symptoms and cognitive dysfunction of schizophrenia. Whereas the most effective interventions for schizophrenia have long revolved around the antipsychotic medications that ameliorate positive symptoms, future treatments will likely focus increasingly on the specific behavioral, cognitive, and emotional disturbances that are also an integral part of schizophrenia.

Suggested Readings

Ahmed AO, Green BA, Goodrum NM, et al: Does a latent class underlie schizotypal personality disorder? Implications for schizophrenia. J Abnorm Psychol 122(2):475–491, 2013 PubMed ID: 23713503

Heckers S, Barch DM, Bustillo J, et al: Structure of the psychotic disorders classification in DSM 5. Schizophr Res May 23, 2013 [Epub ahead of print] PubMed ID: 23707641

Tandon R, Gaebel W, Barch DM, et al: Definition and description of schizophrenia in the DSM-5. Schizophr Res June 22, 2013 [Epub ahead of print] PubMed ID: 23800613

Case 2.2 Increasingly Odd

Ming T. Tsuang, M.D., Ph.D., M.Sc.

William S. Stone, Ph.D.

Gregory Baker was a 20-year-old African American man who was brought to the emergency room (ER) by the campus police of the university from which he had been suspended several months earlier. The police had been called by a professor who reported that Mr. Baker had walked into his classroom shouting, "I am the Joker, and I am looking for Batman." When Mr. Baker refused to leave the class, the professor contacted security.

Although Mr. Baker had much academic success as a teenager, his behavior had become increasingly odd during the past year. He quit seeing his friends and spent most of his time lying in bed staring at the ceiling. He lived with several family members but rarely spoke to any of them. He had been suspended from college because of lack of attendance. His sister said that she had recurrently seen him mumbling quietly to himself and noted that he would sometimes, at night, stand on the roof of their home and wave his arms as if he were "conducting a symphony." He denied having any intention of jumping from the roof or having any thoughts of self-harm, but claimed that he felt liberated and in tune with the music when he was on the roof. Although his father and sister had tried to encourage him to see someone at the university's student health office, Mr. Baker had never seen a psychiatrist and had no prior hospitalizations.

During the prior several months, Mr. Baker had become increasingly preoccupied with a female friend, Anne, who lived down the street. While he insisted to his family that they [illegible]ed. Anne told Mr. Baker's sister that they had hardly ever spoken and certainly

were not dating. Mr. Baker's sister also reported that he had written many letters to Anne but never mailed them; instead, they just accumulated on his desk.

His family said that they had never known him to use illicit substances or alcohol, and his toxicology screen was negative. When asked about drug use, Mr. Baker appeared angry and did not answer.

On examination in the ER, Mr. Baker was a well-groomed young man who was generally uncooperative. He appeared constricted, guarded, inattentive, and preoccupied. He became enraged when the ER staff brought him dinner. He loudly insisted that all of the hospital's food was poisoned and that he would only drink a specific type of bottled water. He was noted to have paranoid, grandiose, and romantic delusions. He appeared to be internally preoccupied, although he denied hallucinations. Mr. Baker reported feeling "bad" but denied depression and had no disturbance in his sleep or appetite. He was oriented and spoke articulately but refused formal cognitive testing. His insight and judgment were deemed to be poor.

Mr. Baker's grandmother had died in a state psychiatric hospital, where she had lived for 30 years. Her diagnosis was unknown. Mr. Baker's mother was reportedly "crazy." She had abandoned the family when Mr. Baker was young, and he was raised by his father and paternal grandmother.

Ultimately, Mr. Baker agreed to sign himself into the psychiatric unit, stating, "I don't mind staying here. Anne will probably be there, so I can spend my time with her."

Diagnosis

- Schizophrenia, first episode, currently in acute episode

Discussion

Mr. Baker's case involves an all-too-familiar scenario in which a high-functioning young man undergoes a significant decline. In addition to having paranoid, grandiose, and romantic delusions, Mr. Baker appears to be responding to internal stimuli (i.e., auditory hallucinations) and demonstrating negative symptoms (lying in bed all day). These symptoms have persisted and intensified over the prior year. The history does not indicate medications, substances of abuse, or other medical or psychiatric disorders that could

cause these symptoms. Therefore, he meets DSM-5 criteria for schizophrenia. Although a family history of psychiatric illness is not a requisite for his DSM-5 diagnosis, Mr. Baker's mother and grandmother appear to have also had major mental disorders.

Schizophrenia is, however, a heterogeneous disorder. For example, Mr. Baker's most prominent symptoms are delusions. Another person with schizophrenia might present most prominently with disorganization of speech and behavior and without any delusions. DSM-5 tries to address this heterogeneity by encouraging a dimensional viewpoint rather than a categorical one. In other words, instead of clarifying whether a patient has "paranoid" or "disorganized" schizophrenia, DSM-5 encourages an assessment of a variety of specifiers. One important specifier, the course specifier, requires a longitudinal assessment to determine whether this is a first episode or one of multiple episodes, and whether it is an acute episode, in partial remission, or in full remission.

DSM-5 also encourages specific ratings of symptoms. For example, is this schizophrenic episode accompanied by catatonia? On a 5-point scale (from 0 to 4), how severe is each of the five cardinal schizophrenia symptoms? DSM-5 also encourages an assessment of cognition, mania, and depression domains. For example, some of Mr. Baker's behaviors (e.g., interrupting a class to proclaim his identity as the Joker) may seem to be symptomatic of mania, but they are unaccompanied by disturbances in sleep, mood, or level of activity. Similarly, Mr. Baker said he felt "bad" but not depressed. These clinical observations likely distinguish Mr. Baker from other subcategories of people with schizophrenia.

The schizophrenia diagnosis can be made without assessing these severity specifiers. Nevertheless, the use of dimensional ratings improves the ability to assess Mr. Baker for the presence of core symptoms of schizophrenia in a more individualized manner. The inclusion of dimensions that cut across diagnostic categories will facilitate the development of a differential diagnosis that includes bipolar disorder and schizoaffective disorder. These assessments may clarify Mr. Baker's functional prognosis in major life roles (e.g., living arrangement or occupational status). Finally, repeated dimensional assessments may facilitate a longitudinal understanding of Mr. Baker's symptomatology, development, and likely responses to treatment.

Suggested Readings

Barch DM, Bustillo J, Gaebel W, et al: Logic and justification for dimensional assessment of symptoms and related clinical phenomena in psychosis: relevance to DSM-5. Schizophr Res May 22, 2013 [Epub ahead of print] PubMed ID: 23706415

Cuesta MJ, Basterra V, Sanchez-Torres A, Peralta V: Controversies surrounding the diagnosis of schizophrenia and other psychoses. Expert Rev Neurother 9(10):1475–1486, 2009 PubMed ID: 19831837

Heckers S, Barch DM, Bustillo J, et al: Structure of the psychotic disorders classification in DSM 5. Schizophr Res May 23, 2013 [Epub ahead of print] PubMed ID: 23707641

Tandon R, Gaebel W, Barch DM, et al: Definition and description of schizophrenia in the DSM-5. Schizophr Res June 22, 2013 [Epub ahead of print] PubMed ID: 23800613

Case 2.3 Hallucinations of a Spiritual Nature

Lianne K. Morris Smith, M.D.

Dolores Malaspina, M.D., M.P.H.

Hakim Coleman was a 25-year-old U.S. Army veteran turned community college student who presented to the emergency room (ER) with his girlfriend and sister. On examination, he was a tall, slim, and well-groomed young man with glasses. He spoke softly, with an increased latency of speech. His affect was blunted except when he became anxious while discussing his symptoms.

Mr. Coleman stated that he had come to the ER at his sister's suggestion. He said he could use a "general checkup" because of several days of "migraines" and "hallucinations of a spiritual nature" that had persisted for 3 months. His headache involved "sharp, shooting" sensations in various bilateral locations in his head and a "ringing" sensation along the midline of his brain that seemed to worsen when he thought about his vices.

Mr. Coleman described his vices as being "alcohol, cigarettes, disrespecting my parents, girls." He denied guilt, anxiety, or preoccupation about any of his military duties during his tour in Iraq, but he had joined an evangelical church 4 months earlier in the context of being "riddled with guilt" about "all the things I've done." Three months earlier, he began "hearing voices trying to make me feel guilty" most days. The last auditory hallucination

had been the day before. During these past few months, he had noticed that strangers were commenting on his past sins.

Mr. Coleman believed that his migraines and guilt might be due to alcohol withdrawal. He had been drinking three or four cans of beer most days of the week for several years until he "quit" 4 months earlier after joining the church. He still drank "a beer or two" every other week but felt guilty afterward. He denied alcohol withdrawal symptoms such as tremor and sweats. He had smoked cannabis up to twice monthly for years but completely quit when he joined the church. He denied using other illicit drugs except for one uneventful use of cocaine 3 years earlier. He slept well except occasional nights when he would sleep only a few hours in order to finish an academic assignment.

Otherwise, Mr. Coleman denied depressive, manic, or psychotic symptoms and violent ideation. He denied posttraumatic stress disorder (PTSD) symptoms. Regarding stressors, he felt overwhelmed by his current responsibilities, which included attending school and near-daily church activities. He had been a straight-A student at the start of the school year but was now receiving Bs and Cs.

The patient's girlfriend and sister were interviewed separately. They agreed that Mr. Coleman had become socially isolative and quiet, after having previously been fun and outgoing. He had also never been especially religious prior to this episode. His sister believed that Mr. Coleman had been "brainwashed" by the church. His girlfriend, however, had attended services with Mr. Coleman. She reported that several members of the congregation had told her they had occasionally talked to new members who felt guilt over their prior behaviors, but none who had ever hallucinated, and they were worried about him.

A physical examination of the patient, including a neurological screen, was unremarkable, as were routine laboratory testing, a blood alcohol level, and urine toxicology. A noncontrast head computed tomography (CT) scan was normal.

Diagnosis

- Schizophreniform disorder, provisional

Discussion

The differential diagnosis for a young military veteran with new-onset psychosis and a history of substance abuse is broad. The primary possibilities include a primary psychotic disorder, a psychotic mood disorder, substance-induced psychosis, a psychotic disorder secondary to a general medical condition, a shared cultural syndrome, and PTSD.

Mr. Coleman seems most likely to fit a DSM-5 schizophreniform disorder, a diagnosis that differs from schizophrenia in two substantive ways: the total duration of schizophreniform illness—including prodrome, active, and residual phases—is greater than 1 month but less than 6 months. In addition, there is no criterion that mandates social or occupational impairment. For both schizophreniform disorder and schizophrenia, the patient must meet at least two of five symptomatic criteria. Mr. Coleman describes hallucinations ("hearing voices trying to make me feel guilty") and negative symptoms (blunted affect, avolition, social isolation). The case report does not mention delusions or disorganization of either speech or behavior.

Not relevant to DSM-5 criteria, but of interest, is that Mr. Coleman reports two schneiderian symptoms besides auditory hallucinations: ideas of reference and possible cenesthetic hallucinations based on his description of his atypical headaches ("ringing" in his brain).

DSM-5 indicates that depressive and manic symptoms should be explored as potentially causing the psychosis, and Mr. Coleman denies pertinent mood symptoms. The diagnosis of schizophreniform disorder also requires exclusion of a contributory general medical condition or substance use disorder. Mr. Coleman appears to have no medical complaints, and both his physical examination and laboratory testing are noncontributory.

The patient himself is convinced that his symptoms are due to alcohol. At its worst, however, his drinking appears to have been modest, and he has lately been drinking "a beer or two" every other week. He denies ever having had symptoms of withdrawal or other complications. His hallucinations began months after he cut back on his alcohol use, and they persisted for months. Additionally, his laboratory tests, including a hepatic panel and complete blood count, were normal, which would be unusual in patients with the sort of chronic alcohol use that usually accompanies alcohol-induced psychosis or significant withdrawal. Mr. Coleman's chronic cannabis use could potentially be implicated in the development of psychosis, but not only was his cannabis use sporadic, he apparently had not used for several months prior to the onset of hallucinations, and results of a toxicology

screen were negative. It would appear that Mr. Coleman's concerns about alcohol and cannabis are linked to hyperreligious guilt rather than an actual substance use disorder. The possibility of a general medical condition was considered, but his normal laboratory testing and physical examination results provided no such evidence.

Schizophreniform disorders last at least 1 month but less than 6 months. In regard to Mr. Coleman, his initial 1–2 months of religious preoccupation and guilty ruminations would be considered a prodrome phase. The 3 months preceding presentation to the ER would represent the active phase of psychosis. Because Mr. Coleman's psychotic symptoms have lasted 4–5 months but are ongoing, he would be said to have provisional schizophreniform disorder. Obviously, everyone who goes on to develop schizophrenia has a 6-month period in which they could be said to have schizophreniform disorder, but about one-third of people with schizophreniform disorder do not go on to develop schizophrenia or schizoaffective disorder.

Three other diagnostic possibilities that deserve mention include PTSD, a dissociative disorder, and a shared cultural syndrome. The case does not go into depth about Mr. Coleman's military experience, but simply the experience of being in an active war zone can be a traumatic exposure. He did not report features of PTSD, but it is not clear how extensively possible PTSD symptoms were discussed. Given that avoidance is a cardinal feature of PTSD—making it less likely that he would spontaneously report the symptoms without being prompted—it would be useful to tactfully explore the possibility.

Mr. Coleman's family members indicate that his symptoms began around the time of his initiation into an evangelical church and worry that he has been "brainwashed." DSM-5 includes a possibly pertinent category, listed under "other specified dissociative disorders," within the chapter on dissociative disorders. This disorder is reserved for individuals who experience an identity disturbance due to prolonged and coercive persuasion in the context of such experiences as long-term political imprisonment or recruitment by cults.

It is also possible that Mr. Coleman's unusual beliefs are a nonpathological manifestation of religious beliefs that he shares with other members of his church.

It appears that his psychotic symptoms began prior to his entry into the church, however, and may have been an underlying motivating factor for him to join a church that had previously not been of interest to him. In addition, although he attended church frequently,

there is no evidence that he joined a cult or particularly manipulative religious sect. Furthermore, other congregants viewed his hallucinations as aberrant, indicating that his views were not part of a shared cultural or religious mindset.

The initial diagnosis of provisional schizophreniform disorder is temporary. Longitudinal follow-up will clarify whether Mr. Coleman's symptoms attenuate or progress to a chronic psychotic illness.

Suggested Readings

Bromet EJ, Kotov R, Fochtmann LJ, et al: Diagnostic shifts during the decade following first admission for psychosis. Am J Psychiatry 168(11):1186–1194, 2011 PubMed ID: 21676994

Heckers S, Barch DM, Bustillo J, et al: Structure of the psychotic disorders classification in DSM 5. Schizophr Res May 23, 2013 [Epub ahead of print] PubMed ID: 23707641

Tamminga CA, Sirovatka PJ, Regier DA, van Os J: Deconstructing Psychosis: Refining the Research Agenda for DSM-V. Arlington, VA, American Psychiatric Association, 2010

Case 2.4 Mind Control

Rajiv Tandon, M.D.

Itsuki Daishi was a 23-year-old engineering student from Japan who was referred to his university student mental health clinic by a professor who had become concerned about his irregular school attendance. When they had met to discuss his declining performance, Mr. Daishi had volunteered to the professor that he was distracted by the "listening devices" and "thought control machines" that had been placed in his apartment.

While initially wary of talking to the psychiatrist, Mr. Daishi indicated that he was relieved to finally get a chance to talk in a room that had not yet been bugged. He said that his problems began 3 months earlier, after he returned from a visit to Japan. He said his first indication of trouble was that his classmates sneezed and grinned in an odd way when he entered the classroom. One day when returning from class, he noticed two strangers outside his apartment and wondered why they were there.

Mr. Daishi said that he first noticed that his apartment had been bugged about a week after the strangers had been standing outside his apartment. When he watched television, he noticed that reporters commented indirectly and critically about him. This experience was most pronounced when he watched Fox News, which he believed had targeted him because of his "superior intelligence" and his intention to someday become the prime minister of Japan. He believed that Fox News was trying to make him "go mad" by instilling conservative ideas into his brain, and that this was possible through the use of tiny mind-control devices they had installed in his apartment.

Mr. Daishi's sleep became increasingly irregular as he became more vigilant, and he feared that everyone at school and in his apartment complex was "in on the plot." He became withdrawn and stopped attending classes, but he continued to eat and maintain his personal hygiene.

He denied feeling elated or euphoric. He described his level of energy as "okay" and his thinking as clear "except when they try to put ideas into my head." He admitted to feeling extremely fearful for several hours on one occasion during his recent trip to Japan. At that time, he had smoked "a lot of pot" and began hearing strange sounds and believing that his friends were laughing at him. He denied any cannabis consumption since his return to the United States and denied ever having experimented with any other substances of abuse, saying that he generally would not even drink alcohol. He denied all other history of auditory or visual hallucinations.

When Mr. Daishi's uncle, listed as his local guardian, was contacted, he described his nephew as a healthy, intelligent, and somewhat shy boy without any prior history of any major psychiatric illness. He described Mr. Daishi's parents as very loving and supportive, although his father "might be a little stern." There was no family history of any major mental illness.

On examination, Mr. Daishi was well groomed and cooperative, with normal psychomotor activity. His speech was coherent and goal directed. He described his mood as "afraid." The range and mobility of his affective expression were normal. He denied any ideas of guilt, suicide, or worthlessness. He was convinced that he was being continuously monitored and that there were "mind-control" devices in his apartment. He denied hallucinations. His cognitive functions were grossly within normal limits. He appeared to have no insight into his beliefs.

On investigation, Mr. Daishi's laboratory test results were normal, his head computed tomography scan was unremarkable, and his urine drug screen was negative for any substances of abuse.

Diagnosis

- Delusional disorder, mixed type

Discussion

Mr. Daishi meets criteria for delusional disorder, which requires one or more delusions that persist for greater than 1 month but no other psychotic symptoms. Most of Mr. Daishi's delusions are persecutory and related to monitoring devices. He has delusions of reference (classmates sneezing and grinning at him), persecution ("trying to make me go mad," monitoring devices), and thought insertion ("machines trying to put ideas into my head"). He warrants the "mixed" specifier because the apparent motivation for his having been targeted appears to be grandiose (his "superior intelligence" and plans to be the prime minister of Japan), but he has no other symptoms of mania.

Other psychotic disorders should also be considered. The 3-month duration of symptoms is too long for brief psychotic disorder (no longer than 1 month) and too brief for schizophrenia (no briefer than 6 months) but is an appropriate duration for schizophreniform disorder (between 1 and 6 months' duration). Mr. Daishi does not appear, however, to have a second symptom (e.g., hallucinations, negative symptoms, or disorganization) as required for a schizophreniform diagnosis. In DSM-IV, a single bizarre delusion—the delusion of thought insertion—would have been adequate to reach symptomatic criteria for schizophreniform disorder (or schizophrenia), but bizarre delusions no longer receive special treatment among the DSM-5 schizophrenia spectrum disorders.

The absence of manic or major depressive mood symptoms excludes a diagnosis of bipolar disorder (with psychotic symptoms), major depressive disorder (with psychotic symptoms), or schizoaffective disorder.

Substance-induced psychotic disorder should be considered in view of Mr. Daishi's recent, significant cannabis consumption. His symptoms do seem to have developed soon after consumption of a substance known to cause psychosis (cannabis, with or without

adulteration with another substance such as phencyclidine), and cannabis might be considered a trigger that Mr. Daishi should avoid in the future. DSM-5 specifically excludes the diagnosis of substance-induced psychotic disorder, however, when symptoms persist for a substantial period of time (e.g., 1 month) following the discontinuation of the substance.

Suggested Readings

Cermolacce M, Sass L, Parnas J: What is bizarre about bizarre delusions? A critical review. Schizophr Bull 36(4):667–679, 2010 PubMed ID: 20142381

Nordgaard J, Arnfred SM, Handest P, Parnas J: The diagnostic status of first-rank symptoms. Schizophr Bull 34(1):137–154, 2008 PubMed ID: 17562695

Tandon R: The nosology of schizophrenia: toward DSM-5 and ICD-11. Psychiatr Clin North Am 35(3):557–569, 2012 PubMed ID: 22929866

Tandon R, Carpenter WT: DSM-5 status of psychotic disorders: 1 year prepublication. Schizophr Bull 38(3):369–370, 2012 PubMed ID: 22504739

Case 2.5 Sad and Psychotic

Stephan Heckers, M.D., M.Sc.

John Evans was a 25-year-old single, unemployed white man who had been seeing a psychiatrist for several years for management of psychosis, depression, anxiety, and abuse of marijuana and alcohol.

After an apparently normal childhood, Mr. Evans began to show dysphoric mood, anhedonia, low energy, and social isolation by age 15. At about the same time, Mr. Evans began to drink alcohol and smoke marijuana every day. In addition, he developed recurrent panic attacks, marked by a sudden onset of palpitations, diaphoresis, and thoughts that he was going to die. When he was at his most depressed and panicky, he twice received a combination of sertraline 100 mg/day and psychotherapy. In both cases, his most intense depressive symptoms lifted within a few weeks, and he discontinued the sertraline after a few months. Between episodes of severe depression, he was generally seen as sad, irritable, and amotivated. His school performance declined around tenth grade and remained

marginal through the rest of high school. He did not attend college as his parents had expected him to, but instead lived at home and did odd jobs in the neighborhood.

Around age 20, Mr. Evans developed a psychotic episode in which he had the conviction that he had murdered people when he was 6 years old. Although he could not remember who these people were or the circumstances, he was absolutely convinced that this had happened, something that was confirmed by continuous voices accusing him of being a murderer. He also became convinced that other people would punish him for what had happened, and thus he feared for his life. Over the ensuing few weeks, he became guilt-ridden and preoccupied with the idea that he should kill himself by slashing his wrists, which culminated in his being psychiatrically hospitalized. Although his affect on admission was anxious, within a couple of days he also became very depressed, with prominent anhedonia, poor sleep, and decreased appetite and concentration. With the combined use of antipsychotic and antidepressant medications, both the depression and the psychotic symptoms remitted after 4 weeks. Thus, the total duration of the psychotic episode was approximately 7 weeks, 4 of which were also characterized by major depression. He was hospitalized with the same pattern of symptoms two additional times before age 22, each of which started with several weeks of delusions and hallucinations related to his conviction that he had murdered someone when he was a child, followed by severe depression lasting an additional month. Both relapses occurred while he was apparently adherent to reasonable dosages of antipsychotic and antidepressant medications. During the 3 years prior to this evaluation, Mr. Evans had been adherent to clozapine and had been without hallucinations and delusions. He had also been adherent to his antidepressant medication and supportive psychotherapy, although his dysphoria, irritability, and amotivation never completely resolved.

Mr. Evans's history was significant for marijuana and alcohol abuse that began at age 15. Before the onset of psychosis at age 20, he smoked several joints of marijuana almost daily and binge drank on weekends, with occasional blackouts. After the onset of the psychosis, he decreased his marijuana and alcohol use significantly, with two several-month-long periods of abstinence, yet he continued to have psychotic episodes up through age 22. He started attending Alcoholics Anonymous and Narcotics Anonymous groups, achieved sobriety from marijuana and alcohol at age 23, and had remained sober for 2 years.

Diagnoses

- Schizoaffective disorder, depressive type
- Alcohol use disorder, in remission
- Marijuana use disorder, in remission

Discussion

Mr. Evans has struggled with depression and anxiety since adolescence, worsened by frequent use of marijuana and alcohol. At first, his treaters diagnosed him with depression and panic disorder and treated him accordingly. He did not enter college, as his family had expected, and he has not been employed since graduation from high school. At age 20, psychosis emerged and he required psychiatric hospitalization.

His major psychotic symptom is paranoia, with persecutory delusions and paramnesias of homicide. The delusions are worsened by auditory hallucinations, which he experiences as confirmation of his delusions. The delusions and hallucinations occurred almost daily between ages 20 and 22, until they resolved with clozapine treatment. Although he reports difficulties with his memory, he has not displayed marked cognitive impairment or disorganization of thought. He is socially isolated and minimally able to interact with others. The extent, severity, and duration of his psychotic symptoms are consistent with the diagnosis of a schizophrenia spectrum disorder.

Mr. Evans's psychosis emerged after several years of depression, anxiety, and panic attacks. Since the onset of his psychotic illness, he has experienced multiple episodes of depression, which emerge after periods of delusion and hallucinations and feature overwhelming guilt, prominent anhedonia, poor sleep, and occasional bursts of irritability. He can become suicidal when psychosis and depression reach peak intensity.

Mr. Evans meets criteria, therefore, for DSM-5 schizoaffective disorder. He has had an uninterrupted period in which his major depressive symptoms were concurrent with his schizophrenia symptoms. He has had several-week periods of hallucinations and delusions without prominent mood symptoms. Since the onset of the active and residual portions of his schizophrenia, the major depressive symptoms have been present most of the time.

Mr. Evans also used marijuana and alcohol for 8 years. Although these might have contributed to the emergence of his mood and psychotic symptoms, he continued to

experience significant delusions, hallucinations, and depression between ages 20 and 22, when he stopped using marijuana and alcohol for several months. An alcohol- or marijuana-induced depressive, anxiety, or psychotic disorder might have been considered at various times in Mr. Evans's life, but the persistence of his mood and psychotic symptoms for months after the discontinuation of marijuana and alcohol indicates that he does not have a substance-induced psychiatric disorder.

His response to treatment with antipsychotic, antidepressant, and mood-stabilizing medication is typical: several failed attempts with antipsychotic drugs, the need for combined treatment during periods of exacerbations, and failed attempts to taper either the antidepressant or the antipsychotic medication.

One complicating factor in regard to diagnosing a DSM-5 schizoaffective disorder is the reality that although DSM-5 requires that the mood disorder be present for the majority of the active and residual portions of the schizophrenia, mood and psychotic disorders tend to vary significantly in regard to treatment response and clinical course. For example, whereas depressive and bipolar disorders tend to run in cycles, schizophrenia—once it unfolds—tends to persist. Furthermore, depressive and bipolar disorders tend to be more amenable to treatment than schizophrenia, especially because the diagnostic time frame for the latter includes the residual phase of schizophrenia, which can be largely resistant to psychiatric interventions. It remains to be seen how this tightening of the criteria for schizoaffective disorder will affect the identification and treatment of this cluster of patients.

Suggested Reading

Heckers S: Diagnostic criteria for schizoaffective disorder. Expert Rev Neurother 12(1):1–3, 2012 PubMed ID: 22243039

Case 2.6 Psychosis and Cannabis

Melissa Nau, M.D.

Heather Warm, M.D.

Kevin Foster, a 32-year-old white man with a history of bipolar disorder, was brought to the emergency room (ER) by police after his wife called 911 to report that he was

threatening to jump out of their hotel window.

At the time of the episode, Mr. Foster and his wife were on vacation, celebrating their fifth anniversary. To commemorate the event, they decided to get tattoos. Afterward, they went to a nearby park, where Mr. Foster bought and smoked a marijuana cigarette. During the ensuing hour, Mr. Foster began to believe that the symbols in his tattoo had mysterious meaning and power. He became convinced that the tattoo artist was conspiring with others against him and that his wife was cheating on him. After returning to the hotel, the patient searched his wife's phone for evidence of her infidelity and threatened to jump out the window. The patient's wife, an ER physician, successfully convinced the patient to go to sleep, thinking that the episode would resolve.

The following day, the patient remained paranoid and delusional. He again threatened to jump out the window, and indicated that he would have no choice but to kill his wife the next time she slept. She called 911, and her husband was brought to the ER of a large nearby hospital. Later that day, he was admitted to an acute inpatient psychiatric unit with a diagnosis of unspecified psychotic disorder.

The patient had smoked cannabis sporadically from age 18 but began to smoke daily 5 years prior to this admission. He and his wife denied that he had ever used other illicit substances, and the patient indicated that he rarely drank alcohol. Until 1 year earlier, he had never seen a psychiatrist or been viewed by his friends and family as having significant psychiatric issues.

In the past year, however, Mr. Foster had been hospitalized four times for psychiatric problems. He had been hospitalized twice with classic manic symptoms and once for a suicidal depression. In addition, 7 months prior to this presentation, the patient had been hospitalized for a 6-week episode of cannabis-induced psychosis, which responded well to risperidone. At that time, his main symptom was paranoia. Two months prior to this admission, he entered a 1-month inpatient substance abuse treatment program for cannabis use disorder. Until the weekend of this admission, he had not used marijuana, alcohol, or any other substances since discharge from the rehabilitation facility. He had also been functioning well while taking lithium monotherapy for 3 months.

Mr. Foster had been steadily employed as a film editor since graduating from college. His father had a bipolar disorder, and his paternal grandfather committed suicide via gunshot

but with an unknown diagnosis.

On the second day of hospitalization, the patient began to realize that his wife was not cheating on him and that the symbols in his tattoo were not meaningful. By the third day, he spontaneously said the paranoia was the result of cannabis intoxication. He declined further risperidone but continued lithium monotherapy. He was discharged with an appointment to follow up with his outpatient psychiatrist.

Diagnoses

- Cannabis-induced psychotic disorder
- Bipolar disorder, in remission

Discussion

Soon after smoking a marijuana cigarette, Mr. Foster began to believe that the symbols of his new tattoo had mysterious meaning and power. Within hours, he became paranoid about the tattoo artist and delusionally jealous. He threatened to kill himself and his wife. He was admitted to a psychiatric unit. The psychotic symptoms cleared within a few days, and the patient regained appropriate insight. This symptom trajectory fits DSM-5 substance/medication-induced psychotic disorder, which requires delusions or hallucinations that develop during, or soon after, a substance intoxication (or withdrawal or medication exposure).

An additional DSM-5 diagnostic criterion for cannabis-induced psychotic disorder revolves around whether Mr. Foster's delusions might not be better explained by a primary psychotic disorder such as schizophrenia or psychotic symptoms within depression or mania. His symptoms resolved within 3 days, which is typical for a cannabis-induced psychosis but not for an independent psychotic disorder. The rapid resolution of symptoms would support the likelihood that the cannabis caused his symptoms.

Mr. Foster's psychiatric history complicates the diagnosis in two different ways. First, of the four psychiatric hospitalizations Mr. Foster has had in the past year, one was for paranoid delusions in the context of cannabis use, leading to a 6-week hospitalization. The duration of the actual paranoid delusions is not entirely clear, but they appear to have lasted far longer than would be typical for a cannabis-induced psychosis. DSM-5

specifically cautions that persistence of a psychosis beyond 1 month after the exposure implies that the psychosis may be independent rather than substance induced.

Second, of Mr. Foster's three other psychiatric hospitalizations, two were for "classic" mania and one was for "suicidal depression." It is not clear whether paranoia or psychosis was part of these episodes. DSM-5 points out that a history of recurrent non-substance-related psychotic episodes would make a substance-induced psychosis less likely.

It is not clear whether these psychiatric episodes can be brought together under a single diagnostic umbrella. For example, Mr. Foster could have bipolar disorder with recurrent episodes of depression and mania. The cannabis might help him sleep—which might reduce the mania—but could possibly trigger episodes. If manic and depressive episodes (with or without psychosis) are triggered by a substance but symptoms persist for an extended period of time, then the most accurate diagnosis would be the bipolar disorder. This would be especially true if similar symptoms develop in the absence of substance use. Mr. Foster has a family history significant for bipolar disorder, which could further support such a diagnosis. On the other hand, Mr. Foster did not endorse any mood symptoms during this most recent psychotic episode, and psychotic symptoms resolved within 2–3 days. This history would seem to indicate that although Mr. Foster has historically met criteria for bipolar disorder, it seems to be currently in remission.

Multiple schizophrenia spectrum disorders might be considered. Given a 3-day duration of symptoms, however, most diagnoses are quickly eliminated as possibilities. In addition, Mr. Foster appears to have only one affected domain (delusions). Delusional disorder involves only delusions, but the minimum duration is 1 month. Brief psychotic disorder also requires only one of the four primary schizophrenia spectrum symptoms, but it does require an evaluation as to whether the precipitant might be a substance or medication.

At the moment, then, a cannabis-induced psychotic disorder appears to be the most likely diagnosis for Mr. Foster's particular episode. Clarification might be possible through more thorough investigation of prior medical records, but even more helpful will be ongoing, longitudinal follow-up.

Suggested Readings

Caton CL, Hasin DS, Shrout PE, et al: Stability of early-phase primary psychotic disorders with concurrent substance use and substance-induced psychosis. Br J Psychiatry

190:105–111, 2007 PubMed ID: 17267925

Ekleberry S: Treating Co-Occurring Disorders: A Handbook for Mental Health and Substance Abuse Professionals. Binghamton, NY, Haworth, 2004

Grant BF, Stinson FS, Dawson DA, et al: Prevalence and co-occurrence of substance use disorders and independent mood and anxiety disorders: results from the National Epidemiologic Survey on Alcohol and Related Conditions. Arch Gen Psychiatry 61(8):807–816, 2004 PubMed ID: 15289279

Pettinati HM, O'Brien CP, Dundon WD: Current status of co-occurring mood and substance use disorders: a new therapeutic target. Am J Psychiatry 170(1):23–30, 2013 PubMed ID: 23223834

Case 2.7 Flea Infestation

Julie B. Penzner, M.D.

Lara Gonzalez, a 51-year-old divorced freelance journalist, brought herself to the emergency room requesting dermatological evaluation for flea infestation. When no corroborating evidence was found on skin examination and the patient insisted that she was unsafe at home, she was admitted to an inpatient psychiatric service with "unspecified psychotic disorder."

Her concerns began around 1 week prior to presentation. To contend with financial stress, she had taken in temporary renters for a spare room in her home and had begun pet sitting for some neighbors. Under these conditions, she perceived brown insects burrowing into her skin and walls and covering her rugs and mattress. She threw away a bag of clothing, believing she heard fleas "rustling and scratching inside." She was not sleeping well, and she had spent the 36 hours prior to presentation frantically cleaning her home, fearing that her tenants would not pay if they saw the fleas. She showered multiple times using shampoos meant to treat animal infestations. She called an exterminator who found no evidence of fleas, but she did not believe him. She was upset about the infestation but was otherwise not troubled by depressive or manic symptoms, or by paranoia. She did not use drugs or alcohol. No one in the family had a history of psychiatric illness. Ms. Gonzalez had

had depression once in the past and was briefly treated with an antidepressant. She had no relevant medical problems.

Her worries about infestation began in the setting of her sister's diagnosis with invasive cancer, the onset of her own menopause, financial strain that was likely forcing her to move from the United States back to Argentina (her country of origin), and a recent breakup with her boyfriend. At baseline, she described herself as an obsessive person who had always had contamination phobias, which historically worsened during times of anxiety.

On mental status examination, Ms. Gonzalez was calm and easily engaged, with normal relatedness and eye contact. She offered up a small plastic bag containing "fleas and larvae" that she had collected in the hospital while awaiting evaluation. Inspection of the bag revealed lint and plaster. Her speech had an urgent quality to it, and she described her mood as "sad right now." She was tearful intermittently but otherwise smiling reactively. Her thoughts were overly inclusive and intensely focused on fleas. She expressed belief that each time a hair fell out of her head, it would morph into larvae. When crying, she believed an egg came out of her tear duct. She was not suicidal or homicidal. She expressed an unshakable belief that lint was larvae, and that she was infested. She denied hallucinations. Cognition was intact. Her insight was impaired, but her judgment was deemed reasonably appropriate.

Dermatological examination revealed no insects or larvae embedded in Ms. Gonzalez's skin. Results of neurological examination, head computed tomography scan, laboratory tests, and toxicology data were normal. She was discharged on a low-dose antipsychotic medication and seen weekly for supportive psychotherapy. Her preoccupation improved within days and resolved entirely within 2 weeks. She developed enough insight to refer to her belief that fleas were in her skin as a "crazy thought." She attributed her "break from reality" to multiple stressors, and was able to articulate that she relied on her delusion as a way to distract herself from real problems. Her family corroborated her quick return to baseline.

Diagnosis

- Brief psychotic disorder with marked stressors

Discussion

Ms. Gonzalez's delusions with quick return to full premorbid functioning suggest a diagnosis of brief psychotic disorder with marked stressors. Formerly called "brief reactive psychosis," a brief psychotic disorder (with or without marked stressors) may not be diagnosed until return to baseline has occurred. The differential diagnosis of this condition is important.

At the time of admission, the patient was diagnosed with "unspecified psychotic disorder," a term often used when psychosis is present but information is incomplete. Only after her symptoms rapidly resolved could she be diagnosed with a brief psychotic disorder. Ms. Gonzalez's insight returned quite quickly, and she was able to link her symptoms to antecedent stressors. Although treatment is likely to shorten the duration of an acute psychotic episode, DSM-5 specifically does not factor treatment into the requirement that the episode last less than 1 month.

It is worth noting that stressors can be positive (e.g., marriage, new job, new baby) or negative, as in Ms. Gonzalez's case. A favorable prognosis is often associated with a history of good premorbid functioning, significant acute stressors, and a lack of family or personal history of psychiatric illness.

Ms. Gonzalez's sleeplessness, behavioral agitation, and premorbid depressive history might also suggest bipolar episode, but there are no other symptoms to support this diagnosis. Similarly, her delusional obsession with flea infestation suggests a possible delusional disorder, but Ms. Gonzalez's symptoms resolved far too quickly for this to be likely. Patients with personality disorders can have "micropsychoses," but Ms. Gonzalez does not appear to have a personality disorder or particular personality vulnerability. Malingering and factitious disorder appear unlikely, as do delirium and other medically mediated illnesses.

Brief psychotic episodes have a low prevalence in the population, which could indicate that brief psychoses are unusual. It could also indicate that people with a very short duration of psychotic symptoms may not seek psychiatric help. The brevity and unpredictability of symptoms also makes it difficult to do research and for any particular clinician or institution to develop an expertise. Brief psychotic episodes are also noted to have a relatively low stability over time, which makes sense given that—unlike schizophrenia—brief psychotic episodes are, by definition, of short duration and cannot even be diagnosed without both remission of symptoms and careful follow-up.

Suggested Readings

Jørgensen P, Bennedsen B, Christensen J, Hyllested A: Acute and transient psychotic disorder: comorbidity with personality disorder. Acta Psychiatr Scand 94(6):460–464, 1996 PubMed ID: 9021000

Salvatore P, Baldessarini RJ, Tohen M, et al: McLean-Harvard International First-Episode Project: two-year stability of DSM-IV diagnoses in 500 first-episode psychotic disorder patients. J Clin Psychiatry 70(4):458–466, 2009 PubMed ID: 19200422

SOURCE:
Psychological testing: History, principles and applications (7th ed.). Boston, MA: Pearson

CHAPTER 10

Neuropsychological Assessment and Screening

TOPIC 10A Neurobiological Concepts and Behavioral Assessment

10.1 The Human Brain: An Overview
(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/ch10lev1sec1#ch10lev1sec1)

10.2 Structures and Systems of the Brain
(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/ch10lev1sec2#ch10lev1sec2)

10.3 Survival Systems: The Hindbrain and Midbrain
(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/ch10lev1sec3#ch10lev1sec3)

10.4 Attentional Systems
(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/ch10lev1sec4#ch10lev1sec4)

10.5 Motor/Coordination Systems
(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/ch10lev1sec5#ch10lev1sec5)

10.6 Memory Systems
(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/ch10lev1sec6#ch10lev1sec6)

10.7 Limbic System
(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/ch10lev1sec7#ch10lev1sec7)

10.8 Language Functions and Cerebral Lateralization
(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/ch10lev1sec8#ch10lev1sec8)

10.9 Visual System
(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/ch10lev1sec9#ch10lev1sec9)

10.10 Executive Functions
(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/ch10lev1sec10#ch10lev1sec10)

10.11 Neuropathology of Adulthood and Aging
(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/ch10lev1sec11#ch10lev1sec11)

10.12 Behavioral Assessment of Neuropathology
(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/ch10lev1sec12#ch10lev1sec12)

In the practice of assessment, psychologists often discover that their clients need assistance with various problems that are best understood from a neurobiological standpoint. These problems typically

arise as a consequence of head injury, learning disability, memory impairment, language disorder, or attentional difficulties, to list just a few examples. Tens of millions of individuals are affected. For example, in the United States an estimated 5 to 8 million children struggle with a learning disability (**Dey, Schiller, & Tai, 2004** (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib411)), about 13 to 16 million adults live with memory loss and other symptoms related to dementia (**Alzheimer's Disease and Related Disorders Association, 2000** (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib25)), and approximately 1.7 million people experience a head injury *each year* (**Faul, Xu, Wald, & Coronado, 2010** (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib488)).

These numbers are staggering, and they provide an ongoing mandate for psychologists to develop specialized tests and procedures at the interface of psychology and medicine. The purpose of this chapter is to summarize pertinent tests, concepts, methods, and issues encountered in neuropsychological assessment and ancillary areas of appraisal such as substance abuse evaluation and screening for dementia. In **Topic 10A** (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/ch10#ch10box1) , Neurobiological Concepts and Behavioral Assessment, we provide a condensed review of neurobiological concepts relevant to psychological testing and assessment. The emphasis in this topic is upon the various brain systems that underlie effective cognitive and emotional functioning. Understanding these brain systems is essential for those who study or use psychological tests. In this primer, the reader also will encounter several of the simpler approaches to assessment used by neuropsychologists. In the process, a good foundation will be set for **Topic 10B** (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/ch10lev1sec12#ch10box2) , Neuropsychological Tests, Batteries, and Screening Tools, which reviews prominent neuropsychological instruments, test batteries, and screening tools.

10.1 THE HUMAN BRAIN: AN OVERVIEW

By convention the nervous system is divided into the central nervous system consisting of the brain and spinal cord, and the peripheral nervous system that includes the cranial nerves and the network of nerves emanating from the spinal cord. The brain is intimately involved in thinking, feeling, and behaving. For these reasons, our focus in this topic is the structure and function of the brain.

The brain is beyond doubt the most protected organ in the human body. The first line of defense against physical trauma is the skull, consisting of several intermeshed, rigid bones that almost completely encase the brain. Beneath the skull, the brain is also surrounded by the **meninges** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm01#bm01gloss197)** , a thin layering of three tough membranes that encases the brain and spinal cord, providing additional protection. The middle spongy layer of the meninges is filled with another form of protection, cerebrospinal fluid, which buffers the brain against sudden acceleration and deceleration, such as from a blow to the head. The brain literally floats in a snugly fitting bath of cerebrospinal fluid. Buoyancy reduces the effective weight of the organ to a few ounces, vastly reducing pressure upon the base of the brain. Without the protection of this fluid, the brain would bruise easily from any rapid movement of the head.

When unbouyed, the brain weighs less than three pounds. It is composed principally of five elements: gray matter, white matter, glial cells, cerebrospinal fluid (CSF), and the blood vessels of the vascular system that provide the brain with oxygen and nutrients.

The 10^{11} or 100 billion neurons in the brain are arranged in complex networks that largely have defied understanding. In part, the inscrutability of the brain derives from its computational complexity. Neurons communicate by sending all-or-none electrochemical impulses to one another. Each neuron might send transmissions to thousands, perhaps tens of thousands, of other neurons at near and distant sites called synapses. Chemical communications across the synapses can occur up to a thousand times a second. Even if we use a conservative estimate of a thousand synapses per neuron, in theory the number of neural transmissions that could occur in just one second is a staggering 10^{17} or 100,000,000,000,000,000 (one hundred quadrillion). No wonder that staid neuroscientists such as Sir John Eccles (who received a Nobel Prize for his work in neurophysiology) resort to hyperbole and describe the brain as "without qualification the most highly organized and most complexly organized matter in the universe" (**Eccles, 1973** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib448)**). Considering how little we know of the universe, the truth of this statement is open to question. But it does effectively underscore the point that neuroscientists approach the study of the human brain with a sense of awe.

Cerebrospinal Fluid and the Ventricular System

Cerebrospinal fluid **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm01#bm01gloss42)** (CSF) is a clear liquid that is continuously produced and replenished within the ventricles. The **ventricles** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm01#bm01gloss346)** are hollow, interconnected chambers found in the middle of the brain. There are four ventricles: two side-by-side ventricles, called the lateral ventricles, and two midline ventricles known as the third and fourth ventricles.

In rare cases, the normal flow of CSF can become constricted, such as when the aqueduct leaving the third or fourth ventricle becomes too small. This can be a congenital condition present at birth or a

of the ventricles and compression of the brain against the skull. In time, the skull can even enlarge. This condition is known as hydrocephalus or, literally, "water on the brain." Untreated, the consequence of hydrocephalus can be mental retardation and early mortality. Fortunately, effective treatments are available, including the insertion of a shunt to drain the excess fluid from the ventricles—usually into the child's abdomen.

The Vascular System of the Brain

Metabolically, the brain is a highly active organ, needing substantial supplies of oxygen and glucose to function effectively. These energy sources are supplied by the flow of blood through the cardiovascular system. Hence, the general physical health of the client and the specific condition of his or her vascular system in the brain are essential to highlevel cognitive functioning.

Two pairs of arteries carry blood to the brain. These are the left and right internal carotid arteries, found in the front of the neck, and the left and right vertebral arteries, found in the back of the neck. The vertebral arteries come together just below the base of the brain to form a single artery, the basilar artery. These three arteries—the left and right internal carotids and the basilar artery—all feed into a circular arterial structure at the base of the brain known as the circle of Willis. This circular network ensures that the brain receives a continual supply of blood, even if one of the input arteries is compromised.

From this circular arterial system at the base of the brain, three arteries branch upward on each side to the roughly symmetrical cerebral hemispheres of the brain. The anterior cerebral arteries supply blood to the left and right frontal lobes and some midline structures. The middle cerebral arteries provide blood to the vast majority of the lateral surface of each hemisphere, including the frontal, parietal, and temporal lobes, and to some internal structures as well. Finally, the posterior cerebral arteries supply blood to the left and right occipital lobes and to additional subcortical structures.

Especially with advancing age, it is not unusual for one or more arteries in the brain to become completely obstructed by a condition known as atherosclerosis, the gradual buildup of fatty plaque. When an artery in the brain becomes completely obstructed—whether gradually or suddenly from a piece of dislodged plaque—the brain tissue supplied by that vessel dies because it is deprived of oxygen. This event is called an infarct, which is one kind of stroke or **cerebrovascular accident** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm01#bm01gloss43)** (CVA). Another kind of CVA occurs when a bulging area of arterial weakness, called an aneurysm, bursts open, allowing blood to spurt directly into the brain tissue. This is technically known as an arterial rupture. The effects of a CVA depend upon the size and location of the resulting damage to the brain. For example, an infarct occurring at the base of the left middle cerebral artery would have calamitous generalized effects (e.g., right-sided paralysis of the body, loss of speech), whereas an infarct occurring higher up, in a smaller offshoot from the artery, might have limited effects or even go unnoticed. One form of vascular impairment known as **multi-infarct dementia** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm01#bm01gloss211)** (MID) occurs when the hardly noticeable individual effects of many small infarcts accumulate over a number of years. The symptoms of MID are varied but often impact the ability to perform everyday activities such as eating, dressing, and shopping. The symptoms might include forgetfulness, vague or circumstantial speech, lack of concentration, loss of balance, physical weakness, difficulty following instructions, and problems handling money. Often the onset of MID is so gradual and insidious that relatives recognize only in retrospect that something has been wrong for months after the onset of problems.

10.2 STRUCTURES AND SYSTEMS OF THE BRAIN

The organization of the human brain is difficult to comprehend because important structures are interwoven and folded over upon one another. As noted, the brain also contains an intricate system of fluid-filled caverns, the ventricles, further complicating the spatial arrangement of important brain structures. In addition, functional brain systems rarely obey any simple structural organization—they typically meander their way from one part of the brain to another. Hence, we will focus mainly on a functional systems approach to explaining the operation of the brain, alluding to structures when appropriate.

We begin with a quick overview of the central nervous system and its primary subdivisions. The most basic element of the nervous system is the **cerebrum** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm01#bm01gloss44)** , consisting of the left and right cerebral hemispheres, which are connected by the corpus callosum, a band of fibers that transfers information from one hemisphere to the other. From the standpoint of evolution, the cerebrum is the most recent part of the brain to develop. This is where thought, perception, imagination, judgment, and decision occur. Some essential structures located beneath the cerebrum are the basal ganglia and the cerebellum (both important in coordinated movement), the diencephalon (including the thalamus), the midbrain (consisting of the cranial nerves and other important relay stations), the pons (connecting the cerebrum with the cerebellum and the spinal cord), and the medulla (mediating essential bodily functions).

Corpus Callosum

The **corpus callosum** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm01#bm01gloss71)** is the major commissure that serves to integrate the functions of the two cerebral hemispheres. This large bundle of subcortical nerve fibers is about four inches long and a quarter inch thick. The corpus callosum spans the brain from side to side just above the level of the thalamus. Although there are exceptions, the corpus callosum generally connects homologous brain sites in the left and right hemispheres.

The function of the corpus callosum was poorly understood until the 1960s when Sperry, Gazzaniga, and others initiated sophisticated laboratory studies of so-called split-brain patients (**Sperry, 1964** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1531)** ; **Gazzaniga, 1970** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib573)** ; **Gazzaniga & LeDoux, 1978** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib574)**). These patients were persons with epilepsy whose corpus callosa had been severed to prevent the transport of epileptic discharges from one hemisphere to the other. Although outwardly normal, split-brain patients revealed a striking isolation of consciousness when visual information was restricted to one hemisphere or the other. For example, when a picture of an apple was tachistoscopically presented to the left side of the examinee's fixation point, this stimulus was processed only in the right hemisphere (on account of the normal crossing over of neural connections). Furthermore, because the corpus callosum was severed, the image of the apple remained trapped in the right hemisphere. As the reader will discover later, the right hemisphere is usually mute and does not subserve important language functions. Thus, when asked, "What did you see?" the examinees, responding from the verbal left hemisphere, would honestly reply, "Nothing." Yet, these patients could readily identify the object by pointing to it with the left hand (which is under the neural control of the right hemisphere). This suggests that although the right hemisphere cannot talk, it has a separate and independent capacity to perceive,

In a normal individual with intact corpus callosum, consciousness appears unitary because the two halves of the brain can communicate and forge a compromise as regards perception, thought, and action. Much of our knowledge of hemispheric specializations, discussed later, has been garnered from the detailed study of split-brain patients. Further insight has been gained from studies of persons living with the congenital absence of this structure, a condition known as agenesis of the corpus callosum (ACC). Present in about 1 in 4,000 live births, ACC manifests with a variety of deficits, superbly summarized by Paul, Brown, Adolphs, and others (**2007** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1272)**). Even though overall IQ is minimally impacted, impairments are observed in abstract reasoning, problem solving, and category fluency (e.g., the ability to list multiple items in a category such as animals). One intriguing symptom that bears on current understanding of language function is that persons with ACC show marked difficulty in the verbal expression of emotional experience. Parents of children with the disorder consistently describe conversations that are meaningless or out of place (**Paul et al., 2007** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1272)**). This corresponds well with known lateralization of brain function, in which logical components of language are underwritten by the left hemisphere, whereas the emotional aspects of language are subserved by the right hemisphere. In the absence of a corpus callosum, individuals with ACC find it particularly difficult to synthesize these two elements of language.

Cerebral Cortex

The **cerebral cortex** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm01#bm01gloss41)** , the outermost layer of the brain, is the source of the highest levels of sensory, motor, and cognitive processing. Also called the neocortex, the cerebral cortex is a very recent evolutionary development. It is the functional capacity of this brain system—a uniform six layers deep—that most dramatically separates humans from the lower animals.

The tissue of the cerebral cortex is folded over into elaborate convolutions consisting of bulges and grooves. The prominent bulges are called gyri (singular *gyrus),* whereas the clefts, fissures, and grooves are called sulci (singular *sulcus).* This arrangement allows the brain to have a great deal more cerebral cortex than if the surface were smooth. Although the pattern of gyri and sulci is subtly unique for each person, certain major landmarks such as the central sulcus and the lateral sulcus (**Figure 10.1** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/ch10lev1sec2#ch10fig1)**) are always discernible in a normal brain.

A small portion of the cerebral cortex is committed cortex. These sites are dedicated to basic sensory processing of vision, hearing, touch, and motor control. Nonetheless, the specificity of committed cortex is relative, not absolute. For example, the precentral gyrus classically is regarded as the motor cortex (see **Figure 10.1** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/ch10lev1sec2#ch10fig1)**), but only a fraction of the neurons subserving voluntary movement are located there. This has been demonstrated through neurosurgical investigations of the exposed cortex in persons with epilepsy, beginning with the pioneering work of Wilder Penfield (**1958** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1278)**). The fully conscious patient received local anesthesia while surgeons opened a skull flap to expose one side of the brain. Then a stylus was used to deliver a small, brief, harmless electrical charge to specific sites in the sensory, motor, and language areas. The purpose of this procedure was to map the topography of the cortex so that vital brain sites were not excised. Using this approach, Uematsu, Lesser, Fisher, and others (**1992** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1677)**) reconfirmed that a significant proportion—more than one-third—of motor responses originate outside

adjoining brain sites. Furthermore, the motor strip contains a sizeable proportion of sensory cells, too. Thus, cells that subserve each specific sensory or motor function are highly concentrated in the respective committed area, but also thin out and overlap with nearby brain sites. In brief, the committed cortex of the **frontal lobe** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm01#bm01gloss128)** is dedicated to motor control, the **parietal lobe** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm01#bm01gloss234)** is concerned with the processing of touch and other somatosensory information, the occipital lobe is involved in visual perception, and the **temporal lobe** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm01#bm01gloss324)** is essential to the processing of auditory information. Of course, these brain regions serve other functions as well, but part of each major lobe is dedicated to a specific motor or sensory function (**Figure 10.2** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/ch10lev1sec2#ch10fig2)**).

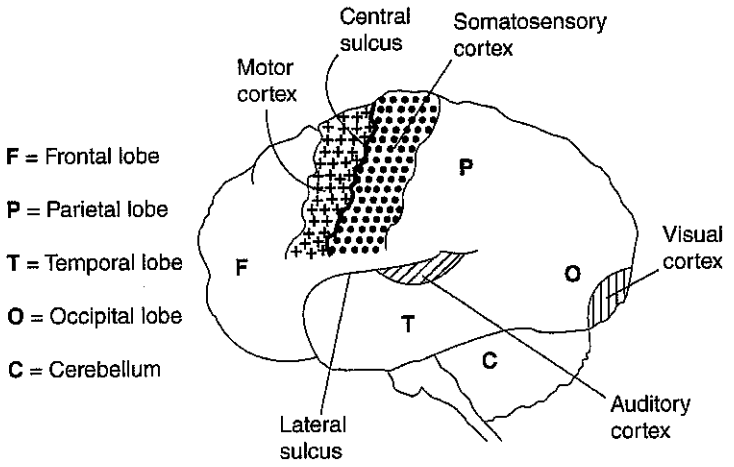


FIGURE 10.1 Major Landmarks of the Left Cerebral Hemisphere

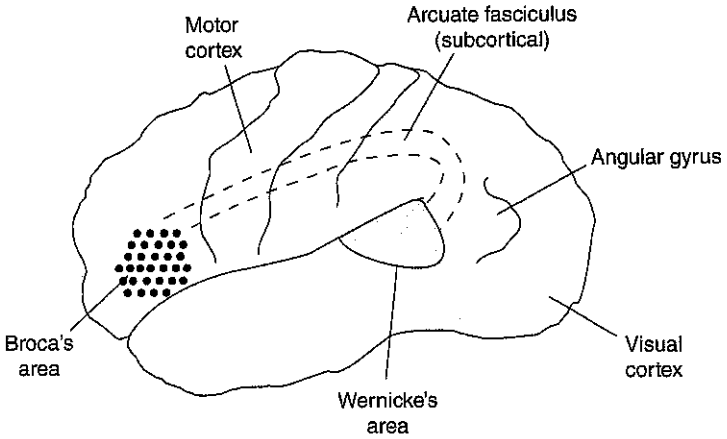


FIGURE 10.2 The Structural Model of Left Hemisphere Language Functions

10.3 SURVIVAL SYSTEMS: THE HINDBRAIN AND MIDBRAIN

The lowest part of the brain, located at the top of the spinal cord, consists of the **hindbrain** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm01#bm01gloss144)** , which includes the medulla oblongata, the pons, the reticular formation, and the cerebellum. From the standpoint of evolution, the hindbrain was the first brain system to develop, which explains why so many vital bodily functions are governed by this brain area. For example, the automatic control of breathing is mediated here—we breathe even when asleep, or for that matter, when in a deep coma.

The lowest section of the hindbrain is the medulla oblongata, which mediates several essential bodily functions: breathing, swallowing, vomiting, blood pressure, and, partially, heart rate (**Kandel, Schwartz, & Jessell, 1995** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib857)**). Aspects of talking and singing also are governed here, although higher brain sites are intimately involved in these functions as well.

Significant damage to the medulla usually is fatal. In rare cases, a small stroke in the medulla causes one or more of the following symptoms: opposite-sided paralysis, partial loss of pain and temperature sense, clumsiness, dizziness, partial loss of the gag reflex, and same-sided paralysis and atrophy of the tongue. Thus, one reason why neurologists ask patients to stick out their tongue and move it from side to side is specifically to check for neurological damage in and around the medulla. The polio virus—rampant in the 1950s but now well controlled—may attack the medulla, shutting down the neural control of breathing and necessitating a mechanical respirator.

The pons and cerebellum are the highest structures in the hindbrain. Together they help coordinate muscle tone, posture, and hand and eye movement. The role of the cerebellum in motor control is discussed later. Lesions of the pons may render the individual incapable of making coordinated lateral eye movements. For this reason, neurologists and neuropsychologists commonly ask patients to demonstrate left-right and up-down eye movements.

Located just above the hindbrain is the **mid-brain** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm01#bm01gloss204)** , which includes a number of important relay stations involved in hearing and vision. In addition, the midbrain contains nuclei for many of the cranial nerves (some of which also emanate from the hindbrain). The 12 paired **cranial nerves** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm01#bm01gloss75)** are major neural tracts whose functions are well understood and easily tested. Some are exclusively sensory, relaying information from the external world to the brain; some are exclusively motor, serving to execute commands from the brain; about a third of the cranial nerves possess both sensory and motor functions. Neurologists refer to the cranial nerves by number. The numbers correspond roughly to the top to bottom sequence of the nerves' emergence from the brain (**Table 10.1** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/ch10lev1sec3#ch10tbl1)**). The reader will notice that many cranial nerves mediate aspects of vision and eye movement, basic sensory functions, and movement of jaw, tongue, face, and head. Over the centuries, neurologists have devised a variety of simple confrontational techniques to assess the cranial nerves. As peculiar as it may appear, asking the patient to stick out his or her tongue and move it left, right, up, or down can provide important information about the functioning of the hypoglossal (12th) cranial nerve. In like manner, various simple tests of hearing, balance, eye movement, and so on are used to complete the examination of the cranial nerves.

TABLE 10.1 The Cranial Nerves and Their Functions

1. Olfactory	Sense of smell
2. Optic	Vision
3. Oculomotor	Horizontal and vertical eye movement
4. Trochlear	Vertical eye movement
5. Trigeminal	Facial sensation, jaw movement
6. Abducens	Horizontal eye movement
7. Facial	Facial movement and taste
8. Auditory/vestibular	Hearing and balance
9. Glossopharyngeal	Taste, swallowing
10. Vagus	Visceral reflexes
11. Accessory	Head movement
12. Hypoglossal	Tongue movement

10.4 ATTENTIONAL SYSTEMS

Attention has been likened to a "spotlight" that our brain uses to identify what is relevant and ignore what is irrelevant (**Andreasen, 2001** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib51)**). Attention is often a primitive, automatic cognitive system that is essential for survival. Consider the variety of competing stimuli encountered when you drive a car down the highway, perhaps with a friend sitting next to you. A realistic scenario is that your friend asks a question, an airplane flies low in the distant horizon, a billboard on the left lures your visual focus, a siren blares in the distance, your back aches from a strenuous workout—all these sources of stimulation compete for your attention. Then a car swerves into your lane. Instantly, without conscious forethought, your brain focuses every last fragment of attention on this one looming threat, ignoring all else.

Neuropsychologists have identified several kinds of **attention** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm01#bm01gloss20)** , including the following types:

- Orienting
- Selective
- Divided
- Sustained

Orienting attention is the simplest and most primitive form, related to the "fight" or "flight" reflex. This is the straightforward direction of all attentional resources to a single threatening stimulus, such as a car swerving into your lane. Selective attention refers to the identification of a single, personally relevant stimulus embedded within a flow of extraneous information. This is exemplified when, for example, a young boy who seems absorbed in solitary play nonetheless turns his head when he overhears his name spoken quietly in the background. Divided attention, also known as distributed attention, pertains to the ability to shift back and forth between two or more tasks. An example might be when a partygoer tries to follow two conversations at the same time. Sustained attention, also known as vigilance, refers to the ability to sustain attention over relatively long periods of time. This involves the capacity to resist distraction and stay on task for a prolonged period. A good example is the air traffic controller who must monitor radar images carefully to keep airplanes at a safe distance from one another.

The exact neurological mechanisms of attention are not well understood. Kandel, Schwartz, and Jessell (**1995 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib857)**) note that the "neuronal mechanisms of focused attention and conscious awareness are now emerging as one of the great unresolved problems in perception and indeed in all of neurobiology" (p. 402). Neurologically, attention is a complex function that involves the collaborative effort of several brain sites. Furthermore, different forms of attention appear to invoke different brain systems. For example, sustained attention or vigilance is mediated by the reticular formation, a network of ascending and descending nerve cell bodies and fibers, which begins in the spinal cord and extends through the medulla all the way up to the thalamus. Specific nuclei within the reticular formation project through the thalamus to wide areas of the brain and thereby help mediate attention. Based upon the classic studies of Moruzzi and Magoun (**1949** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1181)**) demonstrating that ascending nerve tracts within the **reticular formation** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm01#bm01gloss279)** govern general

chronic drowsiness to stupor or coma (**Carpenter, 1991** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib271)**).

Selective attention appears to invoke brain sites in addition to the reticular formation. For example, based upon functional imaging studies that highlight active brain sites, it appears that the cingulate gyrus is essential for focusing upon relevant aspects of the environment while ignoring irrelevant information. One finding is that, when asked to perform complex attentional tasks, persons who suffer from schizophrenia and who, therefore, reveal deficits in selective attention also show dysfunction in the cingulate gyrus (**Carter, Mintun, Nichols, & Cohen, 1997** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib275)**).

10.5 MOTOR/COORDINATION SYSTEMS

Although many brain sites are involved in motor control, three areas are of special significance: the cerebellum, the basal ganglia, and the motor cortex. The **cerebellum** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm01#bm01gloss40)** sits just below the cerebrum at the back of the brain. Together with other brain structures, it helps coordinate muscle tone, posture, and hand and eye movements. Lesions in or near the cerebellum may render the individual incapable of making coordinated lateral eye movements. For this reason, neurologists and neuropsychologists commonly ask patients to demonstrate left-right and up-down eye movements. An individual with damage to the cerebellum might not be able to move his or her eyes with facility in all directions.

The cerebellum receives sensory information from every part of the body and coordinates the details of automatic skilled movements. Damage to the cerebellum may cause a variety of motor disturbances, depending upon the specific sites affected (**Manto & Pandolfo, 2002** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1040)**). Slurred, hesitant speech known as **dysarthria** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm01#bm01gloss97)** may be a symptom of cerebellar damage. Muscles may become flabby and tire easily. Rapid, coordinated tapping of the index finger may prove difficult. Measures of finger-tapping speed (**Reitan & Wolfson, 1993** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1350)**) are, therefore, an important component of neuropsychological test batteries.

Bodily movements may lose their coordination in cerebellar disease, becoming spasmodic and jerky. Even a simple gesture such as reaching for a cup may result in the inadvertent thrusting of cup and contents halfway across the room. The characteristic wide-based gait of alcoholics—called ataxia—is a consequence of cerebellar degeneration (**Ghez, 1991** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib583)**). Another symptom of cerebellar damage is intention tremor, so named because it is not present at rest but arises during voluntary, intentional movements of the hands. Nystagmus also is common in cerebellar disease. In this symptom, the eyes appear to jitter back and forth even when the individual attempts to hold a steady gaze.

In conjunction with the vestibular center in the inner ear, the cerebellum also helps coordinate the vestibuloocular reflex (VOR). The VOR acts to maintain the eyes on a fixed target when the head is rotated. Without the VOR, vision would be incredibly blurred whenever the head moved even a fraction of an inch. Instead, a small area of the cerebellum coordinates a rapid refixation of the eyes to compensate for head movements.

The **basal ganglia** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm01#bm01gloss23)** consist of a collection of nuclei in the in the forebrain that makes connections with the cerebral cortex above and the thalamus below. The basal ganglia are traditionally considered as part of the motor system. The main constituents of the basal ganglia are three large subcortical nuclei: the caudate, the putamen, and the globus pallidus. Some authorities also consider the amygdala to be part of the basal ganglia (**Carpenter, 1991** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib271)**). These structures are interconnected with and functionally related to the subthalamic nucleus and the substantia nigra. Along with the cerebellum, the corticospinal system, and the motor nuclei in the brain stem, the basal ganglia participate in the control of movement. Unlike the other components of the motor

system, the basal ganglia do not have direct connections with the spinal cord. The motor functions of the basal ganglia are indirect and are mediated via neural connections with the frontal cerebral cortex.

The most common syndrome caused by damage to the basal ganglia is **Parkinson's disease** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm01#bm01gloss235)** (PD) (**Factor & Weiner, 2008** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib479)**). In Parkinson's disease, three characteristic types of motor disturbances are observed: involuntary movement, including tremor; poverty and slowness of movement without paralysis; and changes in posture and muscle tone. In its later stages, this disease is typified by an immobile, masklike facial expression, an extreme difficulty initiating movements, and a fine tremor that may disappear once a movement is under way.

Patients with Parkinson's disease also reveal specific cognitive deficits, suggesting that the basal ganglia contribute not just to movement but to thinking as well. Deficits observed in these patients include problems formulating goals and evaluating progress, difficulties with attention, limitations in word-finding, and slowed thinking. Some patients with PD report that their brain feels "swampy" (**Tröster, 2012** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1667)**). A loss of spontaneity and a lack of initiative also are observed (**La Rue, 1992** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib936)**).

The **motor cortex** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm01#bm01gloss210)** is found on the precentral gyrus of the frontal lobe. Primary motor cells that subserve voluntary movement are located here and in adjoining brain sites. Motor control is substantially but not exclusively contralateral (opposite-sided), meaning that the left precentral gyrus subserves the right side of the body, and vice versa. Thus, when an individual makes a decision, say, to lift his right hand, motor neurons in the left precentral gyrus will be activated. For obvious reasons, this area is also known as the motor strip.

The fact that motor control is substantially opposite-sided is the basis for several neuropsychological procedures that compare the function of the two sides of the body as a means of determining the integrity of the left and right motor strips. Consider the finger-tapping test, employed with many neuropsychological test batteries (e.g., **Reitan & Wolfson, 1993** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1350)**). In a typical finger-tapping procedure, the examiner uses standardized procedures with repeated trials to determine the maximal tapping rates of the left and right index fingers over a 10-second span. Of course, the preferred hand will have a slight advantage, with a normative expectation of a rate that is 10 percent higher. For example, in a right-handed person, a tapping rate of 55 for the right index finger and 50 for the left index finger might be typical.

Any significant deviation from this expected pattern may suggest impairment in the opposite-sided motor strip. For example, suppose a right-handed examinee has a tapping rate of 47 for the right index finger and 50 for the left index finger. Because the right-sided tapping rate is comparatively slower than expected (i.e., 6 percent slower instead of 10 percent faster than the left-sided tapping rate), this would suggest impairment in the left motor strip.

10.6 MEMORY SYSTEMS

Although the lay public thinks of memory as a single thing, psychologists have known for more than a century that there are many types of memory and also several stages of memory (**Ebbinghaus, 1885/1913** (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib446)). We can provide only a cursory review here. The importance of reviewing these basic distinctions is that different brain systems may be involved in different kinds of memory.

As to types of memory, Andreasen (**2001** (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib51)) posits the existence of at least four different polarities of **memory** (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm01#bm01gloss196) : episodic versus semantic, working versus associative, declarative versus procedural, and explicit versus implicit. To this list, we would add a fifth dimension: short-term versus long-term memory. These dimensions are not completely separate and distinct from one another. Episodic memory refers to memory of events or experiences, such as recalling that you had oatmeal for breakfast. In contrast, semantic memory is general knowledge not tied to a specific learning experience, such as knowing that a butterfly is an insect, not a bird. Working memory is the retention of information that we need only briefly, such as remembering the digits of a phone number just long enough to complete the call. Associative memory involves memories that are invoked because of their association with particular cues, for example, recalling the smell and taste of popcorn when hearing the sound of it popping in the microwave. Declarative memory involves the "what" of memory (e.g., knowing that a bicycle has two wheels) whereas procedural memory involves the "how" of memory (e.g., knowing how to ride a bicycle). Another way of dividing memory is explicit versus implicit, which defines the difference between memories that are immediately accessible and obvious (e.g., knowing your name) compared to those that are latent, beneath the surface (e.g., surprising yourself when you are able to recall the name of your first-grade teacher).

Another important distinction is between short-term and long-term memory. Short-term memory is synonymous with working memory and is very short in duration, lasting from perhaps 10 seconds to a minute. If short-term memories are not "refreshed" through rehearsal, they disappear after this brief duration. Long-term memory refers to memories that have been consolidated in some way so that they are more lasting in duration—hours or years—although not necessarily permanent.

Describing the brain systems involved in memory is challenging because multiple brain sites are typically involved and different types of memory utilize different pathways. Even so, there is substantial evidence that structures within the temporal lobes are essential to many important features of memory. In particular, the **hippocampus** (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm01#bm01gloss145) and the **amygdala** (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm01#bm01gloss11) appear to be involved in various aspects of memory and learning. Specifically, these brain sites are involved in the consolidation of short-term memories into long-term memories. The amygdala may play a special role in integrating memories from different modalities and, especially, in consolidating memories with strong emotional meaning (**Andreasen, 2001** (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib51)).

Humans have both a left hippocampus and right hippocampus (plural: *hippocampi)*, located subcortically within the left and right temporal lobes. The same is true for the amygdala (plural: *amygdalae)*, which is

of the forward section of the temporal lobe on both sides of his brain (**Milner, 1968** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1163)**). Prior to this case, many individuals with epilepsy had been successfully treated by the removal of the diseased portion of one temporal lobe. The goal of this kind of surgery is to remove the diseased brain areas that serve as the "trigger" or focus point for seizure activity. The cognitive consequences of single-sided temporal lobe surgery had proved to be minimal. H.M. was the first carefully studied case of bilateral temporal lobe surgery.

The consequences of his surgery were devastating, which was a shocking revelation to everyone involved. Put simply, H.M. proved incapable of forming any new memories from the point of the surgery onward (**Milner, 1968** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1163)**). His old long-term memories remained intact, so he could recall where he attended high school, and so forth. And his short-term memory was intact, so he could remember a phone number briefly, for example. But his ability to consolidate new long-term memories was completely annihilated. He could read the same magazine from day to day, unaware that he had read it, cover to cover, the day before. A new doctor remained a new doctor on each new visit. He was essentially a prisoner of the moment, able to converse and interact with apparent normality but unable to remember anything new for more than a few minutes.

Structured testing of H.M. confirmed that different forms of memory are subserved by different brain systems. Consider procedural memory, for example, the recollection of how to do something. H.M. was asked to undertake repeated trials of mirror drawing—a complex procedural task in which the examinee traces a path on a sheet of paper while looking in a mirror. This is a daunting assignment in which directionality—left and right—are effectively reversed. With practice, normal individuals typically show slow improvement, tracing the path more quickly and with fewer errors. Intriguingly, H.M. likewise showed normal improvement on this task from day to day—indicating that his procedural memory remained intact—even though he had no realization that he had seen the puzzle before (**Corkin, 1968** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib350)**). Most likely, this kind of procedural memory is subserved by the cerebellum. Clearly, it is not underwritten by the temporal lobes.

10.7 LIMBIC SYSTEM

The **limbic system** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm01#bm01gloss188)** is a "primitive" central brain system that is involved in emotions and basic survival drives. This system overlaps with other brain sites, especially those involved in memory. The structures of the limbic system are involved in emotions, such as fear and aggression, as well as in the acquisition of memory. The pleasure centers of the brain are located here, too, within the nucleus acumbens. In addition to the hippocampus and amygdala, other limbic structures are the cingulate gyrus, mammillary bodies, and the fornix. Andreasen (**2001** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib51)**) points out that the exact boundaries of what constitutes the limbic system are not well established because our understanding of this brain system has been steadily growing.

In evolutionary terms, the limbic system is very old and, consequently, involved in primitive survival functions. Because of its proximity to and connections with the hypothalamus, the limbic system indirectly exerts autonomic nervous system control over crucial bodily functions needed for continued existence.

The **hypothalamus** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm01#bm01gloss148)** is a deceptively small structure that sits just below and in front of the thalamus. Even though it composes only about 0.3 percent of the brain's weight, the hypothalamus is involved in numerous aspects of motivated behavior and bodily regulation: blood pressure, feeding, sexual behavior, sleep/wake cycle, temperature regulation, emotional behavior, and movement. Well studied in lower animals, the functions of the hypothalamus are less well known in humans (**Kolb & Whishaw, 2011** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib912)**). It is known that the hypothalamus exerts proprietary control over the pituitary gland, thereby modulating a wide range of endocrine functions. The most common cause of a hypothalamic lesion is a severe head injury. Hypothalamic lesions often lead to disturbances of pituitary function, including excessive or deficient intake of food or water and temperature and blood pressure dysregulation (Kupfermann, 1991a). Dysfunction of the hypothalamus also can lead to emotional dysregulation (especially fear or rage) and sleep disturbance (hypersomnolence or insomnia).

10.8 LANGUAGE FUNCTIONS AND CEREBRAL LATERALIZATION

Language Functions of the Left Hemisphere

Language is primarily (but not exclusively) a left hemisphere function that involves widely separated cortical and subcortical structures. Because so many regions of the left hemisphere are involved in language, virtually any significant left hemisphere lesion will produce some kind of disturbance in the production or comprehension of language. For this reason a detailed profile of language skills offers a window to the integrity and functioning of the left hemisphere.

Yet, we need to keep in mind that virtually any high-level intellectual activity, including language expression and comprehension, requires the synthetic interaction of the entire brain. Speech is a case in point. While primarily subserved by the left hemisphere in most individuals, the right cerebral hemisphere does provide the intonation patterns for speech. As a result, patients with right-sided lesions (particularly in the frontal area) may speak in an eerie monotone (**Kalat, 2012** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib852)**).

Modern conceptions of brain-language correlations actually stem from the late nineteenth century. In 1861, Paul Broca observed that damage to a small region just in front of the motor cortex of the left hemisphere caused a language disorder originally called expressive aphasia and now more typically known as nonfluent aphasia. Persons with damage to this left hemisphere premotor area—aptly named Broca's area—speak in a slow, labored manner. They have difficulty enunciating words correctly; the act of speaking seems to be torturous for them. Speech takes on a frankly telegrammatic nature; adjectives, adverbs, articles, and conjunctions—the words that add color to speech—frequently are omitted. Writing also is difficult for these persons. Fortunately, persons who experience **Broca's aphasia** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm01#bm01gloss37)** have little difficulty understanding either spoken or written language. In its pure form, the disorder involves expressive language only.

In 1874, Wernicke announced that damage to the upper and rearward portion of the left temporal lobe—a region now known as Wernicke's area—was linked to a language disorder originally called receptive aphasia and now more typically known as fluent aphasia. Affected individuals appear unable to comprehend spoken or written language. Apparently, persons with **Wernicke's aphasia** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm01#bm01gloss350)** have no difficulty perceiving words but cannot associate the words with their underlying meaning. As a consequence, the written and verbal expressions of persons with this aphasia are fluent but meaningless. For example, when asked to define *book*, a patient might respond, "Book, a husbelt, a king of prepator, find it in front of a car ready to be directed." The same person might define *scarecrow* as, "We'll call that a three-minute resk witch, you'll find one in the country in three witches" (**Williams, 1979** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1759)**).

Building on the observations of Broca and Wernicke, Geschwind (**1972** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib580)**) proposed a structural, neurological model of left hemisphere language functions that has been highly influential in neuropsychological assessment. This model bears directly upon the assessment of language skills; the major elements are outlined next and depicted in **Figure 10.2** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/ch10lev1sec2#ch10fig2)** . Geschwind postulated the following:

- Spoken language is perceived in the left auditory cortex at the top of the temporal lobe and then transferred to Wernicke's area.
- In Wernicke's area, the meanings of words are activated and the auditory codes are transported to a subcortical bundle of transmission fibers called the arcuate fasciculus.
- The arcuate fasciculus sends the auditory codes directly to Broca's area.
- Upon reaching Broca's area, the auditory code activates the corresponding articulatory code that specifies the sequence of muscle actions required to pronounce a word.
- In turn, the articulatory code is transmitted to the portions of the motor cortex governing tongue, lips, larynx, and so forth in order to produce the desired spoken word.

Comprehending or speaking a written word involves most of the previously outlined pathways, but with a different starting point:

- Written words are first registered in the visual cortex, then relayed through the visual association cortex to the angular gyrus.
- In the angular gyrus, the visual form of the word is mapped into the auditory code stored in Wernicke's area, thereby gaining access to the meaning of the written word, which can also be spoken (steps 2 through 5 previously).

The Geschwind model is helpful in explaining a number of clinical syndromes caused by discrete left hemisphere brain damage (**Gregory, 1999 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib650)**):

- Lesions to Broca's area will cause slow, labored, telegraphic speech, but the comprehension of spoken or written language will not be affected.
- Damage to Wernicke's area will have more serious and pervasive implications for language comprehension; namely, the patient will be unable to understand spoken or written communications.
- Damage to the angular gyrus will cause serious reading disability, but there will be little problem in comprehending speech or in speaking.
- Impairment limited to the left auditory cortex will result in serious disruption of verbal comprehension. However, such persons will be able to speak and read normally.

In practice, few patients reveal aphasic symptoms that fall neatly into one or another of the preceding categories. Furthermore, modern conceptions of aphasia point to weaknesses in the classical model (e.g., its overly simplistic view of the structure of language) and propose a complex, nonlinear model of aphasia that is beyond the scope of coverage here (**Bonner, Ash, & Grossman, 2010 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib178)**). Nonetheless, a thorough assessment of language functions is an essential part of every neuropsychological evaluation and the classical model of Broca, Wernicke, and Geschwind provides a useful starting point. Additional perspectives on aphasia and the structural model of language can be found in Benson (**1994 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib140)**) and Mayeux and Kandel (**1991 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1077)**).

Specialized Functions of the Right Hemisphere

Based on thousands of studies of normal and braindamaged persons, it is now well established that the right hemisphere is dominant for a variety of cognitive and perceptual skills. However, a detailed discussion of specialized right hemisphere functions is beyond the scope of this section. Competent reviews of the extensive literature on this topic can be found in Bradshaw and Mattingley (**1995**

de Oliviera, and others (**2009 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib521)**), Springer and Deutsch (**1997 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1542)**), and Witelson (**2007 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1776)**). In general, the right hemisphere appears to be dominant for the analysis of geometric and visual space, the comprehension and expression of emotion, the processing of music and nonverbal environmental sounds, the production of nonverbal and spatial memories, and the tactual recognition of complex shapes.

A frequent symptom of right hemisphere damage is **constructional dyspraxia (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm01#bm01gloss65)** , the impaired ability to deal with spatial relationships either in a two- or three-dimensional framework (**Reitan & Wolfson, 1993 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1350)**). This symptom is commonly exhibited by an impaired ability to copy simple shapes such as a cross. Left hemisphere lesions can also cause constructional dyspraxia, but the correlation is less consistent. Most neuropsychological test batteries include one or more copying tasks to screen for constructional dyspraxia. We include a summary of findings on cerebral lateralization in **Table 10.2 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/ch10lev1sec9#ch10tbl2)** .

10.9 VISUAL SYSTEM

The primary sensory areas for vision are located in the occipital lobes; much of this projection area is on the mesial or midline surface that separates the two cerebral hemispheres. Each occipital lobe sees the opposite side of the visual world. Thus, all visual stimuli to the left of the reader's fixation point are ultimately processed in the right occipital lobe, and vice versa. The split visual world is shared across the splenium, the rearward portion of the corpus callosum, producing a unified perception of the entire visual field. Damage to the primary visual area produces a corresponding loss of visual field on the opposite side. For example, an extensive lesion in the left occipital lobe would render a person blind to the right half of the visual world. A very small lesion might produce a scotoma or blind spot.

TABLE 10.2 A Summary of Findings on Cerebral Lateralization

Functional System	***Left Hemisphere Dominance***	***Right Hemisphere Dominance***
Vision	Processing of the right visual field Recognition of letters, words	Processing of the left visual field Recognition of faces
Audition	Processing of right ear Processing of language-related sounds	Processing of left ear Processing of music and environmental sounds
Somatosensory	Sensory input from the right side	Sensory input from the left side
Movement	Motor output to the right side Complex voluntary movement, including speech	Motor output to the left side
Language	Speech, reading, writing, and arithmetic	Intonation and emotional patterning to speech
Memory	Verbal memory	Pictorial memory
Spatial processes		Analysis of geometric and visual space
Emotion		Comprehension and expression of emotion
Olfaction	Smell in left nostril	Smell in right nostril

The forward portion of each occipital lobe is unimodal association cortex. These regions synthesize visual stimuli and produce meaning from them. This is where the high-level processing of visual information occurs. Damage to the association cortex of the occipital lobes may cause **visual agnosia** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm01#bm01gloss349)** , a difficulty in the recognition of drawings, objects, or faces (**Kandel, 1991** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib856)**). Luria (**1973** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1019)**) described a typical case of a patient with such a lesion:

> The patient carefully examines the picture of a pair of spectacles shown to him. He is confused and does not know what the picture represents. He starts to guess. "There is a circle . . . and another circle . . . and a stick . . . a crossbar . . . why, it must be a bicycle?"

The visual agnosias are especially linked to right-sided lesions of occipital association cortex, but may

agnosia is prosopagnosia, the inability to recognize familiar faces. Benson (**1994 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib140)**) cites the example of a 70-year-old man who suffered a series of strokes affecting the forward portions of the occipital lobes. The patient's chief complaint was that he could not recognize his wife or his daughter by sight, although he immediately recognized them by their voices. In another case of visual agnosia known as object agnosia, a patient reproduced a drawing of a train with great skill but had no idea what he had drawn. Benson (**1988 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib139)**) describes the many fascinating symptoms of visual agnosia.

10.10 EXECUTIVE FUNCTIONS

The **executive functions** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm01#bm01gloss102)** of the brain provide the ability to respond to novel situations in an adaptive manner. Lezak, Howieson, and Loring **(2004 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib984)**) propose that the executive functions consist of four components:

- Volition
- Planning
- Purposive action
- Effective performance

Volition is the capacity for intentional behavior, the ability to conceptualize a goal. Planning is the identification of the steps needed to achieve the goal. Purposive action is the capacity to take action and sustain it in an orderly manner. Effective performance requires the ability to monitor one's activities in light of the original goals and shift strategies as needed. Thus, executive functions are implicated in a wide range of cognitive, emotional, and social skills.

An intriguing paradox of psychological testing is that few instruments are sensitive to impairments of executive functions. When provided with the structure of a typical psychological test, individuals with impaired executive functions often rise to the occasion and perform well. However, in the perplexity of real life, personal functioning may reveal catastrophic disability. For example, a successful financial planner who sustained a brain injury

> . . . can no longer formulate plans well because of an inability to take all aspects of a situation into account and integrate them. This disability is further aggravated by his lack of awareness of his mistakes. Problems occasioned by the man's emotional lability and proneness to irritability are overshadowed by the crises resulting from his efforts to carry out inappropriate and sometimes financially hazardous plans. (**Lezak, 1995** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib981)** , p. 650)

Yet, cognitive test scores for this individual—and others like him with impaired executive functions—might well be normal.

Executive functions are substantially but not exclusively underwritten by the frontal lobes. Although it is true that disturbances in executive functions can arise from a variety of neurological conditions that involve diverse brain sites, in the vast majority of cases damage to the frontal lobes is implicated. It is with the frontal lobes that humans create intentions, form plans, and regulate their behavior by comparing the effects of their actions with their original intentions. In short, the frontal lobes are essential for the programming, regulation, verification, and motor performance of executive functions.

Enacting a plan requires a bodily movement of some kind. People pursue their goals by physically manipulating the environment, whether with their hands or through the motor activity of speech. It is not surprising, then, to find that the primary motor cortex is located in the frontal lobes—where plans and intentions are also formed.

The primary motor cortex is found on the precentral gyrus, at the rear of the frontal lobe, just in front of

movements on the right, and vice versa. The topical organization of the motor strip was first mapped by Penfield (**1958** (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1278)) during a series of operations to remove damaged cortical tissue in persons with epilepsy. He stimulated different areas of the motor cortex with a harmless electrical current to map the correspondence between cortex and different body parts. Penfield found that those areas of the body requiring precise control, such as fingers and mouth, occupy a disproportionately large amount of cortical space.

Just in front of the primary motor cortex is the supplementary motor cortex. The supplementary motor cortex is involved in the serial ordering of complex motor chains, that is, movement programming. A portion of the frontal lobes just below the supplementary motor cortex is involved in the control of voluntary eye gaze. The left frontal lobe also mediates expressive language, discussed in detail later.

Damage to the primary motor cortex causes opposite-sided deficits in fine motor control and also reduces the speed and strength of limb movements. These effects are easily detected with simple motor tests such as finger-tapping speed. Severe damage to the motor cortex causes total paralysis of the affected bodily parts. Damage to the supplementary motor cortex causes deficits in the execution of motor sequences such as copying a series of arm or facial movements (**Kolb & Milner, 1981** (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib910)).

The most common cause of frontal lobe damage is closed head injury, which is one type of traumatic brain injury. In a closed head injury, acceleration/deceleration forces are instantly applied to the entire brain, as when a person's head strikes the dashboard in an automobile accident. Because of the irregular surfaces of the surrounding skull, the forward underside surfaces of the frontal lobes are almost always damaged (**Jennett & Teasdale, 1981** (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib826)).

The front ends of the temporal lobes also are highly vulnerable in closed head injury.

Nauta (**1971** (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1218)) summarizes the effects of frontal lobe dysfunction as a "derangement of behavioral programming." Lezak (**1983** (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib980) , **1995** (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib981)) has catalogued the behavioral disturbances that can result from generalized, bilateral frontal lobe damage:

- Motivational-like problems involving decreased spontaneity, decreased productivity, reduced rate of behavior, and lack of initiative
- Difficulties in making mental shifts and perseveration of activities and responses
- Problems in stopping that are often described as impulsivity, overreactivity, and difficulty in holding back a wrong or unwanted response
- Deficits in self-awareness resulting in an inability to perceive performance errors or to size up social situations appropriately
- A concrete attitude (**Goldstein, 1944** (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib614)) in which objects, experiences, and behavior are all taken at their most obvious face value

Curiously, frontal lobe lesions may have little effect on old learning and well-established skills. Both Hebb and Penfield reported that surgical removal of frontal lobe tissue caused little change in IQ scores (**Hebb, 1939** (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib729) ; **Penfield & Evans, 1935** (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1279)). Early

studies of prefrontal lobotomy demonstrated much the same finding: no change in IQ or even a slight improvement after disconnection of the frontal lobes.

Devising adequate measures of frontal lobe function has proved to be difficult. Lezak et al. (**2004 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib984)**) note that frontal lobe disorders change how a person responds, whereas most tests measure what a person knows. Lezak (**1982 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib979)**) has devised an ingenious method called the Tinkertoy® Test, discussed in the next topic, to assess the programming difficulties experienced by persons with frontal lobe lesions. More commonly, clinicians rely upon observation and checklists to diagnose frontal lobe dysfunction. A generic example of a checklist for executive functions is provided in **Figure 10.3 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/ch10lev1sec10#ch10fig3)** .

Awareness						
Is unaware of limitations	1	2	3	4	5	Has insight into limitations
Goal Selection						
Sets no goals	1	2	3	4	5	Sets suitable long-term goals
Logical Analysis						
Is disorganized	1	2	3	4	5	Plans thought-fully
Action Orientation						
Needs prompting	1	2	3	4	5	Takes decisive action
Self-Monitoring						
Is unable to identify errors	1	2	3	4	5	Detects and corrects mistakes
Impulse Control						
Is highly impulsive	1	2	3	4	5	Thinks before acting
Flexibility						
Is inflexible in approach	1	2	3	4	5	Learns from feedback

1 = profoundly deficient
2 = severely deficient
3 = moderately deficient
4 = mildly deficient
5 = normal

FIGURE 10.3 Example of a Structured Checklist for the Assessment of Executive Functions

10.11 NEUROPATHOLOGY OF ADULTHOOD AND AGING

Although most individuals age gracefully and maintain good health into old age, an unfortunate minority experience one or more neurological syndromes such as brain injury, dementia, or Parkinson's disease. In this section we provide a brief synopsis of a number of more common neurological problems encountered in adulthood and old age. Because neuropsychological tests excel in the evaluation of these syndromes, a brief survey will provide an important backdrop to the selected instruments discussed in the second half of the chapter.

Traumatic Brain Injury

Traumatic brain injury or TBI is an inclusive term that encompasses everything from a "mild" concussion to severe brain injury (**Silver, McAllister, & Yudofsky, 2011 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1498)**). TBI is most commonly the consequence of a blow to the head, and concussion is probably the most common form of TBI. The classic example of a concussion is the football player who receives a hard hit ("sees stars"), is rendered briefly unconscious and immobile, and then gradually walks off the field with assistance. Within hours or a few days, he is back to normal. The symptoms of **concussion (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm01#bm01gloss61)** include a brief loss of consciousness followed by a low-grade headache, difficulty concentrating, fatigue, irritability, and other emotional symptoms. Although some concussions can have serious, lasting effects, most patients appear to make a full recovery in a few days or weeks. A concussion is one example of a closed head injury (CHI)—a trauma to the head and brain in which the skull remains intact. But *closed head injury* is a broader term than *concussion* and potentially signifies a greater level of impairment than typically found in a concussion. Closed head injury is often contrasted with open head injury or OHI—a trauma to the head and brain in which the skull is penetrated. OHI is also known as penetrating head injury. Typically, the consequences of OHI are focal or localized in and near the site of impact, whereas the effects of CHI are more diffuse, affecting areas throughout the brain.

The neurological consequences of TBI depend upon the nature and severity of the injury, but any or all of the following are possible:

- a contusion or bruising of the brain underneath the site of impact known as a coup injury
- a contusion opposite the side of the impact, caused by rebound, and known as a contrecoup injury
- frequent contusions in the undersurfaces of the frontal lobes and the tips of the temporal lobes because of the bony skull protrusions located there
- diffuse axonal injury or nonspecific brain cell damage from shear-strain effects on neural pathways
- brain tissue damage due to obstructed blood flow when cerebral arteries are ruptured
- hematoma or bleeding into the brain between the skull and the surface of the brain
- edema or swelling of the brain, which can lead to secondary brain damage
- in the long term, possible shrinkage of the brain and enlargement of the ventricular system

As to the neurobehavioral effects of TBI, the most common and reliable complaints are of concentration and memory problems. This is why tests of concentration and memory are found in virtually every test battery used in neuropsychological assessment. Other generalizations about TBI are difficult because the nature and severity of the brain damage will not be the same in any two patients. Focal damage may lead

Many studies suggest that TBI patients are more seriously handicapped by personality and emotional disturbances than by cognitive and physical disabilities (**Lezak & O'Brien, 1990** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib982)**).

Modern warfare constitutes a major source of TBI cases. Beginning just after the stunning and devastating attacks of September 11, 2001, more than two million U.S. troops have been deployed to Afghanistan and Iraq. Almost half of these soldiers have been deployed more than once, totaling in excess of three million tours of duty *(Marine Corps Times*, December 18, 2009). In these contemporary war theatres, blast injuries from roadside bombs known as improvised explosive devices (IEDs) comprise a common source of TBI. The detonation of an IED produces a pressure shock wave that reverberates through the brain and body, often causing neuronal changes that include diffuse axonal injury. TBI from these deadly devices is recognized as the "signature injury" of the wars in Afghanistan and Iraq (**Dixon, 2011** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib421)**). Even a "mild" blast can produce subtle deficits that are difficult to detect and measure.

The prevalence of troop exposure to IED blasts is not well appreciated by the public. In a study of 2,525 U.S. Army infantry soldiers conducted three to four months after a year-long deployment to Iraq (**Hoge, McGurk, Thomas, and others, 2008** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib759)**), *fully 62 percent* of the sample reported that an IED had exploded near them *on two or more* occasions! From the large subsample of IED-exposed soldiers *(N* = 1,556), 7 percent reported an injury with loss of consciousness, 15 percent told of injury with altered mental status, and 18 percent reported other injury. Emotional and health consequences likewise were common, with many troops demonstrating Post-Traumatic Stress Disorder (PTSD), depression, and health problems such as stomach pain, headache, fatigue, and sleep disturbance. Overall, 15 percent of the original sample met the criteria for mild TBI (mTBI). The presence of mTBI was especially correlated with IED blasts that caused a loss of consciousness.

Neoplastic Disease (Tumor)

Neoplastic disease or brain tumor encompasses many different forms of tumorous growth (**Reitan & Wolfson, 1993** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1350)**). For example, gliomas are tendrillike tumors of the glial cells that infiltrate the brain over a period of weeks or months; meningiomas are slower-growing, globular-shaped tumors of the meninges (membranes encasing the brain) that press down upon the brain.

Brain tumors produce a variety of effects, depending upon their location, size, and rate of growth. A rapidly infiltrating tumor such as a glioma quickly may compromise many skills. For example, if the tumor is on the left side of the brain, motor and sensory functions on the right side of the body may be severely impacted, as well as language and problem-solving abilities. If the tumor is on the right side of the brain, constructional abilities (e.g., drawing, assembling three-dimensional objects) will be impaired as well as motor and sensory functions on the left side. A slower-growing meningioma may produce no symptoms for years and then create focal symptoms that relate to the site of encroachment on the brain. For example, if the right parietal area is affected, deficits in spatial ability may be observed.

Chronic Alcohol Abuse

Chronic alcohol ingestion leads to neuronal changes that include a loss of dendritic branches and dendritic spines, especially in areas important for memory such as the hippocampus. Over time, enlargement of the ventricles and widening of the cerebral sulci also are observed. In severe cases, atrophy of the medial thalamus and mamillary bodies is found, leading to the pronounced memory

problems that characterize Wernicke-Korsakoff's syndrome (Davila, Shear, Lane, Sullivan, & Pfefferbaum, 1994). The neuropathology of alcoholism often is exacerbated by vitamin and nutritional deficiencies.

In those tragic cases of severe alcohol abuse in which the medial thalamus and mamillary bodies are damaged, the profound anterograde amnesia of Wernicke-Korsakoff's syndrome is noted. Patients show an inability to retain memory of events for more than a short time even though immediate memory is intact and remote memory is only mildly impaired. The falsification of memory known as confabulation, in the presence of clear consciousness, is noted. Other symptoms of severe abuse include gait disturbance and gaze difficulties. In neurologically intact alcoholics, neurobehavioral effects are more elusive and controversial but may include subtle memory deficits and difficulties with novel problem solving (e.g., Waugh, Jackson, Fox, Hawke, & Tuck, 1989).

Recent research indicates that the brain changes and neurocognitive impairments caused by prolonged alcohol abuse can be partially reversed. A common problem observed in chronic alcoholics is, literally, shrinkage of brain tissue and enlargement of the ventricles. The ventricles are fluid-filled caverns at the center of the brain. The relationship is linear, with greater alcohol intake predicting greater brain shrinkage and larger ventricular enlargement (Anstey, Jorm, Reglade-Méslin, and others, 2006). Using sophisticated imaging techniques, Bartsch, Homola, Biller, and others (**2007 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib105)**) studied longitudinal changes in brain volume in 15 alcoholics and 10 matched controls. After 6-7 weeks of abstinence, the alcoholics revealed a 2 percent gain in volume of brain tissue, compared to no change among the controls. While a 2 percent improvement may not seem like much, it could foretell even more dramatic gains with long-term abstinence. The common metric among substance abuse professionals is that full cognitive recovery takes at least a year. In the Bartsch et al. study (**2007 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib105)**), pretest versus posttest scores on the d2-test, a measure of attention and concentration, also improved in the recovery group but showed no change in the control group. Several other studies confirm improvement in neuropsy-chological test results after abstinence in recovering alcoholics, as summarized by Walker (**2006 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1705)**).

Normal Pressure Hydrocephalus

Hydrocephalus is a build-up of cerebral spinal fluid (CSF) inside the skull, which causes brain swelling. In normal pressure hydrocephalus (NPH), which mainly affects individuals aged 60 or older, there is an increase in CSF, but the pressure of the fluid remains normal. Even so, brain function is affected, leading to a classic triad of symptoms: gait ataxia, incontinence, and dementia. Conn (**2011 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib337)**) describes his own case of NPH from a unique perspective (he is a physician) and suggests that many cases of dementia caused by NPH are misdiagnosed with potentially tragic consequences. NPH is highly treatable, whereas other forms of dementia resist intervention. His story is a warning against complacency and fatalism among health care workers who deal with assessment and diagnosis, including psychologists. His case of NPH

> . . . began in about 1992 as a trivial abnormality of gait that was misdiagnosed as Parkinson's disease (PD). Over the next 10 years, during which I was being unsuccessfully treated with dopaminergic drugs for PD, the illness gradually progressed until I could barely walk with a walking frame, had become incontinent of urine and, sometimes, faeces and began to show signs of cognitive loss. In the process of obtaining a motorised wheelchair I was referred to a younger neurologist who recognised that I had run the whole classic course of NPH, a disease of which I had never heard. I

> had a ventriculoperitoneal shunt (VPS) implanted in 2003 and was miraculously restored virtually to normal (p. 162).

A VPS shunt is a catheter extending beneath the skin from the ventricles of the brain to the abdominal cavity, allowing excess CSF to drain off.

The prevalence of NPH is difficult to ascertain because it resembles other forms of diffuse dementia. Many cases likely are overlooked. Based on his evaluation of published studies, Conn (**2011** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib337)**) estimates that 1 percent of the population will develop NPH by the age of 80.

Alzheimer's Disease

The most common degenerative neurological disease is **Alzheimer's disease** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm01#bm01gloss09)** (AD), which features an insidious degeneration of the brain. The pathophysiology includes clumplike deposits in the brain known as neuritic plaques and neurofibrillary tangles (Koss, 1994). Additional brain changes include neuronal loss, shrinkage or atrophy of the brain, depletion of acetylcholine neurotransmitters involved in memory, and accumulation of foreign deposits in the cerebral vasculature; the course of the disease invariably is downhill. First described in 1907, Alois Alzheimer portrayed his initial case as follows:

> The first noticeable symptom of illness shown by this 51-year-old woman was suspicious-ness of her husband. Soon, a rapidly increasing memory impairment became evident; she could no longer orient herself in her own dwelling, dragged objects here and there and hid them, and at times, believing that people were out to murder her, started to scream loudly. On observation at the institution, her entire demeanor bears the stamp of utter bewilderment. She is completely disoriented to time and place. (**La Rue, 1992** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib936)**)

Although Alzheimer's disease is not part of normal aging, advanced age is an important risk factor. Rare before age 65, the disease afflicts 3 percent of persons 65 to 74 years of age, 18 percent of persons 75 to 84 years of age, and nearly half of those 85 years and older (**Evans, Funkenstein, Albert, and others, 1989** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib466)**). Symptoms and examples suggestive of Alzheimer's disease are listed in **Table 10.3** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/ch10lev1sec11#ch10tbl3)** . These examples characterize other forms of dementia as well.

TABLE 10.3 General Symptoms and Specific Examples Suggestive of Alzheimer's Disease

Significant memory problems that extend beyond benign forgetfulness	
Fails to recall what was eaten for breakfast	
Difficulty with everyday tasks and commonplace activities	
No longer balances the checkbook, prepares the same meal	
Loss of orientation to date, time and/or place	
Significantly off as to date or time, loses the way going home	
Gradual and insidious onset	
Onset is hard to identify, problem is recognized in retrospect	
Language and word finding difficulties	
Conversation characterized by circumlocution and vagueness	
Problems with abstract thinking	
Difficulty following the rules of simple card games Deterioration of social judgment	
Dresses inappropriately, neglects personal hygiene	
Misplaces or loses important items	
Car keys disappear, eyeglasses are found in a kitchen drawer	
Changes in Personality:	
Onset of suspiciousness, periods of agitation, mood changes	
Loss of Initiative	
Absence of self-initiation, needs prompting to become involved	

Note: These examples characterize other forms of dementia as well.
Source: A synthesis based on Alzheimer's disease websites.

As detailed by Storandt and Hill (**1989 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1583)**), difficulty with the acquisition of new information (short-term memory dysfunction) is generally the most salient symptom in the early stages. As the disease progresses, patients may also show a prominent language dysfunction (e.g., pronounced word finding difficulty) or a striking visuospatial disturbance. Reports of personality change, including delusions and agitation, also are common. The late stages are characterized by severe, pervasive disability.

Vascular Dementia (Stroke)

The second most common cause of dementia in the elderly is vascular dementia, caused by blockage of an artery and subsequent death of brain tissue due to insufficient blood supply (infarction) or bleeding into or around the brain (hemorrhage). Sudden onset is the rule, but the accumulation of small strokes over time, known as multi-infarct dementia (MID), may produce an apparently progressive disorder. The Hachinski Ischemic Score was developed to distinguish multi-infarct dementia from Alzheimer's disease **(Hachinski, Iliff, Zilha, and others, 1975 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib680)**). Using this index, MID is indicated by the presence of several of the following factors: abrupt onset, somatic complaints,

confusion, history of strokes, personality preserved, atherosclerosis present, depression, and focal neurological signs. Because MID may be treatable to some degree, the differential diagnosis of MID versus Alzheimer's disease is more than academic.

The stroke syndrome is defined by the acute onset of a focal deficit involving the central nervous system. The specific symptoms depend upon the site of infarction but may include motor weakness and impaired sensibility in the limbs on the opposite side; nonfluent aphasia if the dominant hemisphere is affected; partial loss of the visual field if the stroke occurs in the rear of the brain. The acute symptoms of stroke often subside in some measure and lead to a plateau of stable functioning.

Parkinson's Disease (PD)

Parkinson's disease (PD) is almost nonexistent before age 40 and affects only 1 or 2 in 1,000 persons ages 70 and over (**La Rue, 1992 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib936)**). Primarily identified as a movement disorder, cognitive and emotional problems are common in PD. In fact, the late stages of PD may entail a clear dementia. The symptoms include slowness of movement (bradykinesia), tremor at rest, shuffling gait, and postural rigidity. The neuropathology includes depletion of dopamine and neuron loss in the basal ganglia.

Tremor is the most common and the least debilitating early symptom in PD. The rate of progression is quite variable, but movement disability in PD can become pronounced and lead to confinement; 10 to 20 percent of PD patients develop a clear dementia. Patients with PD reveal a deficit on neuropsychological tests requiring speed (e.g., Digit Symbol, Trail Making, reaction time measures). Surprisingly, tests of visual discrimination and paired-associate learning—which do not require speed—also differentiate patients with moderate to severe PD from matched controls (**Pirozzolo, Hansch, Mortimer, Webster, & Kuskowski, 1982 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1306)**). About 40 to 60 percent of PD patients also experience depression (**La Rue, 1992 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib936)**).

10.12 BEHAVIORAL ASSESSMENT OF NEUROPATHOLOGY

Psychological testing can be essential in the evaluation of neuropathology, as we will see in the next topic. Yet, it is easy for psychologists to become enamored of tests and to overlook the value of simple observation, interview, and behavioral evaluation. In medicine, the field of behavioral neurology has recognized the merit of these straightforward approaches for at least 150 years, dating back to the pioneering observations of Paul Broca and Carl Wernicke on syndromes of aphasia (Pincus & Tucker, 2003). Psychologists make use of this long-established tradition when they conduct a mental status examination at the beginning of assessment (**Sonne, 2012 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1522)**).

Assessment of Mental Status

The mental status examination (MSE) is a loosely structured interview that usually precedes other forms of psychological and medical assessment. The purpose of the evaluation is to provide an accurate description of the patient's functioning in the realms of orientation, memory, thought, feeling, and judgment. The MSE is the psychological equivalent of the general physical examination: Just as the physician reviews all the major organ systems, looking for evidence of disease, the psychologist reviews the major categories of personal and intellectual functioning, looking for signs and symptoms of psychopathology (**Gregory, 1999 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib650)**). Although there is some latitude as to the scope of the MSE, certain mental functions are almost always investigated. A typical evaluation touches upon the areas listed in **Table 10.4 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/ch10lev1sec12#ch10tbl4)** .

TABLE 10.4 Major Areas of a Typical Mental Status Exam

Appearance and Behavior

Grooming

Facial

expressions

Gross motor behavior

Eye contact

Speech and Communication Processes

Speech content, rate, tone, volume

Word difficulty, confusion, misuse

Thought content

Logic, clarity, appropriateness

Delusions

Cognitive and Memory Functioning

Calculating ability

Immediate recall

Recent and remote memory

Fund of information

Abstracting ability

Emotional Functioning

Predominant mood

Appropriateness of affect

Insight and Judgment

Awareness of problems

Orientation

Day, date, time, location

Source: Based on Gregory, R. J. (**1999 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib650)**). *Foundations of intellectual assessment: The WAIS-III and other tests in clinical practice.* Boston: Allyn and Bacon.

Some of the elements in this list can be assessed with short screening tests. In particular, cognition, memory, and orientation are intellectual functions that can be tested in a formal, structured manner (Hodges, 1994). These measures are most commonly used in the mental status evaluation of the elderly, especially when the client appears to have a dementia such as Alzheimer's disease, as discussed later in this chapter. Formal tests of mental status are also helpful in the assessment of certain brain-impairing conditions such as head injury, schizophrenia, severe depression, and drug-induced delirium. It is important to emphasize that screening tests are supplementary—they do not replace clinical judgment in

example, the evaluation of a patient's insight requires keen observation and sensitive interviewing skills. An MSE screening test for insight does not exist.

Behavioral Rating Scales

Another approach in the behavioral tradition is to utilize observations from persons familiar with the patient, such as a spouse, parent, close friend, or caretaker. Asking them questions about the patient is a good starting point. But a more efficient method is to employ a relevant behavior rating scale tied to the specific behaviors of the individual. This allows for reliable assessment and provides access to normative data. Hundreds of behaviorally based scales exist (**Tate, 2010 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1614)**). These can be broad-spectrum (such as establishing the likelihood of dementia) or narrow in focus (such as verifying the presence of the syndrome of disinhibition). For purposes of illustration, we will summarize two instruments here, one for the evaluation of dementia in general, and another for the appraisal of specific frontal lobe syndromes.

The Behavioral and Psychological Assessment of Dementia (BPAD) is a proxy-report rating scale designed to assess dementia-related changes in behavior among adults 30 years of age and older (**Schmidt & Gallo, 2007 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1467)**). In completing the BPAD, the informant rates the client on 78 items *Within the past four weeks (current)*, and also *five years ago (past).* Items are rated on a four-point scale. The BPAD assesses the symptoms for each of the two time periods (current and past) and also computes a change score. The change score reflects changes in mood and behavior that might signal the onset of dementia. Thus, three sets of scores emerge: Current, Past, and Change.

For each of the three sets of scores, the BPAD yields a total score and seven domain scores. All scores are reported as T-scores with a mean of 50 and standard deviation of 10, relative to the standardization sample. The test was standardized and validated on a large sample of men and women 30 to 90 years of age. The sample was matched to U.S. census proportions in regard to racial/ethnic makeup, educational backgrounds, and geographic regions.

The seven domains of the test are grouped into three clusters, as follows:

- Psychopathological Symptom Cluster
- Perceptual Delusions
- Positive Mood/Anxiety
- Negative Mood/Anxiety
- Behavioral Symptom Cluster
- Aggressive
- Perseverative/Rigid
- Disinhibited
- Biological Symptom Cluster
- Biological Rhythms

The instrument also yields a total score based on the sum of all seven domains. The BPAD items are at a grade 6 reading level. The test can be used in a variety of settings (inpatient, outpatient, assisted living) with patients suspected of having Alzheimer's disease, vascular dementia, and psychiatric problems.

The BPAD is a promising test, but there is scant validity research at this time. Certainly the domains

and psychological symptoms of dementia. For example, a prominent international group provides the following authoritative statement on the behavioral manifestations of dementia:

- *Behavioral symptoms:* Usually identified on the basis of observation of the patient, including physical aggression, screaming, restlessness, agitation, wandering, culturally inappropriate behaviors, sexual disinhibition, hoarding, cursing and shadowing.
- *Psychological symptoms:* Usually and mainly assessed on the basis of interviews with patients and relatives; these symptoms include anxiety, depressive mood, hallucinations and delusions. A psychosis of Alzheimer's disease has been accepted since the 1999 conference (**International Psychogeriatric Association, 2002 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib807)**).

Although terminology is not identical, the BPAD domains possess a clear commonality with the above description of dementia.

A test that embodies a more specific application is the Frontal Systems Behavior Scale (FrSBe) (**Grace & Malloy, 2001 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib636)**). The purpose of this instrument is to provide a behaviorally oriented assessment of three frontal lobe syndromes: apathy, disinhibition, and executive dysfunction. The scale consists of 46 items rated on a 5-point Likert scale by either the patient or a family member. Results from a family member are considered more reliable and valid. Items are written at a 6th grade level. Separate norms are provided for the patient and family form. The scale also attempts to quantify behavioral changes over time by including a baseline (retrospective) and a current assessment. A highly desirable feature of the form is that it takes only 10 minutes to administer and 10-15 minutes to score.

The subscales include Apathy (14 items), Disinhibition (15 items), and Executive Dysfunction (17 items), which are reported as T-scores (mean of 50, *SD* of 10) derived from a community-based sample of 436 men and women with two levels of education. Comparison data also are provided for several clinical groups: frontotemporal dementia, frontal lesions, nonfrontal stroke, head injury, Alzheimer's disease, Parkinson's disease, and Huntington's disease.

The construct validity of the FrSBe is firmly upheld by an exploratory factor analytic study of results for 324 neurological patients and research participants, the majority diagnosed with neurodegenerative disorders such as Alzheimer's, Parkinson's, and Huntington's disease (**Stout, Ready, Grace, Malloy, & Paulsen, 2006 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1584)**). The three-factor solution revealed that 83 percent of the items from the Apathy, Disinhibition, and Executive Dysfunction scales loaded prominently on the corresponding factors from the analysis. These results highly support the utility of the scale in assessment of the three frontal syndromes.

In a study of 66 individuals with a history of traumatic brain injury, Reid-Arndt, Nehl, and Hinkebein (**2007 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1346)**) found that the FrSBe was a better predictor of community integration than neuropsychological tests. Mendez, Licht, and Saul (**2008 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1129)**) reported that the scale differentiates patients with frontotemporal dementia (FTD) from those with Alzheimer's disease (AD) and vascular dementia (VaD). Specifically, the FTD patients had significantly greater scores on Disinhibition than the AD patients and the VaD patients. Chiaravalloti and DeLuca (**2003 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib301)**) testify that the FrSBe is sensitive to the behavioral changes observed in patients with Multiple Sclerosis. In sum, this simple, brief scale is an excellent measure for use with patients who display frontal lobe manifestations

TOPIC 10B Neuropsychological Tests, Batteries, and Screening Tools

The purpose of this topic is to review a diverse collection of neuropsychological tests, batteries, and screening tools. We focus here on representative tests, prominent batteries, and useful screening tools, recognizing that comprehensive coverage is well beyond the scope of the book. For a complete treatment of neuropsychological assessment, the reader is referred to the authoritative tome amassed by Lezak, Howieson, Bigler, and Tranel (**2012** (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib983)), which runs to an amazing 1,200 pages in length. By necessity, the coverage here is more discerning and emphasizes better-known tests and batteries.

Neuropsychologists and other clinicians often encounter clients who struggle with alcoholism or other types of substance abuse. For this reason, we also review a few simple but practical tools for rapid screening of clients with possible alcohol problems. This issue is vital because, at any given time, 10 percent of the adult population manifests an alcohol disorder (**Yalisove, 2004** (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1793)). Although it might

how often they drink—in reality this is a vexing diagnostic challenge due to the active façade of denial maintained by most alcoholics. However, a number of screening tools summarized later are useful for this task.

Finally, it is important to emphasize that neuropsychological assessment involves more than the administration and scoring of specialized tests and screening tools. An essential component of any assessment is the evaluation of a client's mental status. This is particularly true with elderly clients who may experience Alzheimer's disease or other forms of dementia. Accordingly, we close this chapter with a focus upon mental status assessment in the elderly. In this concluding topic, we pay special attention to the Mini-Mental Status Exam (**Tombaugh, McDowell, Kristjansson, & Hubley, 1996 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1654)**), one of the most widely used screening tools in existence.

Neuropsychological tests and procedures encompass an eclectic assortment of methods and purposes. At one end of the spectrum are simple, 10-minute screening tests used to probe the need for further assessment. At the other end of the spectrum are exhaustive, six-hour test batteries designed to provide a comprehensive assessment. In between are hundreds of specialized instruments developed to measure particular neuropsychological abilities. At first glance, this multitude of tests would appear to resist simple categorization, as if researchers in this area had followed an incoherent philosophy of trial and error in the development of new instruments and procedures. However, with closer scrutiny it is evident that most neuropsychological tests fit within a simple, logical model of brain-behavior relationships. We will use this model as a framework for discussing well-known neuropsychological tests and procedures.

10.13 A CONCEPTUAL MODEL OF BRAIN-BEHAVIOR RELATIONSHIPS

Bennett (**1988** (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib136)) has proposed a simplified model of brain-behavior relationships that is helpful in organizing the seemingly chaotic profusion of neuropsychological tests (**Figure 10.4** (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/ch10lev1sec13#ch10fig4)). His conceptualization is a slight expansion of the model presented by Reitan and Wolfson (**1993** (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1350)). According to this view, each neuropsychological test or procedure evaluates one or more of the following categories:

- Sensory input
- Attention and concentration
- Learning and memory
- Language
- Spatial and manipulatory ability
- Executive functions:
  - Logical analysis
  - Concept formation
  - Reasoning
  - Planning
  - Flexibility of thinking
- Motor output

The order of the categories listed corresponds roughly to the order in which incoming information is analyzed by the brain in preparation for a response or motor output.

In the remainder of this topic, the discussion of neuropsychological tests and procedures is organized around these seven categories. Within each category we will review established tests and also introduce new instruments that show promise of extending the horizons of neuropsychological assessment. However, the reader needs to know that neuropsychological assessment commonly involves a battery of tests. One approach is flexible or patient-centered testing in which an individualized test battery is fashioned for each client. These batteries are based upon the presenting complaints, referral issues, and an initial assessment (**Kane, 1991** (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib859) ; **Larrabee, 2008** (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib951)). More typically, neuropsychologists employ a fixed battery of tests for most referrals. One of the most widely used fixed batteries, the Halstead-Reitan Neuropsychological Battery, is outlined in **Table 10.5** (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/ch10lev1sec14#ch10tbl5) . Even though the HRNB is an old test—the elements of the battery have not been changed since its inception in the 1950s—many neuropsychologists still regard this battery as the "gold standard" in the field (**Horton, 2008** (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib789) ; **Sweeney, et al., 2007** (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1605)). In large measure, this is because of the steadily accumulating body of affirming research on the battery, which includes 267 publications by its developer, Ralph Reitan, and literally hundreds of additional articles from the dozens of neuropsychologists mentored by him. Yet, the HRNB is not without competition. The chapter closes with a presentation of two other batteries, namely, the Neuropsychological Assessment Battery and the Luria-Nebraska Neuropsychological Battery.

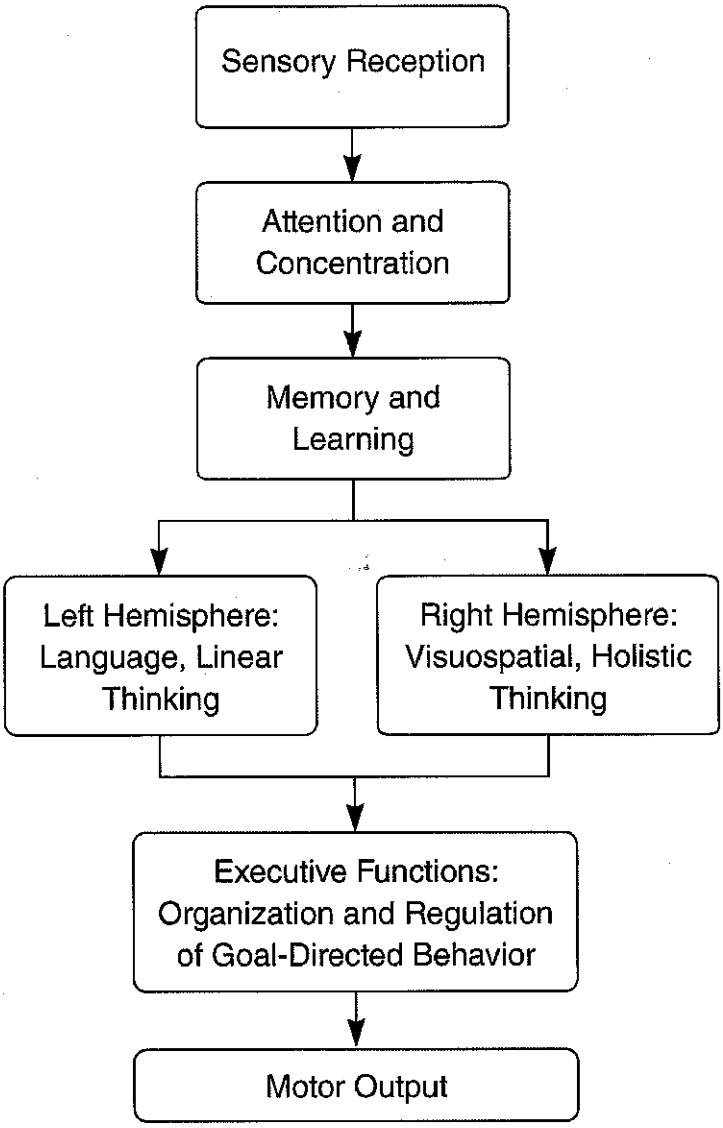


FIGURE 10.4 Conceptual Model of Brain Behavior Relationships

Source: Adapted with permission from Reitan and Wolfson (**1993** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1350)**).

10.14 ASSESSMENT OF SENSORY INPUT

The accuracy of sensory input is crucial to the proficiency of perception, thought, plans, and action. An individual who does not see stimuli correctly, hear sounds accurately, or process touch reliably may encounter additional handicaps at higher levels of perception and cognition. Neuropsychological assessment always incorporates a multimodal examination of sensory capacities.

TABLE 10.5 Tests and Procedures of the Halstead-Reitan Test Battery

Test	*Description*
Category Test* **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/ch10lev1sec14#ch10fn5)**	Measures abstract reasoning and concept formation; requires examinee to find the rule for categorizing pictures of geometric shapes
Tactual Performance Test* **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/ch10lev1sec14#ch10fn5)**	Measures kinesthetic and sensorimotor ability; requires blindfolded examinee to place blocks in appropriate cutout on an upright board with dominant hand, then nondominant hand, then both hands; also tests for incidental memory of blocks

Test	*Description*
Speech Sounds Perception Test* **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/ch10lev1sec14#ch10fn5)**	Measures attention and auditory-visual synthesis; requires examinee to pick from four choices the written version of taped nonsense words
Seashore Rhythm Test* **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/ch10lev1sec14#ch10fn5)**	Measures attention and auditory perception; requires examinee to indicate whether paired musical rhythms are same or different
Finger Tapping Test* **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/ch10lev1sec14#ch10fn5)**	Measures motor speed; requires examinee to tap a telegraph keylike lever as quickly as possible for 10 seconds
Grip Strength	Measures grip strength with dynamometer; requires examinee to squeeze as hard as possible; separate trials with each hand

Test	*Description*
Trail Making, parts A, B	Measures scanning ability, mental flexibility, and speed; requires examinee to connect numbers (part A) or numbers and letters in alternating order (part B) with a pencil line under pressure of time
Tactile Form Recognition	Measures sensory-perceptual ability; requires examinee to recognize simple shapes (e.g., triangle) placed in the palm of the hand
Sensory-Perceptual Exam	Measures sensory-perceptual ability; requires examinee to respond to simple bilateral sensory tasks, e.g., detecting which finger has been touched, which ear has received a brief sound; assesses the visual fields

Test	*Description*
Aphasia Screening Test	Measures expressive and receptive language abilities; tasks include naming a pictured item (e.g., fork) repeating short phrases; copying tasks (not a measure of aphasia) included here for historical reasons
Supplementary	WAIS-III, WRAT-3, MMPI-2, memory tests such as Wechsler Memory Scale-III or Rey Auditory Verbal Learning Test

*Strictly speaking, these five measures constitute the Halstead-Reitan Test Battery. However, in common parlance reference to the Halstead-Reitan includes all of the measures listed in the table.

Sensory-Perceptual Exam

The procedures developed by Reitan and Klove are entirely typical of sensory-perceptual procedures (**Reitan, 1984** (**http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1349**), 1985). The Reitan-Klove Sensory-Perceptual Examination consists of several methods for delivering unilateral and bilateral stimulation in the modalities of touch, hearing, and vision. The tasks are so simple that normal persons seldom make any errors at all. For example, the examinee is asked to say which hand has been touched (with eyes closed), or to report which ear has received a barely audible finger snap, or to identify which number has been traced on the fingertip. The results of this test are especially diagnostic if the examinee consistently makes more errors on one side of the body than the other. The reader will recall from the previous chapter that neural innervation is almost exclusively opposite-sided. Furthermore, certain areas of the cerebral cortex are devoted to primary processing of touch, hearing, and vision. Thus, an examinee who finds it difficult to process touch in the right hand may have a lesion in the postcentral gyrus of the left parietal lobe. Similarly, difficulty processing sound in the right ear may indicate a lesion in the superior portion of the left temporal lobe, and right-sided visual defects may indicate brain impairment in the left occipital lobe.

Finger Localization Test

Finger localization is a venerable procedure developed by neurologists to evaluate possible sensory losses caused by impairment of brain functions. Most neuropsychological test batteries employ a variant of this test, in which examinees must identify those fingers that have been touched (without benefit of sight). Benton has developed a well-normed 60-item test of finger localization that consists of three parts: (1) with the hand visible, identifying single fingers touched by the examiner with the pointed end of a pencil (10 trials each hand); (2) with the hand hidden from view, identifying single fingers touched by the examiner (10 trials each hand); (3) with the hand hidden from view, identifying pairs of fingers simultaneously touched by the examiner (10 trials each hand). The method of response is left to the patient: naming, touching, or pointing to fingers on a diagram (**Benton, Sivan, Hamsher, Varney, & Spreen, 1994** (**http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib143**)). Each stimulus presentation is scored right or wrong, and normal adults typically make very few errors in the 60 trials. Mean scores for normal adults are near perfect, ranging from 56 to 60 in various samples. In contrast, patients with brain disease find finger localization to be a challenging task, particularly on the second and third parts of the test.

10.15 MEASURES OF ATTENTION AND CONCENTRATION

The attentional capacity of the brain makes it possible to attend to meaningful stimuli, screen irrelevant sensory input from the profusion of incoming stimuli, and flexibly shift to alternative stimuli when conditions demand it (**Kinsbourne, 1994** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib894)**). While in theory it might be possible to make subtle distinctions between simple attention, concentration, mental shifting, mental tracking, vigilance, and other variants of attention/concentration, in practice these skills are difficult to separate. Only one attentional measure—the Test of Everyday Attention (TEA)—has succeeded in partitioning attention into its component sources. We discuss the TEA and other prominent measures of attentional impairment in the following sections.

Test of Everyday Attention

The Test of Everyday Attention (TEA) is a promising measure devised in Great Britain by Robertson, Ward, Ridgeway, and NimmoSmith (**1994** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1374)** , **1996** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1375)**). The TEA measures the subcomponents of attention, including sustained attention, selective attention, divided attention, and attentional switching. The subtests of the TEA are outlined in **Table 10.6** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/ch10lev1sec15#ch10tbl6)** . The test has three parallel versions and has been well validated with closed head injury clients, stroke patients, and persons with Alzheimer's disease. Normative data are based upon the performance of 154 healthy individuals between the ages of 18 and 80. Examinees enjoy the real-life scenarios of the TEA, which adds to the ecological validity of the instrument. The TEA is highly sensitive to normal age effects in the general population and is, therefore, well suited to geriatric assessment. With the exception of the Elevator Counting subtest, the eight subtests were standardized to yield equivalent scores with a common mean of 10 and standard deviation of 3. Thus, the TEA allows for subtest analysis as a means of identifying an individual's particular strengths and weaknesses (**Crawford, Sommerville, & Robertson, 1997** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib369)**). The TEA is highly sensitive to the effects of closed head injury (**Chan, 2000** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib294)**), with the Map Search and Telephone Search subtests revealing the largest deficits from brain injury (**Bate, Mathias, & Crawford, 2001** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib106)**). Chan and his colleagues developed a Cantonese version of the TEA and report favorably on its use with clinical and nonclinical Chinese participants (**Chan, Lai, & Robertson, 2006** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib295)** ; **Chan & Lai, 2006** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib295)**). A children's version (TEA-ch) is also available (**Manly, Nimmo-Smith, Watson, and others, 2001** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1038)**).

Continuous Performance Test

The Continuous Performance Test (CPT) is not really a single test but rather a family of similar procedures that dates back to the pathbreaking research of Rosvold, Mirsky, Sarason, and others (**1956** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1408)**). These authors devised a measure of sustained attention (also called vigilance) that involved continuous presentation of letters on a screen. In some cases, examinees were to press a key when a certain letter appeared (e.g., *x*). In some variations, examinees were to press a key when a certain letter appeared *after* another letter

(e.g., *x* when it occurs after *a*). Errors of omission are noted when the examinee fails to press for a target stimulus. Errors of commission are noted when the examinee presses the key for a nontarget stimulus. Normal subjects make few errors.

TABLE 10.6 Subtests of the Test of Everyday Attention (TEA)

Map Search: A two-minute speeded search for 80 symbols on a colored map; measures selective attention.
Elevator Counting: Simulation of elevator floor counting from tape-presented tones; measures sustained attention.
Elevator Counting with Distraction: Same as above but with auditory distractors; measures sustained attention.
Visual Elevator: Visual simulation of elevator floor counting with up-down reversals; measures attentional switching.
Auditory Elevator with Reversal: Same as visual elevator, except it is presented on tape; measures attentional switching.
Telephone Search: Search for key symbols while searching entries in a simulated classified telephone directory; measures divided attention.
Telephone Search Dual Task: Combines Telephone Search with simultaneous counting of auditory tones; measures divided attention.
Lottery: Subject listens for winning numbers known to end in 55 and then writes down preceding stimuli; measures sustained attention.

Although CPT tests are sensitive to a wide variety of brain-impairing conditions including hyper-activity, drug effects, schizophrenia, and overt brain damage, these tests are not a panacea for the diagnosis of attention-deficit disorders. For example, in one study of the popular Conners (**1995 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib340)**) CPT, children with diagnosed Attention-Deficit/Hyperactivity Disorder (ADHD) did not score worse than clinical controls; on the other hand, children with diagnosed reading disorders showed impaired performance on the CPT (**McGee, Clark, & Symons, 2000 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1091)**). In general, reviewers recommend that CPT tests should be interpreted in the context of a comprehensive test battery, especially when they are used in the assessment of persons with suspected attentional problems (**Riccio, Reynolds, & Lowe, 2001 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1364)**).

The CPT is ideal for computerized adaptation, and dozens of different versions of it have appeared in the literature (e.g., **Conners, 1995 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib340)** ; **Gordon & Mettelman, 1988 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib622)**). Unfortunately, the proliferation of similar but not identical tests has hindered research on the practical utility of this promising measure of attention. Sandford and Turner (**1997 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1434)**) have published a computerized CPT that uses both visual and auditory stimuli. The Intermediate Visual and Auditory Continuous Performance Test (IVA) is normed on 781 normal persons ranging from 5 to 90 years of age

analysis, the IVA showed 92 percent sensitivity (i.e., an 8 percent rate of false negatives) and 90 percent specificity (i.e., a 10 percent rate of false positives) in differentiating children diagnosed with Attention-Deficit/Hyperactivity Disorder (ADHD) from normal children. Research by Tinius (**2003 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1652)**) further endorses the validity of the IVA. He found that adults with mild traumatic brain injury or ADHD performed significantly lower than normal controls on IVA subtests assessing reaction time, inattention, impulsivity, and variability of reaction time. This instrument is just one of many promising neuropsychological tests that takes advantage of microcomputer technology.

10.16 TESTS OF LEARNING AND MEMORY

Learning and memory are intertwined processes that are difficult to discuss in isolation. Learning new material usually requires the exercise of memory. Furthermore, many tests of memory incorporate a learning curve through repeated administrations. The separation of learning and memory processes is theoretically possible but of little practical value in clinical assessment. We make no tight distinction between these processes.

Memory tests can be categorized according to several dimensions, including short term versus long term, verbal versus pictorial, and learning curve versus no learning curve. These dimensions reflect neurological factors discussed in the previous section. For example, verbal memory is significantly lateralized to the left hemisphere, whereas pictorial memory is largely underwritten by the right hemisphere. The interested reader can consult Lezak et al. (**2012 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib983)**) for more detailed analyses of the neural substrates for different types of memory. Here we will concentrate on the psychometric characteristics of four quite dissimilar memory tests.

Wechsler Memory Scale-IV

The Wechsler Memory Scale-IV (Wechsler, 2009) is a monumental revision of the previous edition. The latest version is barely recognizable as the offspring of the original one-page test published more than 60 years ago (Wechsler, 1945). The fourth edition is an extensive, multiphasic memory test consisting of nine subtests, although seven are sufficient for the Standard Battery. The nine subtests are described in **Table 10.7 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/ch10lev1sec16#ch10tbl7)** . The first seven subtests constitute the basis for obtaining age-adjusted scaled scores (mean of 100 and SD of 15) for five standard indices:

- Immediate Memory Index
- Delayed Memory Index
- Auditory Memory Index
- Visual Memory Index
- Visual Working Memory Index

If the ancillary subtests (Logos and Names) are employed, five additional index scores can be computed. We confine our discussion here to the Standard Battery, although it is worth noting that the WMS-IV provides for five flexible batteries (e.g., Older Adult/Abbreviated Battery, Logical Memory/Designs Battery) using different combinations of the nine subtests. The standard battery requires about 75 minutes to administer, while the abbreviated battery can be completed in 35-40 minutes.

TABLE 10.7 WMS-IV Subtests

Immediate Recall Subtests

Immediate Recall Subtests

Brief Cognitive Status Exam: Brief assessment for significant cognitive impairment.

Logical Memory I: Verbal recall of essential elements from brief stories read to the examinee.

Verbal Paired Associates I: Verbal recall for a list of 10 to 14 paired terms (e.g., bicycle—arrow) when only the first term is presented (e.g., bicycle—?).

Designs I: Visual recall for specific elements in a 4 × 4 puzzle grid exposed for 10 seconds; examinee must select small cards with the proper designs and place them correctly within a blank 4 × 4 grid.

Visual Reproduction I: Visual reproduction (drawing) of five (easy to hard) simple geometric designs each exposed for 10 seconds.

Spatial Addition: Visual spatial recall for locations of dots on two separate 4 × 4 grids, adding or subtracting the locations.

Symbol Span: Visual recall for symbols viewed briefly by selecting correct options in the proper order from a large array of symbols.

Logos I: Visual recognition for unique logos paired with fictitious company names by selecting the correct logo from an array when only the company name is provided.

Names I: Recall of names and relevant information about a person from facial images by selecting named persons from a group picture.

Delayed Recall Subtests*

(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/ch10lev1sec16#ch10fn6)

Logical Memory II
Verbal Paired Associates II
Designs II
Visual Reproduction II
Logos II
Names II

*30-minute delayed recall for stimuli in administration I.

The WMS-IV was co-normed with the WAISIV in 2009. The standardization of the new instrument is superb. Based on 2005 census data, the 2,200 participants were stratified as to age (13 age bands spanning 16 to 90), gender, race/ethnicity, educational level, and geographic region.

Because the WMS-IV is a relatively new version, there is currently little external research on its reliability and validity. Even so, the *WMS-IV Technical and Interpretive Manual* (Pearson, 2009) provides a mountain of supportive data. Subtests internal consistencies range from a low of .74 (Visual Reproduction I) to highs of .94 to .97 (Verbal Paired Associates I and Visual Reproduction II, respectively). Internal consistencies of the index scores were excellent, consistently in the mid-to high-90s. Test-retest reliabilities for the index scores were lower, in the low .80s.

Validity of the battery appears strong, based on a variety of approaches, including confirmatory factor

general, the index scores reveal good convergent validity (high correlations with similar measures) and good discriminant validity (low correlations with dissimilar measures). Test profiles for special groups (e.g., intellectual disability, traumatic brain injury, Alzheimer's disease, and schizophrenia) likewise make theoretical sense in light of the aims of the test battery.

An important disclaimer with any multiphasic battery like the WMS-IV is that distinctive profiles should not be used in isolation for diagnosis. If A implies B, it does not follow that B implies A. This is a logical fallacy. A specific example will illustrate the point. If Alzheimer's disease, on average, yields a distinctive WMS-IV profile, it does not follow that the presence of that profile in a new patient signifies that the patient has Alzheimer's disease. Proper diagnosis always entails the synthesis of many sources, including interview with patient and informants.

Likewise, isolated low scores on a WMS-IV index should not be overinterpreted. Accessing the original standardization data, Brooks, Holdnack, and Iverson (**2011** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib223)**) found that healthy people often obtain low scores on one or more index scores, especially when they had lower education levels or intelligence. Moderating influences need to be considered in test interpretation.

Rey Auditory Verbal Learning Test

In the early 1900s, the Swiss psychologist Edouard Claparede (1873-1940) proposed a memory test consisting of the free-recall of a 15-item word list. This test evolved into the Rey Auditory Verbal Learning Test (RAVLT), making it one of the oldest mental tests in continuous use (**Boake, 2002** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib170)**). The test first appeared in French (**Rey, 1964** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1357)**), but an English-language adaptation has been provided by Lezak (**1983** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib980)** , **1995** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib981)**) and others. The RAVLT is a very popular test of memory, especially for purposes of clinical research. A search of PsychINFO from 1950 onward revealed more than 400 published articles using this simple instrument.

In administering the RAVLT, the examiner reads a list of 15 concrete nouns at the rate of one per second. The examinee recalls as many as possible in any order. Forewarning the examinee to recall all the words, including those previously recalled, the examiner reads the entire list a second time. A third, fourth, and fifth administration and recall then ensue; these are followed by an interference trial with a new list of words. Next, immediate recall of the original list is tested (without benefit of a new presentation). Finally, a recognition trial is included in which the examinee must underline the administered words from a longer written paragraph. The test yields a number of scores, including the number recalled (of 15) for each of the initial five trials, the total for the five trials (75 possible), the immediate recall after the distractor list is read, and the recognition score.

Rosenberg, Ryan, and Prifitera (**1984** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1405)**) concluded that the RAVLT performs well in the identification of patients known to be memory impaired by other criteria. In addition to an overall reduction in performance, memory-impaired patients showed a reduced rate of improvement across the five learning trials. Abundant norms for the RAVLT can be found in Strauss, Sherman, and Spreen (**2006** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1585)**). Schoenberg, Dawson-Duff, and others (**2006**

(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1472)) provide normative data for 392 individuals with documented neurological dysfunction.

The RAVLT is available in at least seven parallel versions, which is both a strength and a weakness of the test (**Hawkins, Dean, & Pearlson, 2004** (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib721)). It is a strength because clinicians often employ repeat testing as they follow patients with memory difficulties. Of course, this raises the specter of practice effects: examinees will do better on second, third, and ensuing administrations to some degree because of their prior exposure to the specific items, regardless of whether their clinical condition is improving or getting worse. With parallel versions of a test, the impact of practice effects can be mitigated by using a different form for each administration. Yet, this is a potential weakness, too, because the equivalence of the seven parallel forms is not well established. In reviewing studies of the seven forms of the RAVLT, Hawkins, Dean, and Pearlson (**2004** (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib721)) could locate only six studies, and four of these were limited to comparisons of the original test against one other form. Although differences between forms likely are minor, their exact magnitude is simply unknown.

Fuld Object-Memory Evaluation

The Fuld Object-Memory Evaluation is a useful test of memory impairment in the elderly (**Fuld, 1977** (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib550)). The test begins by presenting the examinee with a bag containing 10 common objects (ball, bottle, button, etc.). The task is not described as a memory test. The examinee is asked to determine whether he or she can identify objects by touch alone. Each object is felt and then named; the examinee then pulls it out of the bag to see if he or she was right. After all 10 items have been correctly identified, a distractor task is administered: rapidly naming words in a semantic category (e.g., names, foods, things that make people happy, vegetables, or things that make people sad). Then the examinee is asked to recall as many of the objects as possible. After each recall, the subject is slowly and clearly reminded verbally of each item omitted on that trial, a procedure called selective reminding (**Buschke & Fuld, 1974** (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib235)). The examinee is then administered four more chances to recall the list by selective reminding, with a distractor task after each trial. Delayed recall is tested after a 5-minute interval. Finally, the test closes with a multiple-choice recognition test.

The Fuld test is often used to help confirm a diagnosis of **Alzheimer's disease** (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm01#bm01gloss09) , a degenerative neurological disorder described in the previous topic. In the early stages of Alzheimer's disease the most prominent symptom is memory loss. Elderly persons with memory impairment not only score lower than control subjects on the Fuld Object-Memory Evaluation, but they also benefit very little from the selective reminding. Fuld (**1977** (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib550)) has provided norms for community-active and healthy nursing-home residents in their 70s and 80s. Fuld, Masur, Blau, Crystal, and Aronson (**1990** (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib551)) describe a prospective study in which the Fuld Object-Memory Evaluation demonstrated promise as a predictor of dementia in cognitively normal elderly. Lichtenberg, Manning, Vangel, and Ross (**1995** (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib985)) describe a program of neuropsychological research using the Fuld test with older urban medical patients.

Chung (**2009 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib305)**) reports very favorably on the validity of the Fuld test as a screening measure of dementia in Chinese elderly. In a sample of 192 community-dwelling individuals, 57 with confirmed dementia, the optimal cut-off on the total retrieval score yielded an amazing 93 percent sensitivity and 90 percent specificity. In other words, 93 percent of the individuals with dementia were correctly spotted, and 90 percent of the normal individuals were appropriately classified. These are impressive findings for a simple screening test. Chung and Ho (**2009 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib305)**) report similarly favorable results in a Chinese nursing-home sample.

Rivermead Behavioral Memory Test

The Rivermead Behavioral Memory Test (RBMT) is a measure of everyday memory such as route finding, remembering names, and recalling information (**Wilson, Cockburn, & Baddeley, 1991 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1763)**). The instrument includes the following subtests:

- *Names:* A photograph is shown along with the first and second names of the person in the photograph. The examinee is tested on both the first and the second names.
- *Belonging:* At the beginning of the test, the examinee is required to hand over a personal belonging (e.g., wallet), which is then hidden while the examinee observes. Later the examinee must remember to ask for the item and then also to find it.
- *Appointment:* The examinee is asked to remember to ask the date of the next appointment when he or she hears the sound of an alarm timer.
- *Pictures:* The examinee is shown 10 cards with simple pictures or drawings and later is asked to recognize them among a set of 20 cards.
- *Immediate Story:* The examiner reads a short paragraph and immediately afterward asks the examinee to recall as many elements of the brief story as possible.
- *Delayed Story:* After completing a number of additional subtests, the examinee is asked to recall as many elements of the story as possible.
- *Faces:* The examinee is shown 5 cards with a face on them and then asked to recognize them among a set of 10 cards.
- *Immediate Route:* The examiner demonstrates a short route with the examinee and leaves an envelope with a written message at the destination. The examinee is asked to reproduce the route and to recall the message.
- *Immediate Message:* This item is linked to Immediate Route (above). The examinee is asked to recall the written message.
- *Delayed Message:* After completing a number of intervening tasks, the examinee is asked to recall the written message again.
- *Orientation:* This subtest consists of 10 items tapping knowledge of personal and societal information.
- *Date:* The examinee is asked the date of the examination.

The RBMT is highly popular in geriatric and rehabilitation settings because of its robust ecological validity —the subtests parallel the tasks and activities of everyday life (**Guaiana, Tyson, & Mortimer, 2004 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib662)**). Another strong point of the instrument is that it assesses many elements of memory. For example, the test evaluates all of the following aspects: short-term, long-term, verbal, spatial, retrospective, and prospective memory. The focus on prospective memory—remembering to do something in the future—is a rare but welcome

Man, Chung, and Mak (**2009**
(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1037)) developed an online version of the RBMT for use with Chinese examinees. They compared scores of 30 stroke patients on the original, face-to-face version of the test versus the online version, and found exceptionally strong correlations on the 12 subtests, with *r*s ranging from .84 to .93. The new version also was highly successful in distinguishing stroke patients from controls. In sum, the online adaptation looks highly promising as a replacement for the more cumbersome face-to-face edition.

Wide Range Assessment of Memory and Learning-2

The original version of the Wide Range Assessment of Memory and Learning (WRAML) was the first comprehensive memory scale designed for use with children (ages 5 to 17 years). The second edition of the test, the WRAML-2 (Sheslow & Adams, 2004), retains the pediatric focus but also extends the norms upward to 90 years of age. The WRAML-2 is, therefore, unique as the only memory scale that can be used with both children and adults. In addition to examiner convenience (no need to buy and learn several memory tests), there is clinical value as well in using a single test across a wide range of ages. Specifically, when clinicians desire to do follow-up testing on a child or teenage client who subsequently transitions into adulthood, using a single test avoids the pitfall of introducing measurement error associated with different tests.

The WRAML-2 consists of six core subtests that contribute to three Index scores: Verbal Memory, Visual Memory, and Attention/Concentration. Collectively, these Index scores establish the overall General Memory Index. A description of the core memory tasks is provided in **Table 10.8**
(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/ch10lev1sec16#ch10tbl8) .

In addition to the core memory subtests, the WRAML-2 also utilizes delayed memory tasks and recognition memory tasks. The delayed memory tasks require free recall of previously presented material whereas the recognition memory tasks involve mere recognition of the material. The two formats (delayed and recognition) help distinguish between storage and retrieval problems in memory. In particular, a client who performs poorly on delayed memory but who excels at recognition memory most likely has a problem with retrieval rather than storage. This is somewhat similar to not remembering a test item when a fill-in-the blank format is used but succeeding when a multiple-choice format is used. In fact, retrieval memory requires a different neurological substrate than recognition memory. Although capable functioning in both retrieval and recognition memory is typical throughout life, distinct differences (favoring recognition) are observed in old age, with certain neurological conditions such as Alzheimer's disease, and in some forms of brain injury.

TABLE 10.8 Description of core WRAML-2 Subtests

Verbal Memory Subtests
Story Memory: Two short stories are read to the participant who, following each, is asked to recall as many parts of the story as can be remembered. This task measures immediate verbal memory.
Verbal Learning: The examinee is read a relatively long list of simple words followed by an immediate free-recall trial. Three additional presentation/recall trials are used. This task evaluates the ability to actively learn verbal information and yields a verbal learning curve over the four trials.
Visual Learning Subtests

Design Memory: A card with a simple geometric array is presented for a 5-second exposure. Following a 10-second delay, the participant is asked to draw what is remembered about the card. This procedure is used for five separate cards of increasing difficulty.

Picture Memory: The examinee visually scans a complex but common meaningful scene for 10 seconds. Then the examinee is presented with a second similar scene and asked to indicate which elements "have been moved, changed, or added" in the second picture. The procedure is used with four separate scenes.

Attention/Concentration Subtests

Finger Windows: The participant demonstrates memory of a visual pattern using a vertically resting card containing asymmetrically located holes or "windows." The examiner points out a sequence of windows, and then the participant is asked to duplicate the sequence.

Number Letter: The examinee is asked to verbally repeat a random series of numbers and letters orally presented at one per second.

Note: All subtests listed contribute to the General Memory Index.

The WRAML-2 also includes optional subtests that can be used to evaluate a relatively new area of memory measurement, namely, working memory (**Baddeley, 1986 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib73)**). Working memory is a complex form of short-term memory. In addition to simply holding on to rote information for several seconds, when using working memory the client is also "working" with a part of the memory trace without distorting the whole trace. For example, try to read the following sentence only once (i.e., do not reread the sentence to answer the question): If in a bag you had two red balls, three yellow balls, and one green ball, what is the probability the ball would be yellow if you reached into the bag and randomly chose one ball? To answer this question, the short-term verbal memory processor must hold on to all the words in the sentence until the last phrase containing the question. Then it must reproduce the sentence, remembering how many red balls there were, and so on, then hold that information secure, returning to accumulate all the numbers in order to compute the answer. There are two working memory subtests on the WRAML-2, one that examines verbal working memory and another that examines a combination of verbal and visual working memory.

The adult standardization age bands used in norming the WRAML-2 are similar to those of the WMS-III, with similar attention given to stratification variables such as age, gender, ethnicity, geographic region, and educational level. "Tighter" age bands exist for the 5- to 14-year-old samples because there is more change in memory abilities across these ages than in adulthood (except for the oldest age groups). For the WRAML-2, factor-analytic studies show strong support for the three discrete domains being measured (Verbal Memory, Visual Memory, and Attention/Concentration) as well as the newly introduced domain of Working Memory. Especially impressive are the analyses showing extremely low item bias for gender as well as ethnicity. As with the WMS-III, validity studies show clinical groups with neurological disorders scoring significantly lower than nonclinical groups on all WRAML-2 Indexes. The correlation of the WRAML-2 with WAIS-III Full Scale IQ is moderate, supporting the claim that it measures something different from, although related to, intelligence. Of interest, though, a much lower correlation with the WISC-III suggests that there is less correlation between intelligence and memory ability among children than among adults.

Because both tests claim to be memory tests and show some similarities across tasks used to assess memory, it is reasonable to wonder if the WMS-III and WRAML-2 yield similar scores (i.e., if there is reasonable concurrent validity). Using 79 adults from ages 17 through 74 years, the test developers

correlations between scores on the two memory instruments ranged from .29 to .60. These moderate correlations suggest that they are measuring somewhat different aspects of memory and are not interchangeable instruments.

Additional Tests of Learning and Memory

Because of space limitations, we can do no more than briefly mention several other useful tests of learning and memory. The California Verbal Learning Test-II is patterned after the Rey AVLT but provides software to quantify and analyze the pattern of results (**Delis, Kramer, Kaplan, & Ober, 2000** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib408)**). The Benton Visual Retention Test is a design-copying test of visual memory (**Sivan, 1991** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1504)**). Good reviews of memory tests can be found in Lezak et al. (**2012** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib983)**) and Strauss, Sherman, and Spreen (**2006** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1585)**).

10.17 ASSESSMENT OF LANGUAGE FUNCTIONS

As noted in a previous section, language functioning offers a window to the integrity of the left cerebral hemisphere. Thus, neuropsychologists are keenly interested in an examinee's ability to speak, read, write, and comprehend what others say. Little wonder that a comprehensive neuropsychological examination always includes one or more methods for assessing language functions.

Neuropsychologists exhibit a special interest in a variety of language dysfunctions known collectively as aphasia. Briefly stated, **aphasia (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm01#bm01gloss13)** is any deviation in language performance caused by brain damage. In testing for aphasia, a neuropsychologist might use any or all of three approaches: (1) a nonstandardized clinical examination, (2) a standardized screening test, or (3) a comprehensive diagnostic test of aphasia. We will provide examples of each in our brief review of assessment methods in aphasia.

Clinical Examination for Aphasia

A clinical examination for aphasia has the advantages of simplicity, flexibility, and brevity. These are important attributes when assessing a severely impaired patient who may require bedside testing. Every practitioner has a slightly different version of the brief clinical exam (**Lezak et al., 2012 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib983)** ; **Reitan, 1984 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1349)** , 1985). Nonetheless, certain elements commonly are assessed:

- *Spontaneous speech:* The examiner looks for distinctive symptoms of aphasia such as word-finding difficulty or neologisms (e.g., referring to a comb as a "planker").
- *Repetition of sentences and phrases:* The examiner asks the patient to repeat stimuli such as "No ifs, ands, or buts," and "Methodist Episcopal." The repetition tasks are so simple that normal subjects almost never fail them.
- *Comprehension of spoken language:* The examiner asks questions ("Does a car have handlebars?") and issues commands ("Take this paper, fold it in half, and put it on the floor"). Again, the tasks are so simple that normal subjects almost never fail them.
- *Word finding:* The examiner points to common, easily recognized objects and asks, "What's this?" Typical items include watch, pen, pencil, glasses, ring, and shoes. The examiner may ask the patient to name numbers, letters, or colors.
- *Reading:* The examiner requests the patient to read and explain a short paragraph suited to prior level of education and intelligence. The examiner may ask the patient to follow written instructions (e.g., "Close your eyes" or "Clap your hands three times").
- *Writing and copying:* The examiner asks the patient to write spontaneously and from dictation. Also, the examiner may ask the patient to copy written matter and geometric shapes. The examiner is interested in grossly ungrammatical written productions and significant distortions in copying.
- *Calculation:* The examiner asks the patient to perform very simple mathematical calculations (e.g., 17 × 3) with and without aid of scratch paper. The tasks are so simple that normal subjects rarely fail.

Based on the clinical assessment, the examiner may fill out a rating scale for severity of aphasia. For example, the rating scale used in the Boston Diagnostic Aphasia Exam (**Goodglass, Kaplan, & Barresi, http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib618)**) includes the

following speech characteristics: melodic line, phrase length, articulatory agility, grammatical form, word finding, and auditory comprehension.

Screening and Comprehensive Diagnostic Tests for Aphasia

Standardized screening tests for aphasia closely resemble the brief clinical exam. The essential difference is that standardized screening tests incorporate objective and precise instructions for administration and scoring. The weakness of screening tests is that they will not detect subtle forms of aphasia.

Comprehensive diagnostic tests for aphasia are quite lengthy and used mainly when a patient is known to experience aphasia. These tests provide a profile of language skills that is helpful in treatment planning. We provide a brief description of several aphasia tests in **Table 10.9 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/ch10lev1sec18#ch10tbl9)** .

10.18 TESTS OF SPATIAL AND MANIPULATORY ABILITY

Tests of spatial and manipulatory ability are also known as tests of constructional performance. A constructional performance test combines perceptual activity with motor response and always has a spatial component (**Lezak et al., 2012** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib983)**). Because constructional ability involves several complex functions, even mild forms of brain dysfunction will result in impaired constructional performance. However, careful observation is needed to distinguish the cause of the failed performance, which may include spatial confusion, perceptual deficiency, attentional difficulties, motivational problems, and apraxias. The term apraxia refers to a variety of dysfunctions characterized by a breakdown in the direction or execution of complex motor acts (**Strub & Black, 2000** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1596)**). For example, a patient who could not demonstrate how to use a key would be diagnosed as suffering from ideomotor apraxia.

TABLE 10.9 Brief Description of Several Aphasia Tests

Multilingual Aphasia Examination (Benton, Hamsher, Rey, & Sivan, 1994
(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib143))
This respected, comprehensive battery consists of 11 subtests and rating scales that assess visual naming, repetition, fluency, articulation, spelling, and other language variables; available in a Spanish edition, too.

Western Aphasia Battery—Revised (Kertesz, 2000)
Comprehensive test of verbal fluency, auditory comprehension, and repetition that aims to identify aphasia syndromes and determine their severity.

Boston Diagnostic Aphasia Examination (Goodglass, Kaplan, & Barresi, 2000
(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib618))
Comprehensive test with 46 subscales that include music, spatial, computation, and seven types of writing skill in addition to traditional aphasia measures, available in French and Hindi versions, too.

Porch Index of Communicative Ability—Revised (porch, 2001
(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1316))
A battery containing eighteen 10-item subtests, four verbal, eight gestural, and six graphic. Very reliable test often used to measure small changes in patient performance.

Token Test (Spreen & Strauss, 1998
(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1541))
An extremely sensitive test that presents little challenge to normal individuals. The examinee must complete oral commands with colored tokens, e.g., "Put the small red token on top of the large square token." Originally devised by Boller & Vignolo (1966), numerous versions of the Token Test are now available.

Tests of constructional performance embrace two large classes of activities: drawing and assembling. Owing to limitations of space, we will review only a few prominent instruments in each category.

Design Copying Tests

Drawing a copy of simple geometric shapes such as two overlapping pentagons is a complex activity that requires accurate visual perception, correct spatial analysis, as well as intact motor functions and the executive ability to make mid-course corrections in the drawing. Because the activity of copying a design involves so many cognitive capacities, it is sensitive to a wide variety of brain impairing conditions. For this reason, design copying has been a mainstay of cognitive screening for brain impairment.

One of the most widely used design copying tests—indeed, one of the most widely used individual tests of any kind—is the Bender Visual-Motor Gestalt Test (**Bender, 1938** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib132)**), more commonly known as the Bender Gestalt Test (BGT). In the last half of the twentieth century, the BGT consistently ranked among the top four or five most frequently used tests in clinical psychology (Piotrowski, 1995). The original version consisted of nine stimulus drawings similar to those in **Figure 10.5** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/ch10lev1sec18#ch10fig5)** . The test is simple to explain and administer. The examinee is instructed to copy one drawing at a time on a sheet of blank paper. Erasures are discouraged. If needed, additional sheets of paper are provided. The examinee is told "this is not a test of artistic ability, but try to copy the drawings as accurately as possible. Work as fast or as slowly as you wish" (**Hutt, 1977** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib804)**). Use of a ruler or straight edge is not permitted.

For the original version of the BGT, a number of complex scoring systems have been developed for adults (**Hain, 1964** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib683)** ; **Hutt & Briskin, 1960** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib805)** ; **Lacks, 1999** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib943)**). In addition, Koppitz (**1963** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib914)** , **1975** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib915)**) produced an elaborate scoring system for children aged 5 to 11. The Koppitz system yielded a raw score (total errors) that could be converted to an age-equivalent score as well. In contrast to the use of the BGT with adults—where the examiner is looking for signs of brain impairment—when used with children, the primary purpose of the test is to assess the level of developmental maturity. Several interesting variations on the original BGT are discussed in Gregory (**1999** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib650)**).

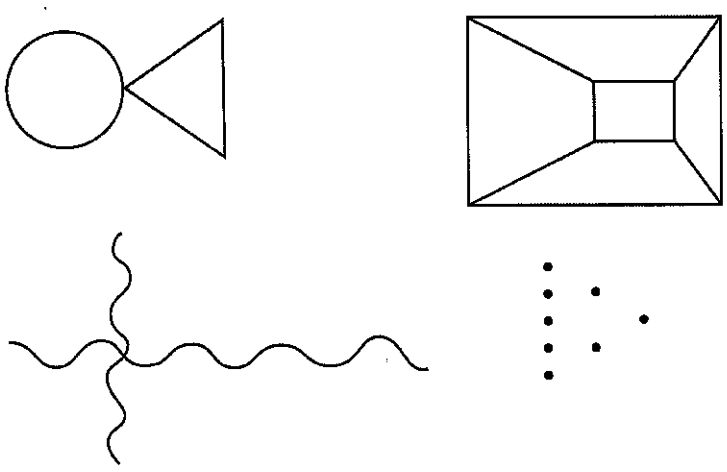


FIGURE 10.5 Stimuli Similar to Those From the Bender Gestalt Test-II.

Note: The Bender-Gestalt-II consists of sixteen stimuli similar to these.

A revised and expanded version of the BGT was published by Brannigan and Decker (**2003** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib209)**). The BGT-II adds to the original test rather than revamping it. Specifically, it includes the original nine stimulus cards supplemented by seven new drawings (four very easy drawings, and three that provide substantial challenge). The four additional "easy" cards are administered only to younger examinees 4 through 7 years of age, whereas the three "difficult" cards are administered only to older examinees 8 through 851 years of age. Unlike previous editions of the test which lacked serious efforts at standardization, the BGT-II norms are based on more than 4,000 individuals, ages 4 through 85, stratified on important demographics according to the 2000 census.

These new stimulus cards are intended to extend the measurement scale at the lower and higher extremes of ability. The authors also provide an explicit scoring system whereby each reproduction is scored on a 5-point scale from 0 (no resemblance) to 4 (nearly perfect). Of course, comprehensive, census-based norms are provided by way of standard scores, *T* scores, percentile ranks, confidence intervals, and classification labels. The standard score is called the Visual Motor Integration (VMI) and is anchored to a mean of 100 and standard deviation (SD) of 15. This is a useful feature of the BG-II because it allows for comparisons of the VMI score with IQs, memory quotients, and other indices normed to mean of 100 and SD of 15. Marnic (**2011** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1045)**) found that the test is valuable in the diagnosis of attention-deficit/hyperactivity disorder in referred children and adolescents. Decker (**2008** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib405)**) provides a sophisticated analysis of subtle changes in BGT-II protocols across the life span, suggesting that visual-motor skills mature rapidly from childhood into middle adolescence, decline steadily through adulthood, and drop steeply in old age.

The Greek Cross (**Reitan & Wolfson, 1993** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1350)**) is a very simple drawing task that is surprisingly sensitive to brain impairment. The examinee is requested to carefully copy the figure without lifting the pencil, that is, by tracing the perimeter. The stimulus figure and examples of defective performance are shown in **Figure 10.6** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/ch10lev1sec19#ch10fig6)** . This test is most often evaluated on a qualitative basis, although scoring guides do exist (**Gregory, 1999** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib650)**).

Assembly Tests

In his classic book on the parietal lobes, Critchley (**1953** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib372)**) provided the rationale for including three-dimensional construction tasks in a neuropsychological test battery:

> It is possible, and indeed useful, to proceed to problems in three-dimensional space though tests of this character are only too rarely employed. This is a more difficult undertaking, and patients who respond moderately well to the usual procedures with sticks and pencil-and-paper may display gross abnormalities when told to assemble bricks according to a three-dimensional pattern.

Benton, Sivan, Hamsher, Varney, and Spreen (**1994 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib143)**) present a three-dimensional block construction test with excellent norms and scoring guide. The two forms of the test (Form A and Form B) consist of three block models that are presented one at a time to the patient. The patient is requested to construct an exact replica of the model by selecting the appropriate blocks from a set of loose blocks on a tray. Based on omissions, additions, substitutions, and displacements, the three models are scored from 0 to 6, 8, and 15 points, respectively. This test is quite sensitive to brain impairment, especially when the left or right parietal area is affected. Lezak et al. (**2012 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib983)**) discusses other assembly tasks. We should mention that the Tactual Performance Test from the Halstead-Reitan battery is, in part, an assembly task that measures spatial and manipulatory abilities (see **Table 10.4 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/ch10lev1sec12#ch10tbl4)**).

10.19 ASSESSMENT OF EXECUTIVE FUNCTIONS

Executive functions **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm01#bm01gloss102)** include logical analysis, conceptualization, reasoning, planning, and flexibility of thinking. The assessment of executive functions presents an unusual quandary to neuropsychologists:

> A major obstacle to examining the executive functions is the paradoxical need to structure a situation in which patients can show whether and how well they can make structure for themselves. Typically in formal examinations, the examiner determines what activity the subject is to do with what materials, when, where, and how. Most cognitive tests, for example, allow the subject little room for discretionary behavior, including many tests thought to be sensitive to executive—or frontal lobe—disorders . . . The problem for clinicians who want to examine the executive functions becomes how to transfer goal setting, structuring, and decision making from the clinician to the subject within the structured examination. (**Lezak, 1995** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib981)**)

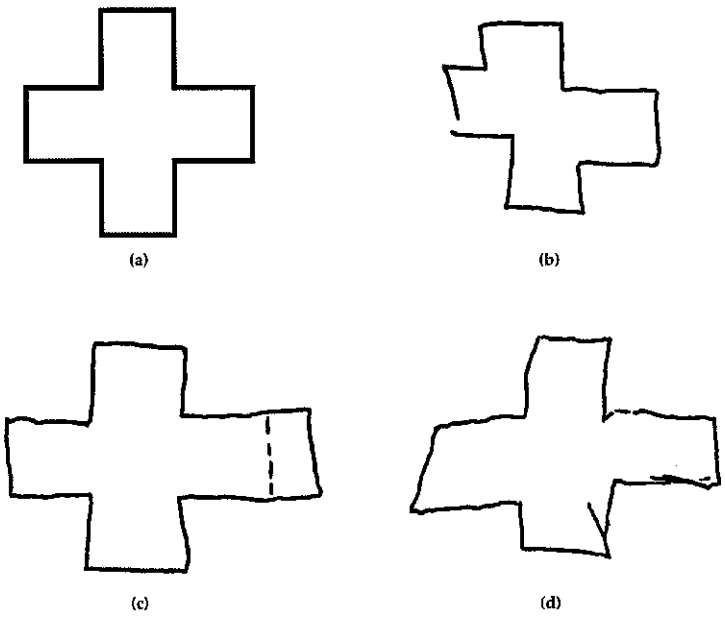


FIGURE 10.6 The Greek Cross Stimulus Figure and Reproductions from Persons with Known Brain Damage

(a) Stimulus figure.
(b) Clerical worker with diffuse right hemisphere dysfunction of unknown origin.
(c) College professor two years after a right hemisphere stroke.
(d) Patient with generalized, diffuse dementia.

Source: From Gregory, Robert J. *Foundations of intellectual assessment: The WAIS-III and other tests in clinical practice*, p. 197. Published by Allyn and Bacon, Boston, MA. Copyright © 1999 by Pearson Education. Adapted by permission of the publisher.

Many neuropsychologists resolve this quandary by using the clinical method to evaluate executive functions rather than administering formal tests (**Cripe, 1996** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib371)**). For example, Pollens, McBratnie, and Burton (**1988** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1312)**) use interview and observations to fill out the structured checklist on executive functions mentioned in the previous topic.

Only a limited number of neuropsychological tests tap executive functions to any appreciable degree. Useful instruments in this regard include the Porteus Mazes, Wisconsin Card Sorting Test, and a novel approach known as the Tinkertoy® Test. We remind the reader that the Category Test from the Halstead-Reitan battery also captures executive functions to some extent (**Table 10.4** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/ch10lev1sec12#ch10tbl4)**).

The Porteus Maze Test was devised as a culture-reduced measure of planning and foresight (**Porteus, 1965** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1318)**). Without lifting the pencil and attempting to avoid dead ends, the examinee must trace a line through a series of increasingly difficult mazes. This underused instrument is quite sensitive to the effects of brain damage, particularly in the frontal lobes (**Smith & Kinder, 1959** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1509)** ; **Smith, 1960** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1507)**).

Krikorian and Bartok (**1998** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib919)**) have published contemporary Porteus Maze norms for children and young adults 7 to 21 years of age; these researchers also demonstrated that test scores are minimally related to IQ scores. Mack and Patterson (**1995** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1027)**) investigated the Porteus test as a useful measure of executive function in elderly patients with Alzheimer's disease. In a study of 276 pediatric patients who had sustained a traumatic brain injury (TBI), Levin, Song, Ewing-Cobbs, and Roberson (**2001** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib972)**) found that the Porteus test was highly sensitive to TBI severity as measured by the volume of tissue damage in the prefrontal areas of the brain.

The Wisconsin Card Sorting Test (WCST) is a good measure of executive functions, although its differential sensitivity to frontal lobe damage is debated (**Mountain & Snow, 1993** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1185)**). The instrument was devised to study abstract thinking and the ability to shift set (**Berg, 1948** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib146)** ; **Heaton, Chelune, Talley, and others, 1993** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib727)**). The examinee is given a pack of 64 cards on which are printed one to four symbols (triangle, star, cross, or circle) in one of four colors (red, green, yellow, or blue). No two cards are identical. Thus, each card embodies a number, a particular shape, and a specific color. The examinee must sort these cards underneath four stimulus cards according to an unknown principle (**Figure 10.7** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/ch10lev1sec19#ch10fig7)**). For example, the unknown principle might be "sort according to color." As the examinee places cards, the examiner says "right" or "wrong." After the examinee has sorted a run of 10 correct placements in a row, the examiner shifts the principle without warning. The test continues until the examinee has made six runs of 10 correct placements. The test can be scored in several different ways, including total number of trials to

criterion (**Axelrod, Greve, & Goldman, 1994** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib70)**). A common use of the WCST is to gauge ongoing recovery in patients with brain trauma of recent onset. Thus, the longitudinal constancy of test scores in patients with stabilized conditions is a reassuring characteristic of this test (**Greve, Love, Sherwin, and others, 2002** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib654)**).

Lezak (**1982** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib979)**) devised the Tinkertoy® Test to give patients the opportunity to demonstrate executive capacities within the structured format of an examination. Fifty pieces of a standard Tinkertoy® set are placed on a clean surface and the examinee is told, "Make whatever you want with these. You will have at least five minutes and as much more time as you wish to make something." The test is scored from – 1 to 112 based on several variables including the number of pieces used, the mobility of the construction, symmetry, and the naming of the construction. Head-injured patients produce impoverished designs consisting of a small number of pieces. These individuals often are unable to provide a name for their constructions.

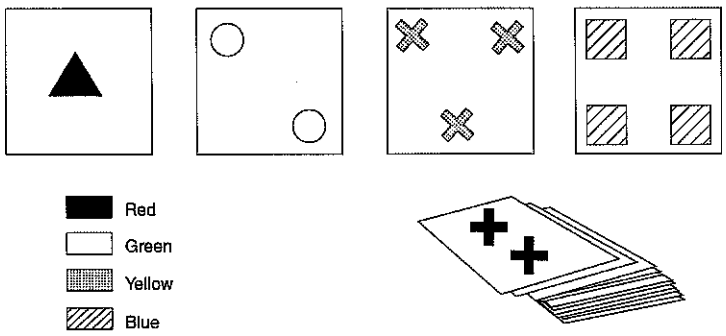


FIGURE 10.7 Cards and Sorting Piles Similar to the Wisconsin Card Sorting Test

Bayless, Varney, and Roberts (**1989** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib111)**) studied the predictive validity of the Tinkertoy® Test by comparing the results of 50 patients with closed-head injuries versus 25 normal controls. Half of the head-injured individuals had returned to work while half had not. Whereas all but one of the head-injured who returned to work scored normally on the Tinkertoy® Test, nearly half of the nonreturnees performed below the level of the worst control subject. The researchers conclude:

> The test seems particularly well suited for demonstrating the presence of deficits in executive functioning, which have proven to be difficult to demonstrate with clinical tests even though they have catastrophic sequelae in daily vocational or psychosocial endeavors. (**Bayless et al., 1989** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib111)**)

The Tinkertoy® Test also shows promise in the assessment of individuals with Alzheimer's disease (**Koss, Patterson, Mack, Smyth, & Whitehouse, 1998** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib916)**).

Neuropsychologists still need additional measures of executive functions. One promising approach in the early stages of development is real-world assessment of route finding. The ability to find an unfamiliar

functions applied to a realistic problem (**Boyd & Sauter, 1993** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib193)**). Another promising approach to the assessment of executive functions is embodied in a recent battery called the Behavioral Assessment of the Dysexecutive System (**BADS; Wilson, Alderman, Burgess, and others, 1996** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1764)**). The BADS battery consists of six novel situational tests that resemble real-life daily activities:

- *Temporal Orientation:* The examinee is asked to estimate how long various common activities take, such as a routine dental checkup.
- *Rule Shift Cards:* This test measures the ability to shift set after establishing a card-sorting pattern according to a simple rule.
- *Action Program:* This test of practical problem solving involves a task in which a cork must be extracted from a test tube by planning the use of available materials.
- *Key Search:* In this analogue test, examinees are required to demonstrate how they would search a field for a set of lost keys.
- *Zoo Map:* This is a test of planning and route finding in which the examinee is asked to plan a route to visit six of a possible 12 locations in a zoo.
- *Six Elements:* This is a multitasking subtest in which the examinee must complete six activities (two naming, two dictation, two mental arithmetic) in 10 minutes.

The battery also includes a 20-item dysexecutive questionnaire with items rated on a 5-point (0 to 4) Likert scale. The items involve likely changes when executive functions are impaired, for example, "I have difficulty thinking ahead and planning for the future." The questions are in four broad areas: personality/emotional changes, motivational changes, behavioral changes, and cognitive changes. Spreen and Strauss (**1998** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1541)**) provide a helpful review of this battery. Norris and Tate (**2000** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1240)**) compared the BADS with six other commonly used tests of executive functioning. In a sample of 36 neurological patients, they demonstrated the ecological superiority of this new instrument in predicting competency in everyday role functioning. Simon, Giacomini, Ferrero, and Mohr (2003) found that the BADS was a fair measure of social adjustment in patients with schizophrenia, correlating $r = .34$ with an index of psychosocial adjustment. The BADS outperformed the Wisconsin Card Sorting Test and the Trail Making Test (part B) in this context. In a study comparing healthy controls, patients with mild cognitive impairment, and patients with mild Alzheimer's disease, the BADS was highly sensitive to the impact of mild Alzheimer's disease, but did not differentiate the other two groups (**da Costa Armentano, Porto, Brucki, & Nitrini, 2009** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib384)**).

10.20 ASSESSMENT OF MOTOR OUTPUT

Most neuropsychological test batteries include measures of manipulative speed and accuracy. Lezak et al. (**2012** (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib983)) provides a comprehensive review. We will briefly summarize three approaches: finger tapping, pegboard performance, and line tracing.

Perhaps the most widely used test of motor dexterity is the Finger-Tapping Test from the Halstead-Reitan battery. This test consists of a tapping key that extends from a mechanical counting device attached to a flat board. With the index finger of each hand, the examinee completes a series of 10-second trials until five trials in a row are within a 5-point range. The score for each hand is the average of these five trials, rounded to the nearest whole number. With the dominant hand, males typically score about 54 taps (SD of 4), whereas females typically score about 51 taps (SD of 5; **Dodrill, 1979** (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib423) ; **Morrison, Gregory, & Paul, 1979** (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1178)).

In general, the absolute level of performance is of less interest than the relative abilities on the two sides of the body. Normative expectation is that the nondominant hand will yield a tapping rate about 90 percent of the dominant hand. Significant deviations from this pattern are thought to indicate a lesion in the hemisphere opposite that of the slowed hand (**Haaland & Delaney, 1981** (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib678)). However, such inferences must be made with great caution owing to the very low reliability of the ratio score. Although test-retest and interexaminer reliabilities for either hand alone approach .80, the reliability of the ratio score is a dismal .44 to .54 (**Morrison, Gregory, & Paul, 1979** (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1178)). The ratio score should be used with extreme caution in making clinical inferences about lateralization of damage.

The Purdue Pegboard Test requires the examinee to place pegs in holes with the left hand, right hand, and then both hands. Each trial lasts only 30 seconds, so the entire test can be administered in a matter of minutes. Tiffin (**1968** (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1651)) reports normative scores for work applicants. Relative slowing in one hand suggests a lesion in the opposite hemisphere, whereas bilateral slowing indicates diffuse or bilateral brain damage. Using the Purdue Pegboard Test in isolation, one study found an 80 percent accuracy in identifying brain impairment among a large group of normal subjects and neurological patients (**Lezak, 1983** (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib980)). Other studies report much less favorable findings (**Heaton, Smith, Lehman, & Vogt, 1978** (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib728)). The Purdue Pegboard Test is a useful addition to a comprehensive battery but should not be used in isolation for screening purposes. Spreen and Strauss (**1998** (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1541)) provide an excellent summary of norms for this widely used test.

Klove has developed a variation on the pegboard test in which the pegs have a ridge along one side (**Klove, 1963** (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib904)). Because each peg must be rotated into position, the Grooved Pegboard requires complex coordination in addition to motor dexterity. The Grooved Pegboard test is an excellent instrument for assessing lateralized brain damage (**Haaland & Delaney, 1981**

Finally, we should mention that useful motor tests need not require sophisticated equipment. Lezak (**1995** (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib981)) recommends a line tracing task to assess difficulties in motor regulation (**Figure 10.8** (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/ch10lev1sec21#ch10fig8)). The examinee is given a brightly colored felt-tipped pen and a sheet of paper with several figures and told to draw over the lines as rapidly as possible. Difficulties with motor regulation show up in overshooting corners, perseveration of an ongoing response, and inability to follow the reduced curves in the bottom figure. Because this task is easily completed by most 10-year-olds, any noticeable deviations are suggestive of difficulties in motor regulation.

10.21 TEST BATTERIES IN NEUROPSYCHOLOGICAL ASSESSMENT

We remind the reader that the Halstead-Reitan Neuropsychological Battery (**Reitan & Wolfson, 1993** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1350)**), discussed earlier, is a respected and widely used battery in neuropsychological assessment. Here we summarize competing approaches.

The Luria-Nebraska Neuropsychological Battery

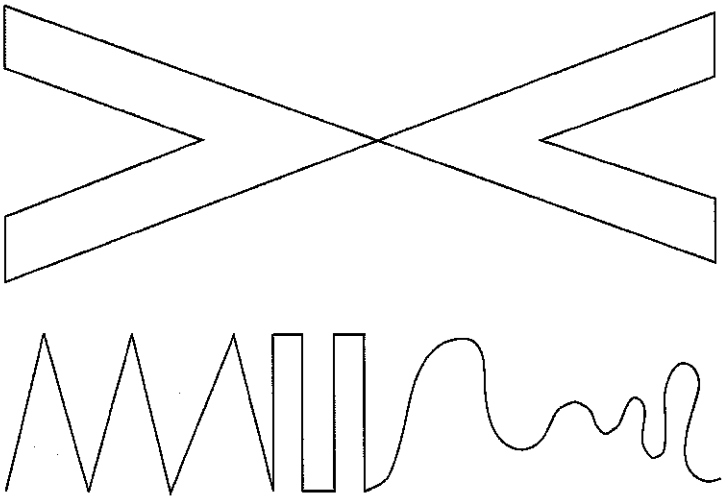


FIGURE 10.8 A Typical Line-Tracing Task (Reduced Size)

Now that we have completed a tour of some individual neuropsychological tests and procedures, it is time once again to remind the reader that many neuropsychologists prefer to use a fixed battery rather than an ever-shifting, individualized assortment of instruments. Certainly, one of the most widely used fixed batteries is the Luria-Nebraska Neuropsychological Battery (LNNB; **Golden, 2004** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib607)** ; **Golden, Purish, & Hammeke, 1980** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib608)** , **1986** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib609)**), now in its third edition (LNNB-III; **Teichner, Golden, Bradley, & Crum, 1999** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1622)**).

The test consists of 269 discrete items, chosen from the work of Luria (**1966** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1017)**) and formally standardized. These items are scored 0, 1, or 2 according to precise criteria in the administration and scoring manual. Similar items are grouped into 11 clinical scales, C1 through C11 (**Table 10.10** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/ch10lev1sec21#ch10tbl10)**). Raw scores on each scale are converted into *T* scores, with a mean of 50 and a standard deviation of 10. Higher scores reflect more psychopathology; scores above 70 are especially suggestive of brain impairment.

TABLE 10.10 Tests and Procedures of the Luria-Nebraska Neuropsychological Battery

Ability Scale: Tasks Included

C1 Motor: Coordination, speed, drawing, complex motor abilities

C2 Rhythm: Attend to, discriminate, and produce verbal and nonverbal rhythmic stimuli

C3 Tactile: Identify tactile stimuli, including stimuli traced on the wrists

C4 Visual: Identify drawings, including overlapping and unfocused objects; solve progressive matrices and other visuospatial skills

C5 Receptive Speech: Discriminate phonemes and comprehend words, phrases, sentences

C6 Expressive Speech: Articulate sounds, words, and sentences fluently; identify pictured or described objects

C7 Writing: Use motor writing abilities in general; copy and write from dictation

C8 Reading: Read letters, words, and sentences; synthesize letters into sounds and words

C9 Arithmetic: Complete simple mathematical computations; comprehend mathematical signs and number structure

C10 Memory: Remember verbal and nonverbal stimuli under both interference and noninterference conditions

C11 Intelligence: Reasoning, concept formation, and complex mathematical problem solving

Three summary scales are also derived from test performance: S1 (Pathognomonic), S2 (Left Hemisphere), and S3 (Right Hemisphere). The Pathognomonic scale reflects the degree of compensation that has occurred since an injury, such as functional reorganization of the brain as well as actual physical recovery. Higher scores reflect less compensation. The Left Hemisphere and Right Hemisphere scales can be used to help determine whether an injury is diffuse or lateralized. A number of other scales and interpretive factors are also available (**Golden, Purish, & Hammeke, 1986** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib609)**).

We cannot review the voluminous literature on the LNNB, but brief mention of a few key studies certainly is merited. The reliability of the LNNB has been evaluated from the usual perspectives (split-half, internal consistency, and test-retest), with excellent results. For example, the mean test-retest reliability for the clinical scales was near .90 (**Bach, Harowski, Kirby, Peterson, & Schulein, 1981** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib72)** ; **Plaisted & Golden, 1982** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1308)** ; **Teichner et al., 1999** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1622)**). In various validity studies of classification of braindamaged persons versus other criterion groups, the LNNB has shown hit rates of 80 percent or better (Golden, Moses, Graber, & Berg, 1981; Hammeke, Golden, & Purish, 1978; Moses & Golden, 1979; Teichner et al., **1999** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1622)**).

In spite of the positive appraisals of the LNNB reported by Golden and his colleagues, some neuropsychologists remain skeptical of the test (e.g., **Lezak, 1995** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib981)**). One concern is that the heterogeneity of the scales is so great that the individual scale scores do not quantify specific neuropsychological deficits but instead serve only to differentiate normal persons from brain-damaged patients (**Snow, 1992** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1516)** ; **Van Gorp, 1992** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1686)**). Early reviewers

also expressed concern that the speech scales were not oriented to syndromes of aphasia and could therefore misdiagnose language deficits (**Delis & Kaplan, 1982** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib407)**). In defense of the LNNB, Purish (**2001** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1326)**) contends that initial criticisms were based on misconceptions as to the theoretical basis for the instrument. Furthermore, in his view, these criticisms have been largely negated by an expanding body of empirical research supporting the test.

Yet, it is possible that the LNNB and its chief rival, the Halstead-Reitan Neuropsychological Battery, have reached their peak of popularity and clinical utility (**Davis, Johnson, and D'Amato, 2005** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib393)**). New batteries emerge every few years. A promising addition is the Neuropsychological Assessment Battery.

The Neuropsychological Assessment Battery (NAB)

The Neuropsychological Assessment Battery or NAB (Stern & White, 2003ab) is a recent and promising entry in the field that is remarkable for its breadth and sophistication. Suitable for adults 18 to 97 years of age, the NAB is a comprehensive battery of 24 individual tests in five modular areas: attention, language, memory, spatial, and executive functions. Twelve of the subtests also can be used as a separate screening module. The instrument comes in two parallel and psychometrically equivalent versions, Form 1 and Form 2. Norms are based on data from 1,448 neurologically healthy individuals matching the U.S. population on educational level, gender, ethnicity, and geographic region.

The five major modules, each consisting of four to six subtests, are listed in **Table 10.11** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/ch10lev1sec21#ch10tbl11)** . Subtests used in the Screening Module are indicated with an asterisk. One feature evident in this table is that each module contains one subtest designed to possess ecological validity as well as psychometric validity. **Ecological validity** refers to the congruence between testing situations and analogous real-world circumstances. A test with strong ecological validity is one that highly resembles practical behaviors required in the real world. Among the NAB subtests with ecological validity are Driving Scenes, Bill Payment, Daily Living Memory, Map Reading, and Judgments. Each resembles a real world situation of importance in daily life. Ecological validity is beneficial because it increases the acceptability of testing to examinees.

The modular nature of the NAB allows for fixed administration of the entire battery (which takes about three hours), or flexible administration of the Screening Module followed by full administration of one or more of the five modules, depending on screening results. Once the test has been administered, software is available to compute a large array of output scores in a highly user-friendly computerized report. The module scores are reported as standard scores (M = 100, *SD* = 15), whereas the subtest scores are rendered as T-scores (M = 50, *SD* = 10).

The reliability of test scores is highly variable across the different modules and subtests, and is influenced by the examinee's age as well. The average coefficient alphas for the subtests in the five major modules revealed the following ranges (Stern & White, 2003b):

Attention Module:	.78 to .79
Language Module:	.48 to .84
Memory Module:	.47 to .86

Executive Functions Module:	.45 to .77

Test-retest reliability was evaluated with 95 individuals who were tested twice over an average span of 6 months. Understandably, these average coefficients were somewhat lower and more variable:

Attention Module:	.44 to .87
Language Module:	.23 to .70
Memory Module:	.41 to .61
Spatial Module	.13 to .68
Executive Functions Module:	.43 to .64

These relationships between test and retest NAB scores are respectable, given the lengthy test-retest interval.

The validity of the NAB is difficult to summarize concisely, because of the complexity of the instrument. The authors provide extensive documentation on validity, as evaluated from the traditional perspectives, including content validity, factor-analytic evidence of construct validity, and convergent and divergent correlations with similar and dissimilar external measures (all supportive). The authors conclude:

> Although the data presented in this chapter support the validity of the NAB, these data and analyses should be considered only the beginning steps in the ongoing process of test validation. (Stern & White, 2003b, p. 141)

TABLE 10.11 Modules and Subtests of the NAB

	Attention
Orientation* **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/ch10lev1sec21#ch10fn10)**	Questions about orientation to self, time, place, and situation
Digits Forward* **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/ch10lev1sec21#ch10fn10)**	Repetition of orally presented digit sequences of increasing length
Digits Backward* **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/ch10lev1sec21#ch10fn10)**	Orally presented digit sequences recalled in reverse order

Dots	Delayed recognition of the "new" dot in visual presentation of dots
Numbers & Letters* **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/ch10lev1sec21#ch10fn10)**	Timed tests of letter cancellation, letter counting, serial addition
Driving Scenes	Recognition of what is "new" in presentation of a second driving scene
	Language
Oral Production	Speech output when the examinee orally describes a picture
Auditory Comprehension	Comprehension of orally presented commands and instructions
Naming* **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/ch10lev1sec21#ch10fn10)**	Ability to name a pictured object, with cues if necessary
Reading Comprehension	Reading comprehension of single words and sentences
Writing	Writing sample scored for delivery, legibility, syntax, spelling
Bill Payment	Real world task of writing a check to pay a utility bill
	Memory

List Learning	Verbal learning of 12-word list with interference trial
Shape Learning* (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/ch10lev1sec21#ch10fn10)	Visual learning of 9 shapes with delayed recognition
Story Learning* (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/ch10lev1sec21#ch10fn10)	Verbal learning of a short narrative story of five sentences
Daily Living Memory	Verbal learning of medication instructions, address, phone number
	Spatial
Visual Discrimination	Matching of stimuli presented visually from an array
Design Construction	Assembling a tangram design from individual pieces
Figure Drawing	Drawing task involving copy and recall of geometric shapes
Map Reading	Answering practical questions based on the map of a city
	Executive Functions
Mazes* (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/ch10lev1sec21#ch10fn10)	Solving paper-and-pencil mazes of increasing complexity

Categories	Classifying and categorizing task based on photos of six people
Word Generation* (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/ch10lev1sec21#ch10fn10)	Creating three-letter words from two vowels and six consonants
Judgment	Answering practical questions about home safety and health

*Subtests used on the Screening Module.

Temple and Zgaljardic (**2009** (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1625)) provide independent evidence for the validity of the Screening Module of the NAB. They note strong associations with a measure of functional independence in a sample of 70 individuals with moderate-to-severe traumatic brain injury at a residential post-acute rehabilitation facility. Yet, Iverson, Williamson, Ropacki, and Reilly (2007) come down on the other side of the fence. In their study of 37 outpatients with neurological problems, results on the Screening Module were better than expected. In other words, in their sample the instrument did not show good sensitivity.

We need to keep in mind that the establishment of test validity is a dynamic process, not something set in stone when a test is released. The meaning of tests scores is sharpened and refined by ongoing research. Several recent reports support the validity of the NAB. For example, in a study of 54 patients with TBI and 54 matched controls, Donders and Levitt (**2012** (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib429)) found that the Attention, Executive Functions, and Memory modules were highly sensitive to brain impairment. Gavett et al. (**2012** (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib572)) reported that the Daily Living memory subtest provided the greatest accuracy in identifying patients with Alzheimer's Disease. It will prove interesting in the years ahead to see how additional studies bear on the validity of the NAB.

Baseline Testing With Brief Neuropsychological Test Batteries

As with most human attributes, variability in neu-rocognitive abilities from one person to the next is substantial. Some people are quick with reaction times, strong in memory skills, and facile with mathematical processing; others innately possess lower levels of ability; and, most of us are somewhere in between. Individual differences present a quandary in assessment, especially when the objective is to identify mild or subtle neuropsychological deficits such as mild traumatic brain injury (mTBI). When do low scores indicate mTBI and when do they signify a typical level of functioning? Access to baseline testing can prove invaluable in making this distinction. For at least two areas of assessment, the acquisition of baseline test data has become the expected practice.

One application of baseline testing is the Automated Neuropsychological Assessment Metrics (ANAM) Traumatic Brain Injury (TBI) Battery used in the armed forces. U.S. military troops deployed to war zones are administered the latest version, the ANAM4 TBI Battery, to obtain baseline neurocognitive performance levels. In situations where a soldier has been exposed to trauma such as an IED blast, retesting with the ANAM4 TBI Battery will help identify the presence of TBI, even if it is mild in severity. The battery was designed to minimize retesting effects by providing a nearly endless source of potential stimuli within each test module. Developed under the guidance of the U.S. Army, the battery is widely available and used in diverse settings worldwide.

The full ANAM4 consists of 22 assessments that can be grouped into flexible or standardized batteries. The subtests include measures of reaction time, learning, memory, mathematical processing, spatial processing, executive functions, and symptoms. Based on decades of study by dozens of neuropsychological and human performance researchers, the subtests are highly sensitive to the impact of brain injury, degenerative disease, toxin exposure, medication effects, and rehabilitation efforts. All modules are administered with a personal laptop computer. For the performance-based measures, stimuli are presented visually, and the left-right mouse buttons are used for the forced-choice options.

The ANAM4 TBI Battery consists of eight assessments that can be administered in about 20 minutes, making it highly feasible as a follow-up test when a soldier has been exposed to trauma such as an IED blast. The eight modules are listed in **Table 10.12** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/ch10lev1sec21#ch10tbl12)** . The ANAM4 software generates a full report providing the examiner with the current neurocognitive status of the soldier, comparisons to previous testing sessions, and comparisons to selected reference and norm groups. Researchers can transfer data in spreadsheet format to preferred statistical packages.

Normative data based on extraordinarily large samples are available for the ANAM4 TBI Battery. Vincent, Roebuck-Spencer, Gilleland, and Schlegel (**2012** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1701)**) collected test data from over 107,500 active duty service members 17 to 65 years of age. The norms are carefully stratified by age and gender. The main criticism of ANAM4 is the lack of research on its effectiveness in identifying mTBI in soldiers (**Kennedy & Moore, 2010** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib877)**). While it is clear that the individual subtests possess strong psychometric qualities, there is surprisingly little research on such matters as sensitivity and specificity of the overall battery in the identification of mTBI.

Another laptop-based neurocognitive battery is ImPACT (Immediate Post-Concussion Assessment and Cognitive Testing), developed in the 1990s by Mark Lovell and Joseph Maroon (**Lovell, 2006** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1008)** ; **Lovell, Iverson, Collins, and others, 2006** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1008)**). ImPACT is intended for sports settings to help in making return-to-play decisions following concussions. The 20-minute battery is widely used in clinical management of concussions for athletes ages 10 through adulthood. The instrument is intended for use when baseline results are available for individual team members. Impact is a highly popular computer-based testing program that is used in high school, college, and professional sports programs. It should be given only by persons trained in its administration and interpretation. The test developers caution that the battery should never be used as a "standalone" device for diagnosis or decision-making.

TABLE 10.12 Subtests of the ANAM4 TBI Battery

Sleepiness scale: A self-assessment of the soldier's sleepiness/fatigue level on a 7-point scale from "very alert" to "very sleepy."

Mood scale: A self-assessment of the user's mood state in seven categories (Vigor, Happiness, Depression, Anger, Fatigue, Anxiety, and Restlessness). A number of adjectives related to these mood categories are rated on a 7-point scale.

Simple reaction time (SRT): The user clicks the left mouse button when an asterisk appears on the screen at random intervals. A measure of attention and reaction time.

Code substitution: A display of digits 1 through 9 appears in a row at the top of the screen with a different symbol above each digit. A series of 72 individual probes appears at the bottom of the screen, each showing a pairing of a digit and symbol. The soldier clicks the left or right mouse button to signify a match or non-match, respectively, with the static display at the top of the screen. A measure of visual search, sustained attention, and encoding.

Procedural reaction time: A series of single digits (2, 3, 4, or 5) is presented in 32 trials. The user clicks the left mouse button to indicate the digit is "low" (2 or 3) or the right mouse button to indicate the digit is "high" (4 or 5). A measure of processing efficiency and rule-following.

Mathematical processing: A series of single-digit arithmetic equations (e.g., 3 + 4 − 1) is presented in 20 trials. The user clicks the left mouse button to indicate the answer is less than 5 or the right mouse button if the answer is greater than 5. A measure of basic computational skills, concentration, and working memory.

Matching to sample: A series of 4 × 4 matrices with cells in a 2-color format are presented in 20 trials. Following each stimulus, a pair of slightly different 4 × 4 matrices appears side-by-side. The user clicks the left or right mouse button to indicate the correct match to the previous stimulus. A measure of spatial processing and visuo-spatial working memory.

Code substitution delayed: A series of 36 probes appears as in the previous code substitution test. The soldier response in the same fashion, but must access memory of the static display, which is not represented. A measure of delayed recall for visual stimuli.

Source: Based on Eonta, S. E., Carr, W., McArdle, J. J., and others (**2011 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib461)**). Automated Neuropsychological Assessment Metrics: Repeated assessments with two military samples. *Aviation, Space, and Environmental Medicine, 82,* 34-39.

ImPACT typically is administered from a laptop computer by an athletic trainer, school nurse, team doctor, or psychologist to help determine when a player is ready to return to the field after a possible concussion from a hard "hit" or other head trauma. The six modules are described in **Table 10.13 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/ch10lev1sec21#ch10tbl13)** .

Dozens of published studies pertain to the reliability and validity of ImPACT. See ***impacttest.com (http://impacttest.com)*** for a listing of references. We will summarize here two studies on the sensitivity and specificity of test scores in predicting certain outcomes. The reader will recall that sensitivity refers to the percentage of respondents with a known condition who are correctly detected, whereas specificity refers to the percentage of respondents without the condition who are correctly designated. Lau, Collins, and Lovell (**2011 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib956)**) followed 108 male high school football players who sustained a concussion and then divided the group into protracted recovery (14 or more days) before returning to play, and short recovery (less than 14 days) before returning to play. A combination of four symptom clusters and four ImPACT scores yielded a

(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1460)) tested 12 recently concussed athletes with ImPACT and compared the data to results for 66 high school athletes with no history of concussion. The best discriminant function analysis correctly classified 82 percent of participants in the concussion group (sensitivity) and 89 percent of participants in the control group (specificity). These two studies support the overall utility of ImPACT.

TABLE 10.13 The Six Modules from the ImPACT Test Battery

Word Discrimination: A measure of attention and verbal recognition memory. Twelve target words are presented for 750 milliseconds each on the computer screen. The list is presented twice. The athlete is tested for recall with the presentation of a 24-word list that includes the 12 target words and 12 non-target words from the same semantic category. For example, if the target word was "carrot" the non-target word might be "celery." Using the mouse, the examinee clicks "yes" or "no" for each of the 24 stimuli.

Design Memory: A measure of attention and visual recognition memory. Twelve target designs are presented for 750 milliseconds each on the computer screen. The designs are presented twice. The athlete is tested for recall with the presentation of 24 designs that include the 12 target designs and 12 non-target designs consisting of the original designs rotated in space. Using the mouse, the examinee clicks "yes" or "no" for each of the 24 designs.

X's and O's: A measure of visual working memory and visual processing speed. The athlete views a screen of randomly placed X's and O's, three of which are illuminated in yellow, for 1.5 seconds. A distractor task ensues (click P key for a red circle, Q key for a blue square). Then, the screen of X's and O's reappears, but no letters are illuminated. The task of the respondent is to click on the stimuli previously illuminated in yellow.

Symbol Matching: A measure of visual processing speed, learning, and memory. The athlete is presented a screen depicting the digits 1 through 9, with a common symbol (circle, square, triangle) above each digit. Below this display a symbol is presented. The examinee clicks the corresponding digit as quickly as possible. If correct, the digit turns green, if wrong, it turns red. Eventually, the symbols above the digits disappear, so that correct responses depend upon memory. Results consist of both reaction time and number of correct responses.

Color Match: A measure of choice reaction time, impulse control, and response inhibition. A brief test of color blindness first is given (ability to perceive the colors red, blue, and green). Next, a color word appears, either in the matching color (e.g., the word RED in red ink) or nonmatching color (e.g., the word BLUE in green ink). The athlete clicks the mouse if the word and color match, otherwise waits for the next stimulus. Both reaction time and errors are assessed.

Three Letter Memory: A measure of working memory and visual-motor response speed. Three consonant letters are displayed on the screen. A distractor task ensues. This consists of the numbers 1 through 25 randomly placed in a 5 × 5. The athlete is instructed to click on the numbers in reversed order, 25 to 1. Then, the examinee is asked to recall the three consonant letters. Five trials are presented. This module yields a memory score (total number of consonants correctly recalled) and a distractor score (total number of digits clicked in the correct order).

Source: Based on descriptions from ***impacttest.com (http://impacttest.com)*** and Lovell (**2006 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1008)**).

But the battery is not without its critics. ESPN contributor Peter Keating (2012) cites a concern about the high false positive rate, and notes the conflict of interest in which the test developers, who have published

the vast majority of research on the battery, also are involved in marketing the battery for profit. Further, he notes that

> . . . in practice, it's hard for any neuropsychological test to get good data. Some athletes intentionally try to perform poorly on baselines so their post-injury tests won't keep them out of play. Peyton Manning [Denver Broncos quarterback] admitted to this practice, which players call sandbagging, in April 2011 (ESPN The Magazine, "Concussion Test May Not Be Panacea," August 26, 2012).

After reviewing the available research, Mayers and Redick (**2012 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1076)**) conclude that the empirical evidence does not support the use of the battery for making return-to-play decisions. ImPACT likely serves a positive purpose by sensitizing players, coaches, and others to the dangers of repeated concussion. But as the test developers acknowledge, test results alone should never be the basis for important decisions like returning to play after head trauma.

The stakes are high for athletes and their families. In the long-term, repeated blows to the head are known to cause chronic traumatic encephalopathy (CTE), a degenerative brain disease associated with memory loss, confusion, aggression, impulse control problems, Parkinsonian symptoms (tremor, gait abnormalities, slurred speech), and, eventually, progressive dementia (**Saulle & Greenwald, 2012 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1440)**). Even "minor" blows to the head that do not result in serious symptoms can lead to CTE if they occur with sufficient frequency, as in boxing or football (**McKee, Cantu, Nowinski, and others, 2009 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1101)**). In a recent post-mortem analysis of brain tissue in 85 former football players, hockey players, and military veterans, McKee, Stein, Nowinski, and others (**2012 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1102)**) concluded that "for some athletes and war fighters, there may be severe and devastating long-term consequences of repetitive brain trauma that has traditionally been considered only mild (p. 20)." As a society, we may want to reconsider the glamorization of contact sports like football, boxing, and hockey.

10.22 SCREENING FOR ALCOHOL USE DISORDERS

The ways in which people can abuse alcohol include a spectrum of misfortune and tragedy ranging from an occasional hangover to, literally, drinking oneself to death. But clinicians and researchers generally recognize two diagnoses: alcohol abuse and alcohol dependence (**American Psychiatric Association, 2000** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib33)**). Loosely speaking, the more generic syndrome of alcoholism refers to either diagnosis. A full discussion of these syndromes is not justified here, but a brief summary is warranted. Interestingly, neither alcohol abuse nor dependence is defined by ingestion of a particular amount of alcohol, although substantial quantities typically are involved. The criteria for alcohol abuse refer to the functional impact of drinking on the life of the patient. In particular, if an individual meets one or more of four criteria, a diagnosis of alcohol abuse is defensible. Briefly, the criteria are:

- Drinking interferes with important life responsibilities at work, home, or school.
- Drinking leads to unsafe behavior such as driving while intoxicated.
- Drinking causes persistent legal problems such as arrests for fighting.
- Drinking leads to conflict with a spouse or significant other.

In addition to meeting one or more of these criteria, the patient must not meet the criteria for a diagnosis of substance dependence, which often entails a more serious and chronic syndrome. Specifically, if a patient meets three or more of seven criteria, a diagnosis of **alcohol dependence** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm01#bm01gloss07)** is warranted. Briefly, the criteria are:

- Tolerance or needing increasingly more alcohol to get the same effect.
- Withdrawal symptoms such as tremor when drinking ceases.
- Drinking in greater quantities or for longer periods than intended.
- Desire to cut down but unsuccessful efforts to control drinking.
- Spending large amounts of time using alcohol or recovering from use.
- Giving up important social, occupational, or recreational activities to drink.
- Continued use in spite of demonstrable health problems such as an ulcer.

Given the high prevalence of alcohol use disorders in the United States, it is nearly inevitable that psychologists and other clinicians will encounter patients who experience problems in this spectrum. Fortunately, there are several simple devices useful for screening and assessment, which we review here. In some cases, these tools are pristinely simple and consist of the clinician casually asking a handful of "yes-no" questions. In other cases, a more traditional paper-and-pencil questionnaire is needed.

The CAGE questionnaire is a short screening instrument that consists of the practitioner asking if the client has thought about Cutting down on drinking, become Annoyed by criticism of his or her drinking, felt Guilty about his or her drinking, or had an Eye-opener drink in the morning. A simple "yes-no" question pertinent to each symptom is asked as part of a general health history. The exact wording of this copyrighted instrument can be found in Ewing (**1984** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib467)**). The endorsement of even a single item suggests the presence of an alcohol use disorder, whereas saying "yes" to two or more items virtually guarantees that the patient will meet the criteria for alcohol abuse or dependence. Research indicates that the tool is more effective when it is not preceded by questions about how much or how often the patient drinks (**Steinweg & Worth, 1993**

questions about quantity and frequency trigger denial in the patient, making accurate assessment nearly impossible. The CAGE questionnaire has proved valuable as a screening tool in numerous locations, including general psychological practice and medical settings. In one study of a "walk-in" or immediate-care Veterans hospital clinic, the test correctly identified 86 percent of patients later confirmed to have alcoholism and accurately ruled out 93 percent of patients later confirmed not to have alcohol problems. Astonishingly, the prevalence rate for alcoholism was determined to be 22 percent for this largely male clinic population (**Liskow, Campbell, Nickel, & Powell, 1995** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1000)**).

A recent epidemiological study conducted in and around Paris, France, casts doubt on the usefulness of the CAGE test as a screening device for alcoholism (**Messiah, et al., 2007** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1133)**). In 2005, the researchers conducted a follow-up to a 1991 study of 1,991 participant responses to the Cut-down, Annoyed, Guilt, and Eye-opener (CAGE) questionnaire through telephone interview of 5,382 residents. The time period in question, 1991 to 2005, was an era in which alcohol consumption was known to be in decline, so it was surprising to the researchers when they found that the percentage of respondents endorsing each of the symptoms had increased substantially. In fact, the magnitude of the paradoxical increase astonished the researchers. For example, when asked whether they had thought about cutting down on their drinking, the percentage of respondents who answered "yes" increased from 4.3 percent in 1991 to 16.6 percent in 2005. The researchers speculate that the results might indicate the emergence of a temperance movement in France. Whether or not this is true, the findings most certainly cast doubt on the value of the CAGE in general population surveys.

Some researchers find that the CAGE questionnaire is more effective for screening with men than women (**Cherpitel, 2002** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib300)**). In response to this shortcoming, a similar instrument called the TWEAK questionnaire was developed specifically for women. The acronym refers to Tolerance for drinking, Worried friends or relatives, Eye-opener to get going in the morning, Amnesia for things done or said while drinking, and feeling the need to Kut down on intake (**Russell, Martier, Sokol, and others, 1994** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1419)**). TWEAK is scored on a 7-point scale, with the first two items earning two points each, the last three items earning one point each. A total score of two or more points indicates the likelihood of an alcohol problem. TWEAK is highly accurate in screening for alcohol problems in women (**Bradley, Boyd-Wickizer, Powell, & Burman, 1998** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib198)**).

CAGE and TWEAK are by no means the only acronymic screening tools for alcohol problems. Other instruments include the five-item RAPS questionnaire or Rapid Alcohol Problems Screen (Cherpitel, 1995) and the 10-item AUDIT questionnaire or Alcohol Use Disorders Identification Test (Saunders, Aasland, Babor, and others, 1993). A huge amount of effort was invested in the development and validation of the AUDIT questionnaire. Research on this instrument was underwritten by the World Health Organization (WHO), and the scale has been translated into many languages.

Dozens of additional screening tests could be mentioned, but we want to close this section by reviewing an interesting scale that embodies some appealing methods of test construction. The Substance Abuse Subtle Screening Inventory-3 or SASSI-3 (**Miller, Roberts, Brooks, & Lazowski, 1997** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1146)**) consists of two types of questions: obvious and subtle. The obvious questions include 26 behaviors that are endorsed on a 4-point Likert-type continuum ranging from *never* to *repeatedly.* These questions embody high face validity and are on a par with "I have taken drugs to improve how I feel" and "I have had more to drink

of the attitudes and behaviors that commonly accompany substance abuse. These questions are on par with "I probably break the law more than others" and "I tend to be a responsible person" [reverse scored]. Both types of items—obvious and subtle—were carefully validated during test construction.

Test construction consisted of administering a large group of preliminary items to three groups of individuals: substance abusers, non-substance abusers, and substance abusers instructed to "fake good." The SASSI-3 emerged after this large pool of items was winnowed down to a smaller number, based on group contrasts. The resulting instrument includes the direct items—those that discriminated substance abusers from non-substance abusers, and the indirect items—those that discriminated the "fake-good" substance abusers from non-substance abusers. In addition to the adult scale, an adolescent version now has been published, and the instrument is available for supervised online administration. A Spanish version also is available.

The test developers report excellent reliability for the SASSI-3, with two-week test-retest stability coefficients for 40 respondents ranging from .92 to 1.00 for the subscales and coefficient alpha of .93 for the test overall. A validity study of 419 respondents revealed a 95 percent rate of correct classification for substance abusers and a 93 percent correct classification rate for non-substance abusers—very impressive results for a short screening test (**Miller & Lazowski, 1999** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1145)**). Laux, Salyers, and Kotova (**2005 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib957)**) found strong test-retest reliability with the SASSI-3 in a sample of 103 college students, reporting $r = .94$ over a one-week period. Feldstein and Miller (**2007** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib494)**) reviewed 36 studies on all editions of the SASSI and weigh in skeptically, citing high rates of false positives. They propose that public domain instruments (e.g., CAGE, AUDIT) perform just as well and have the added advantage of being free.

The SASSI-3 appears to be a capable tool. Yet, given the frequency of its use—the instrument has been administered *millions* of times—it is disconcerting that few independent studies have been published (**Gray, 2001 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib642)**). A search of PsychInfo yielded only 15 studies on the test, and the majority of these were unpublished doctoral dissertations. More research is needed to corroborate the value of this promising inventory.

Mini-Mental State Exam

The most widely used mental status tool with the elderly is the Mini-Mental State Examination (MMSE), a 5- to 10-minute screening test that yields an objective global index of cognitive functioning (**Folstein, Folstein, & McHugh, 1975** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib520)** ; **Tombaugh, McDowell, Kristjansson, & Hubley, 1996** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1654)**). The test contains 30 scorable items having to do with orientation, immediate memory, attention, calculation, language production, language comprehension, and design copying. The items are so easy that normal adults almost always obtain scores in the range of 27 to 30 points (**Figure 10.9** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/ch10lev1sec22#ch10fig9)**).

The reliability of this simple instrument is excellent. Folstein et al. (**1975** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib520)**) report a 24-hour test-retest reliability of .89 for 22 patients with varied depressive symptoms. Reliability over a 28-day period for 23 clinically stable patients with diagnoses of dementia, depression, and schizophrenia was an

impressive .99. Normative data are available from several sources (e.g., **Lindal & Stefansson, 1993** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib992)** ; **Tombaugh, McDowell, Kristjansson, & Hubley, 1996** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1654)**).

Using a cutting score of 23 or below as abnormal and 24 or above as normal, the MMSE is about 80 to 90 percent accurate in identifying elderly patients with suspected Alzheimer's disease or other dementia. This cutting score produces few false-positives (normal patients classified as having dementia). The sensitivity of the instrument depends on a number of factors, including the cutting score used, the educational level of the examinee, the extent of the dementia, the nature of the underlying pathology, and the type of setting in which assessments are undertaken (**Anthony, LeResche, Niaz, Von Korff, & Folstein, 1982** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib57)** ; **Tombaugh, McDowell, Kristjansson, & Hubley, 1996** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1654)** ; **Tsai & Tsuang, 1979** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1670)**). In spite of its limitations, the MMSE remains the most reliable and practical screening test for dementia in the elderly (**Ferris, 1992** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib497)**). Drebing, Van Gorp, Stuck, and others (**1994** **(http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib436)**) recommend its use as part of a short screening battery for cognitive decline in the elderly.

5	Orientation to Time (day, date, month, season, and year)
5	Orientation to Place (floor, building, city area, city, state)
3	Immediate Memory (three words presented orally)
5	Attention and Calculation (serial 7s, five subtractions)
3	Delayed Recall (three words presented orally above)
2	Naming (pencil and watch)
1	Repetition (brief sentence presented orally)
3	Comprehension (follow simple three-part oral command)
1	Reading (read simple command and obey)
1	Writing (compose a simple sentence)
1	Drawing (reproduce two intersecting pentagons)
30	Total

FIGURE 10.9 Scoring Weights and Domains of the Mini-Mental State Examination

Research on the MMSE continues unabated. A search of PsychINFO for articles with "MMSE" in the title yielded 128 hits with 27 of them published since 2010. A final caution is worth mentioning. The MMSE has become so popular that some practitioners use MMSE total scores as a shortcut toward a diagnosis of dementia (**Nieuwenhuis-Mark, 2010 (http://content.thuzelearning.com/books/Gregory.8055.17.1/sections/bm02#bm02bib1234)**). Tests should never be used as a substitute for clinical judgment.

PLEASE USE REQUIRED SOURCES

(ASSIGNMENT 3)

In this assignment you will choose three general topics of interest to you related to psychological assessment from the list below.

Individuals with post-traumatic stress disorder

Individuals with intellectual disabilities

Individuals with schizophrenia spectrum and other psychotic disorders

Research five peer-reviewed articles in the Ashford University Library that were published within the last 15 years, including a minimum of one article for each of your three chosen topics. In your paper, begin with an introduction that describes the role of assessment in diagnosis and treatment. Using your researched articles, compare at least two psychological or educational tests and/or assessment procedures for each of the topics chosen. Analyze and describe the psychometric methodologies employed in the development and/or validation of the tests and/or assessment procedures associated with each of the three topics. Debate any relevant approaches to assessment of the constructs being evaluated by any tests and assessments you described. Include an analysis of any challenges related to assessing individuals from diverse social and cultural backgrounds for each of the three topics. Conclude by evaluating the ethical and professional issues that influence the interpretation of testing and assessment data. You may cite from your textbook to assist you in the development of your introduction and the conclusion of your paper; however, the textbook cannot count as one of the five required peer-reviewed articles.

Writing the Literature Review
The Literature Review:

- Must be 6 to 8 double-spaced pages in length (not including title or references pages) and must be formatted according to APA style as outlined in the Ashford Writing Center (Links to an external site.).
- Must include a title page with the following:
  - Title of paper
  - Student's name
  - Course name and number
  - Instructor's name
  - Date submitted
- Must begin with an introduction that describes the role of assessment in diagnosis and treatment.
- Must address the topic of the paper with critical thought.
- Must end with a conclusion that summarizes your evaluation addressing the current use of psychological tests and the role of assessment in diagnosis and treatment.
- Must use at least five peer-reviewed sources published within the last 15 years, all of which come from the Ashford University Library.
- Must document all sources in APA style as outlined in the Ashford Writing Center (Links to an external site.).

Journal of Clinical Child & Adolescent Psychology, 42(1), 139–159, 2013

ISSN: 1537-4416 print/1537-4424 online
DOI: 10.1080/15374416.2012.736358

Routledge
Taylor & Francis Group

FUTURE DIRECTIONS

Future Directions in Psychological Assessment: Combining Evidence-Based Medicine Innovations with Psychology's Historical Strengths to Enhance Utility

Eric A. Youngstrom

Departments of Psychology and Psychiatry, University of North Carolina at Chapel Hill

Assessment has been a historical strength of psychology, with sophisticated traditions of measurement, psychometrics, and theoretical underpinnings. However, training, reimbursement, and utilization of psychological assessment have been eroded in many settings. Evidence-based medicine (EBM) offers a different perspective on evaluation that complements traditional strengths of psychological assessment. EBM ties assessment directly to clinical decision making about the individual, uses simplified Bayesian methods explicitly to integrate assessment data, and solicits patient preferences as part of the decision-making process. Combining the EBM perspective with psychological assessment creates a hybrid approach that is more client centered, and it defines a set of applied research topics that are highly clinically relevant. This article offers a sequence of a dozen facets of the revised assessment process, along with examples of corollary research studies. An eclectic integration of EBM and evidence-based assessment generates a powerful hybrid that is likely to have broad applicability within clinical psychology and enhance the utility of psychological assessments.

What if we no longer performed psychological assessment? Although assessment has been a core skill and a way of conceptualizing individual differences central to psychology, training and reimbursement have eroded over a period of decades (Merenda, 2007b). Insurance companies question whether they need to reimburse for psychological assessment (Cashel, 2002; Piotrowski, 1999). Educational systems have moved away from using ability-achievement discrepancies as a way of identifying learning disability and decreased the emphasis on individual standardized tests for individual placement (Fletcher, Francis, Morris, & Lyon, 2005). Several traditional approaches to personality assessment, such as the various interpretive systems for the Rorschach, have had their validity challenged repeatedly (cf. Meyer & Handler, 1997; Wood, Nezworski, & Stejskal, 1996). Many graduate-level training programs are reducing their emphasis on aspects of assessment (Belter & Piotrowski, 2001; Childs & Eyde, 2002; Stedman, Hatch, & Schoenfeld, 2001) and psychometrics (Borsboom, 2006; Merenda, 2007a) in their curricula, and few undergraduate programs offer courses focused on assessment or measurement. Efforts to defend assessment have been sometimes disorganized and tepid, or hampered by a lack of data even when committed and scholarly (Meyer et al., 1998).

Is this intrinsically a bad thing? Training programs, systems of care, and providers all have limited resources. Assessment might be a luxury in which some could afford to indulge, paying for extensive evaluations as a way to gain insight into themselves. However, arguments defending assessment as a major clinical activity need to appeal to utility to be persuasive (Hayes, Nelson, & Jarrett, 1987). Here, "utility" refers to adding value to individual care, where the benefits deriving from the assessment procedure clearly outweigh the costs, even when the costs combine fiscal expense with other

Thanks to Guillermo Perez Algorta for comments and suggestions
Correspondence should be addressed to Eric A. Youngstrom, Department of Psychology, University of North Carolina at Chapel Hill, CB #3270, Davie Hall, Chapel Hill, NC 27599-3270. E-mail: eay@unc.edu

factors such as time and the potential for harm (Garb, 1998; Kraemer, 1992; Straus, Glasziou, Richardson, & Haynes, 2011). Although utility has often been described in contexts of dichotomous decision making, such as initiating a treatment or not, or making a diagnosis or not, it also applies to situations with ordered categories or continuous variables. Conventional psychometric concepts such as reliability and validity are prerequisites for utility, but they do not guarantee it. Traditional evaluations of psychological testing have not formally incorporated the concept of costs in either sense—fiscal or risk of harm.

Using utility as an organizing principle has radical implications for the teaching and practice of assessment. Assessment methods can justify their place training and practice if they clearly address at least one aspect of prediction, prescription, or process—the "Three Ps" of assessment utility (Youngstrom, 2008). *Prediction* refers to association with a criterion of importance, which could be a diagnosis, but also could be another category of interest, such as adolescent pregnancy, psychiatric hospitalization, forensic recidivism, graduation from high school, or suicide attempt. For our purposes, the criterion could be continuous or categorical, and the temporal relationship could be contemporaneous or prospective. The goal is to demonstrate predictive validity for the assessment procedure by any of these methods and to make a compelling case that the effect size and cost/benefit ratio suggest utility. *Prescription* refers more narrowly to the assessment providing information that changes the choice of treatment, either via matching treatment to a particular diagnosis or by identifying a moderator of treatment. Similarly, *process* refers to variables that inform about progress over the course of treatment and quantify meaningful outcomes. These could include mediating variables, or be measures of adherence or treatment response. Each of the Three Ps demonstrates a connection to prognosis and treatment. These are not the only purposes that could be served by psychological assessment, but they are some of the most persuasive in terms of satisfying stakeholders that the assessment method is adding value to the clinical process (Meehl, 1997). Many of the other conventional goals of psychological assessment (Sattler, 2002) can be recast in terms of the Three Ps and utility: Using assessment as a way of establishing developmental history or baseline functioning may have predictive value or help with treatment selection, as can assessment of personality (Harkness & Lilienfeld, 1997). Case formulation speaks directly to the process of working effectively with the individual. Gathering history for its own sake is much less compelling than linking the findings to treatment and prognosis (Hunsley & Mash, 2007; Nelson-Gray, 2003).

It was surprising to me as an educator and a psychologist how few of the commonly taught or used techniques can demonstrate any aspect of prediction, prescription, or process—let alone at a clinically significant level (Hunsley & Mash, 2007). Surveys canvassing the content of training programs at the doctoral and internship level (Childs & Eyde, 2002; Stedman et al., 2001; Stedman, Hatch, Schoenfeld, & Keilin, 2005), as well as evaluating what methods are typically used by practicing clinicians (Camara, Nathan, & Puente, 1998; Cashel, 2002), show that people tend to practice similar to how they were trained. There is also a striking amount of inertia in the lists, which have remained mostly stable for three decades (Childs & Eyde, 2002). Content has been set by habits of training, and these in turn have dictated habits of practice that change slowly if at all.

When I first taught assessment, I used the courses I had taken as a graduate student as a template and made some modifications after asking to see syllabi from a few colleagues. The result was a good, conventional course; but the skills that I taught had little connection to the things that I did in my clinical practice as I pursued licensure. Much of my research has focused on assessment, but that created a sense of cognitive dissonance compared to my teaching and practice. One line of research challenged the clinical practice of interpreting factor and subtest scores on cognitive ability tests. These studies repeatedly found little or no incremental validity in more complicated interpretive models (e.g., Glutting, Youngstrom, Ward, Ward, & Hale, 1997), yet they remained entrenched in practice and training (Watkins, 2000). The more disquieting realization, though, was that my own research into assessment methods was disconnected from my clinical work. If conventional group-based statistics were not changing my own practice, why would I put forth my research to students or to other practitioners? Why was I not using the assessments I taught in class? When I reflected on the curriculum, I realized that I was teaching the "same old" tests out of convention, or out of concern that the students needed to demonstrate a certain degree of proficiency with a variety of methods in order to match at a good internship (Stedman et al., 2001).

What was missing was a clear indication of utility for the client. Reviewing my syllabi, or perusing any of the tables ranking the most popular assessment methods, emphasized the disconnect: Does scoring in a certain range on the Wechsler tests make one a better or worse candidate for cognitive behavioral therapy? Does verbal ability moderate response to therapies teaching communication skills? How does the Bender Gestalt test do at predicting important criteria? Do poor scores on it prescribe a change in psychological intervention? . . . or or tell about the process of working with a client? . . . - What about Draw a Person? Our most widely used tools do not have a literature establishing their validity in terms of individual prognosis or treatment, and viewed

through the lens of utility they look superfluous. Yet these are all in the top 10 most widely used for assessing psychopathology in youths, according to practitioner surveys (Camara et al., 1998; Cashel, 2002), even though they do not feature prominently in evidence-based assessment recommendations (Mash & Hunsley, 2005).

Evidence-based medicine (EBM) is rooted in a different tradition, grounded in medical decision making and initially advocated by internal medicine and other specialties bearing little resemblance to the field of psychology (Guyatt & Rennie, 2002; Straus et al., 2011). EBM has grown rapidly, however, and it has a variety of strengths that could reinvigorate psychological assessment practices if there were a way to hybridize the two traditions (Bauer, 2007). The principles of emphasizing evidence, and integrating nomothetic data with clinical expertise and patient preferences, are consistent with the goals of "evidence-based practice" (EBP) in psychology (Spengler, Strohmer, Dixon, & Shivy, 1995; Spring, 2007). Indeed, the American Psychological Association (2005) issued a statement endorsing EBP along the lines articulated by Sackett and colleagues and the Institute of Medicine. However, this is more agreement about a vision; and there is a fair amount of work involved in completing the merger of the different professional traditions. In much of what follows, I refer to EBM instead of EBP when talking about assessment, because EBM has assessment-related concepts that have not yet been discussed or assimilated in EBP in psychology. Key components include a focus on making decisions about individual cases, and knowing when there is enough information to consider something "ruled out" of further consideration or "ruled in" as a focus of treatment. EBM also has a radical emphasis on staying connected to the research literature, including such advice as "burn your textbooks—they are out of date as soon as they are published" (Straus et al., 2011). The emphasis on scientific evidence as guiding clinical practice seems philosophically compatible with the Boulder Model of training, and resonates with recent calls to further emphasize the scientific components of clinical psychology (McFall, 1991).

EBM's focus on relevance to the individual puts utility at the forefront: Each piece of evidence needs to demonstrate that it is valid and that it has the potential to help the patient (Jaeschke, Guyatt, & Sackett, 1994). However, most discussions of EBP in psychology have focused on therapy, with less explication of the concepts of evidence-based assessment (see Mash & Hunsley, 2005, for comment). Despite the shared vision of EBM and the American Psychological Association's endorsement of EBP, most of the techniques and concepts involved in assessment remained in distinct silos. For example, the terms "diagnostic likelihood ratio," "predictive power," "wait-test" or "test-treat threshold," or even "sensitivity" or "specificity" are not included as index terms in the current edition of *Assessment of Children and Adolescents* (Mash & Barkley, 2007; these terms are defined in the assessment context later in this article). A hand search of the volume found five entries in 866 pages that mentioned receiver operating characteristic analysis or diagnostic sensitivity or specificity (excluding the chapter on pediatric bipolar disorder, which was heavily influenced by the EBM approach). Of those five, one was a passing mention of poor sensitivity for an autism screener, and the other four were the exceptions among a set of 77 trauma measures reviewed in a detailed appendix. Discussions of evidence-based assessment have focused on reliability and classical concepts of psychometric validity but not application to individual decision making in the ways EBM proposes (Hunsley & Mash, 2005; Mash & Hunsley, 2005).

Conversely, treatments of EBM barely mention reliability and are devoid of psychometric concepts such as latent variables, measurement models, or differential item functioning (Guyatt & Rennie, 2002; Straus et al., 2011), despite the fact that these methods are clearly relevant to situations where the "gold standard" criterion diagnosis is missing or flawed (Borsboom, 2008; Kraemer, 1992; Pepe, 2003). Similarly, differential item functioning, tests of structural invariance, and the frameworks developed for testing statistical moderation would advance EBM's stated goals of understanding the factors that change whether the research findings apply to the individual patient (i.e., what are the moderating factors?; Cohen, Cohen, West, & Aiken, 2003) and understanding the process of change (i.e., the mediating variables; MacKinnon, Fairchild, & Fritz, 2007).

The two traditions have much to offer each other (Bauer, 2007). Because the guiding visions are congruent, it is often straightforward to transfer ideas and techniques between the EBM and psychological assessment EBP silos. The ideas from EBM have reshaped how I approach research on assessment, and reorganized my research and teaching to have greater relevance to individual cases. Our group has mostly applied these principles to the assessment of bipolar disorder (e.g., Youngstrom, 2007; Youngstrom et al., 2004; Youngstrom, Freeman, & Jenkins, 2009), but the concepts are far more broad. In the next section I lay out the approach to assessment as a general model and discuss the links to both EBM and traditional psychological assessment. This is not an introduction to EBM; there are comprehensive resources available (Guyatt & Rennie, 2002; Straus et al., 2011). Instead, I briefly describe some of the central features from the EBM approach to assessment and then lay out a sequence of steps for integrating these ideas with clinical psychology research and practice. The synthesis defines a set of new research questions and methods that are highly clinically

relevant, and it reorganizes assessment practice in a way that is pragmatic and patient focused (Bauer, 2007). The combination of EBM and psychological assessment also directly addresses the "utility gap" in current assessment practice and training (Hunsley & Mash, 2007). Sections describing research are oriented toward filling existing gaps, not reinforcing any bifurcation of research from practice.

A BRIEF OVERVIEW OF ASSESSMENT IN EBM

EBM focuses on shaping clinical ambiguity into answerable questions and then conducting rapid and focused searches to identify information that addresses each question (Straus et al., 2011). Rather than asking, "What is the diagnosis?" an EBM approach would refine the question to something like, "What information would help rule in or rule out a diagnosis of attention deficit/hyperactivity disorder (ADHD) for this case?" EBM references spend little time talking about reliability and almost no space devoted to traditional psychometrics such as factor analyses or classical descriptions of validity (cf. Borsboom, 2006; Messick, 1995). Instead, they concentrate on a Bayesian approach to interpreting tests, at least with regard to activities such as screening, diagnosis, and forecasting possible harm. The core method involves estimating the probability that a patient has a particular diagnosis, or will engage in a behavior of interest (such as relapse, recidivism, or self-injury), and then using Bayesian methods to combine that prior probability with new information from risk factors, protective factors, or test results to revise the estimate until the revised probability is low enough to consider the issue functionally "ruled out," or high enough to establish the issue as a clear target for treatment (Straus et al., 2011).

Bayes' Theorem, a way of combining probabilities, is literally centuries old (Bayes & Price, 1763). There are two ways of interpreting Bayes' Theorem: A Bayesian interpretation focuses on the degree to which new evidence should rationally change one's degree of belief, whereas a frequentist interpretation connects the inverse probabilities of two events, formally expressed as:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (1)$$

In this formula, P(A) is the prior probability of the condition, before knowing the assessment result; P(A|B) is the posterior probability, or the revised probability taking into account the information value of the assessment result; and P(B|A)/P(B) conveys the degree of support that the assessment result provides for the condition, by comparing the probability of observing the result within the subset of those that have the condition, P(B|A), to the overall rate of the assessment result, P(B). For example, if 20% of the cases coming to a clinical practice have depression—base rate = P(A) = 20%—and the client scores high on a test with 90% diagnostic sensitivity to depression—P(B|A) = 90%, %, or 90% of cases with depression scoring positive—then Bayes' Theorem would combine these two numbers with the rate of positive test results regardless of diagnosis to generate the probability that the client has depression conditional upon the positive test result. If 30% of cases score positive on the test regardless of diagnosis (what Kraemer, 1992, called the "level" of the test, to distinguish it from the false alarm rate), then the probability that the client has depression rises to 60%. Conversely, if the client had scored below threshold on the same test, then the probability of depression drops to less than 3%. The example shows the potential power of directly applying the test results to the individual case but also illustrates the difficulty of combining the information intuitively, as well as the effort involved in traditional implementations of the Bayesian approach.

Luminaries in clinical psychology such as Paul Meehl (Meehl & Rosen, 1955), Robyn Dawes (Dawes, Faust, & Meehl, 1989), and Dick McFall (McFall & Treat, 1999) have advocated incorporating it into everyday clinical practice. Some practical obstacles have delayed the widespread adoption of the method, including that it requires multiple steps and some algebra to combine the information, and the posterior probability is heavily dependent on the base rate of the condition. An innovation of the EBM approach is to address these challenges by offering online calculators or a "slide rule" visual approximation, a probability nomogram (see Figure 1), avoiding the need for computation, albeit at the price of some loss in precision (Straus et al., 2011). The nonlinear spacing of the markers on each line geometrically accomplishes the same effect as transforming prior probabilities (the left-hand line of the nomogram) into odds, then multiplying by the change in the diagnostic likelihood (plotted on the center line) to extrapolate to the posterior probability (the right-hand line), again avoiding the algebra to convert the posterior odds back into a probability (see the appendix, or Jenkins, Youngstrom, Washburn, & Youngstrom 2011, for a worked illustration).

A second, more conceptual innovation developed by EBM is to move past dichotomous "positive test/negative test result" thinking and to suggest a multitiered way of mapping probability estimates onto clinical decision making. In theory, the probability estimate of a target condition could range from 0% to 100% for any given case. In practice, almost no cases would have estimated probabilities of exactly 0% or 100%, and few might even get close to those extremes given the limits of currently available assessment methods. The pragmatic insight is that we do not need such

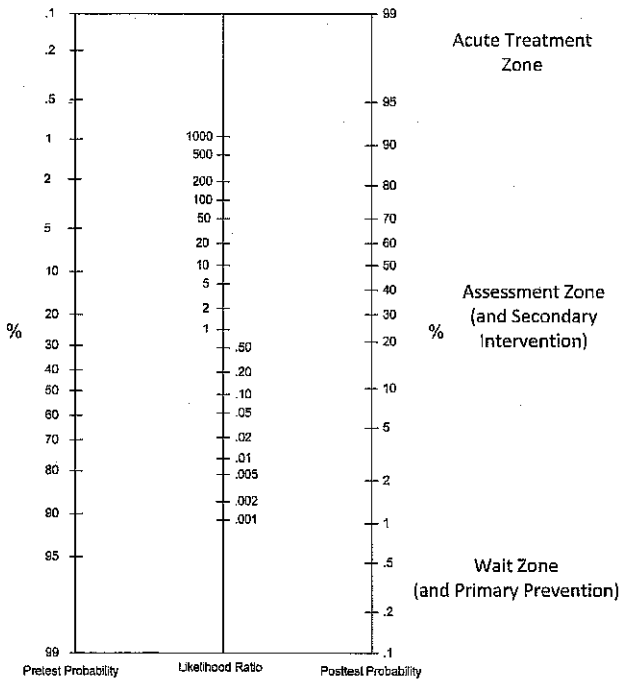


FIGURE 1 Probability nomogram for combining probability with likelihood ratios. *Note:* Straus et al. (2011) provided the rationale and examples of using the nomogram. Jenkins et al. (2011) illustrated using it with a case of possible pediatric bipolar disorder, and Frazier and Youngstrom (2006) with possible attention deficit/hyperactivity disorder.

extreme probability levels in order to make most clinical decisions (Straus et al., 2011). If the revised probability is high enough, then it makes sense to initiate treatment, in the same way that if the weather forecast calls for a 95% chance of showers, then we would do well to dress for rain. EBM calls the threshold where it makes sense to initiate treatment the "test-treat threshold"—probabilities above that level indicate intervention, whereas below that same point suggest continued assessment (Straus et al., 2011). Similarly, there is a point where the probability is sufficiently low to consider the target condition "ruled out" even though the probability is not zero. Below this "wait-test" threshold, EBM argues that there is no utility in continued assessment, nor should treatment be initiated. The two thresholds divide the range of probabilities and map them onto three clinical actions: actively treat, continue assessing, or decide that the initial hypothesis is not supported—and either assess or treat other issues (Guyatt & Rennie, 2002; Straus et al., 2011).

A third innovation in EBM is not to specify the exact locations for the wait-test and test-treat thresholds a priori. Instead, EBM provides a framework for incorporating the costs and benefits attached to the diagnosis, the test, and the treatment, and then using them to help decide where to set the bars for a particular case (Straus et al., 2011). Even better, there are ways of engaging the patient and soliciting personal preferences, including them in the decision-making process. For effective, low-risk, low-cost interventions, the treatment threshold might be so low that it makes sense to skip the assessment process entirely, as happens with routine vaccinations, or with the addition of fluoride to drinking water (Youngstrom, 2008). Conversely, for clinical issues where the treatment is freighted with risks, it makes sense to reserve the intervention until the probability of the target diagnosis is extremely high. For many families, atypical antipsychotics may fall in that category, given the serious side effects and the relative paucity of information about long-term effects on development (Correll, 2008). The EBM method creates a process for collaboratively weighing the costs, benefits, and preferences. This has the potential to empower the patient and customize treatment according to key factors, and it moves decision making from a simple, dichotomous mode to much more nuanced gradations. For the same patient, the test-treat thresholds might be more stringent for initiating medication than therapy, and so based on the same evidence it may make sense to start therapy, and wait to decide about medication until after additional assessment data are integrated.

These three innovations of (a) simplifying the estimation of posterior probabilities; (b) mapping the probability onto the next clinical action; and (c) incorporating the risks, benefits, and patient preferences in the decision-making process combine to restructure the process of assessment selection and interpretation. Assimilating these ideas has led to a multistep model for evaluating potential pediatric bipolar disorder (Youngstrom, Jenkins, Jensen-Doss, & Youngstrom, 2012). This model starts with estimates of the rate of bipolar in different settings, combines that with evidence of risk factors such as familial history of bipolar disorder, and then adds test results from either the Achenbach (Achenbach & Rescorla, 2001) or more specialized mood measures. Our group has published some of the needed components, such as the "diagnostic likelihood ratios" (DLRs; Straus et al., 2011) that simplify using a probability nomogram (Youngstrom et al., 2004), and vignettes illustrating how to combine test results and risk factors for individual cases (Youngstrom & Duax, 2005; Youngstrom & Kogos Youngstrom, 2005). We have tested whether weights developed in one sample generalize to other demographically and clinically different settings (Jenkins, Youngstrom, Youngstrom, Feeny, & Findling, 2012). These methods have large effects on how practicing clinicians interpret information, making their estimates more accurate and consistent, and eliminating a tendency to overestimate the risk of bipolar disorder (Jenkins, et al., 2011).

The methods are not specific to bipolar disorder: The core ideas were developed in internal medicine and have generalized throughout other medical practices (Gray, 2004; Guyatt & Rennie, 2002). These ideas define a set of clinically relevant research projects for each new content area, sometimes only involving a shift in interpretation, but other times entailing new statistical methods or designs. Adopting these approaches redirects research to build bridges to clinical practice and orients the practitioner to look for evidence that will change their work with the patient, thus spanning the research–practice gap from both directions.

TWELVE STEPS FOR EBM, AND A COROLLARY CLINICAL RESEARCH AGENDA

The process of teaching and using the EBA model in our clinic has augmented the steps focused on a single disorder, and no doubt there will be more facets to add in the future. A dozen themes is a good start for outlining a near-future approach to evidence based assessment in psychology. Table 1 lists the steps, a brief description of clinical action, and the corresponding clinical research agenda—reinforcing the synthesis of research and practice in this hybrid approach. Figure 2 lays out a typical sequence of working through the steps, and also maps them onto the clinical decision-making thresholds from EBM and the next clinical actions in terms of assessment and treatment. All of these steps presume that the provider has adequate training and expertise to administer, score, and interpret the assessment tools accurately, or is receiving appropriate supervision while training in their use (Krishnamurthy et al., 2004).

1. Identify the Most Common Diagnoses and Presenting Problems in Our Setting

Before concentrating on the individual client, it is important to take stock of our clinical setting. What are the common presenting problems? What are the usual diagnoses? Are there any frequent clinical issues, such as abuse, custody issues, or self injury?

After making the short list of usual suspects, then it is possible to take stock of the assessment tools and practices in the clinic. Are evidence-based assessment tools available for each of the common issues? Are they routinely used? What are the gaps in coverage, where fairly common issues could be more thoroughly and accurately evaluated? Recent work on evidence-based assessment in psychology has anthologized different instruments and reviewed the evidence for the reliability and validity of each (Hunsley & Mash, 2008; Mash & Barkley, 2007). These can help guide selection. Tests with higher reliability and validity will provide greater precision and more accurate scores for high-stakes decisions about individuals (Hummel, 1999; Kelley, 1927). Factor analyses also help explicate how different scales relate to underlying constructs and to each other, allowing for more parsimony in test selection.

Pareto's "rule of the vital few" is a helpful approximation: It is not necessary to have the resources to address every possible diagnosis or contingency, and pursuing comprehensiveness would yield sharply diminishing returns. Instead, approximately 80% of cases in most clinics will have the same ~20% of the possible clinical issues. Organizing the assessment methods to address the common diagnoses will focus limited resources to address the routine referrals and presenting problems. Making the list of typical issues more explicit also helps trainees and new clinicians to consider their work context, and it turns descriptive data into institutional wisdom that can improve the assessment process through the steps described next. Tests that do not have adequate reliability or evidence of validity cannot have utility for individual decision making. The heuristic of "is this test valid, and will it help with the patient?" (Straus et al., 2011) provides a way of identifying tests that we do not want to use, and should not continue to teach, without new evidence that shows sufficient validity. Thinking about the common presenting problems and the reliable and valid tests that assess them also would help organize a "core battery" if a clinic decides to implement a standardized intake evaluation.

Clinical research agenda. One research approach to identifying the common clinical issues is to conduct clinical epidemiological studies, looking at the rates of diagnoses and key behavioral indicators across a range of service settings. Most epidemiological research focuses on the general population, regardless of treatment status. More relevant to clinicians would be the distributions of diagnoses in outpatient practice, in special education, in residential treatment, and the other settings where we provide services.

A second research project would be to map the relatively short list of families' typical presenting concerns (Garland, Lewczyk-Boxmeyer, Gabayan, & Hawley, 2004) onto the much larger list of diagnostic possibilities. If a family comes in worried about aggression, what is the shortlist of hypotheses to consider? What are the cultural factors and beliefs about causes of behavior that change how families seek help and engage with different treatments (Carpenter-Song, 2009; Yeh et al., 2005)?

2. Know the Base Rates of the Condition in Our Setting

Meehl (1954) advocated "betting the base rate" as a simple strategy to improve the accuracy of clinical

TABLE 1
Twelve Steps in Evidence-Based Assessment and Research

Assessment Step	*Rationale*	*Clinical Research Agenda*
1. Identify most common diagnoses in our setting	Planning for the typical issues helps ensure that appropriate assessment tools are available and routinely used	Clinical epidemiology; mapping presenting problem and cultural factors onto diagnoses and research labels.
2. Know base rates	Base rate is an important starting point to anchor evaluations and prioritize order of investigation	Clinical epidemiology; meta-analyses of rates across different settings and methods.
3. Evaluate relevant risk and moderating factors	Risk factors raise "index of suspicion," enough combined elevate probability into assessment or possibly treatment zones	Compare rates of risk factors in those with versus without target diagnosis; recexpress as DLRs; meta-analyses to identify moderators.
4. Synthesize broad instruments into revised probability estimates	Already widely used; know what the scores mean in terms of changing probability for common conditions	Analyses generating DLRs for popular broad coverage instruments for different clinical targets.
5. Add narrow and incremental assessments to clarify diagnoses	Often more specific measures will show better validity, or incremental value supplementing broad measures	Test incremental validity, or superiority based on cost/benefit ratio.
6. Interpret cross-informant data patterns	Pervasiveness across settings/informants reflects greater pathology. Important to understand typical patterns of disagreement, and not overinterpret common patterns.	Test diagnostic efficiency of each informant separately; test incremental value of combinations.
7. Finalize diagnoses by adding necessary intensive assessment methods	If screening and risk factors put revised probability in the "assessment zone," what are the evidence-based methods to confirm or rule out the diagnosis in question? (e.g., KSADS, neurocognitive testing . . .)	Evaluate tests in sequence in different settings to develop optimal order and weights. Develop highly specific assessments to help rule in diagnoses.
8. Complete assessment for treatment planning and goal setting	Rule out general medical conditions, other medications; Family functioning, quality of life, personality, school adjustment, comorbidities	Develop systematic ways of screening for medical conditions and medication use. Test family functioning, personality, comorbidity, socioeconomic status and other potential moderators of treatment effects.
9. Measure processes ("dashboards, quizzes and homework")	Life charts, mood and energy checkups at each visit, medication monitoring, therapy assignments, daily report cards, three-column and five-column charts . . .	Demonstrate treatment sensitivity; meditational analyses; dismantling studies examining value added.
10. Chart progress and outcome ("midterm and final exams")	Repeat assessment with main severity measures—interview and/or parent report most sensitive to treatment effects	Jacobson and Truax (1991) benchmarks and reliable change metrics; comparison of effect sizes in same trial for different methods; develop low burden methods generalizable across patients, settings, systems. If poor response, revisit diagnoses.
11. Monitor maintenance and relapse	Discuss continued life charting; review triggers, critical events and life transitions	Event history analyses (predictors of relapse, durable recovery), key predictors, recommendations about next action if roughening.
12. Solicit and integrate patient preferences	Patient beliefs and attitudes influence treatment seeking and engagement. Possible to use these preferences to adjust wait-test and test-treat thresholds or utilities.	Qualitative analyses to identify key themes, cultural factors, preferences; studies of how to quantify preferences and add to decision making.

Note: DLR = diagnostic likelihood ratio; KSADS = Kiddie Schedule for Affective Disorders and Schizophrenia.

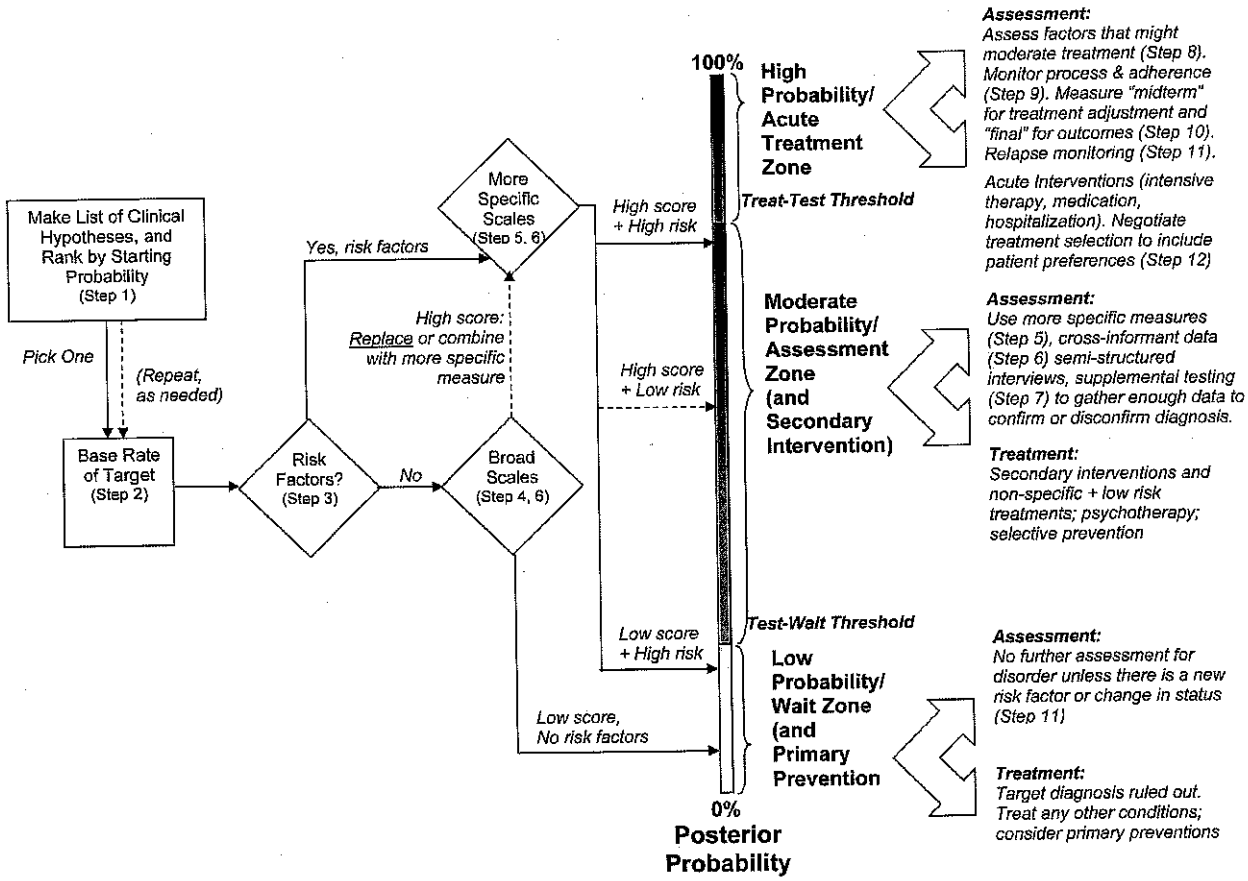


FIGURE 2 Mapping assessment results onto clinical decision making.

assessment, using the base rate as the Bayesian prior probability before adding assessment findings. When the same constellation of symptoms could be explained by an exotic or a quotidian illness, wager on the common cause. A stomachache and fever are more likely to be due to a cold virus than ebola hemhorragic fever, unless there are many other risk factors and signs that point toward the more rare explanation. The clinical epidemiological rates provide a helpful starting point for ranking the potential candidates in terms of probability before considering any case-specific information, organizing a set of potential clinical hypotheses. The prevalence of different conditions also provides a good starting estimate, taking advantage of what cognitive psychologists call the "anchoring heuristic" (Croskerry, 2003; Gigerenzer & Goldstein, 1996). Rather than interpreting case information intuitively, formally thinking about the base rates as a starting point helps increase the consistency of decision making across clinicians (Garb, 1998). Psychology has contributed both to the research about decision making and cognitive heuristics and to descriptive studies of prevalence in different settings.

Clinical research agenda. As more clinical epidemiology studies are published, then meta-analyses could describe general patterns across levels of service and identify moderating variables that change referral patterns. Studies using semistructured or structured interviews provide valuable benchmarks against which to compare local patterns. For example, if studies of urban community mental health centers find that roughly 50% of referrals meet criteria for a diagnosis of ADHD but only 20% of youths at a local center receive clinical diagnoses, or 80% for that matter, then the benchmark raises important questions about whether local assessment practices could benefit from upgrading the evidence based components.

3. Evaluate the Relevant Risk and Moderating Factors

Within the EBM framework, risk factors become data to integrate into the formal assessment process. The DLR central to the EBM method is a ratio of the diagnostic sensitivity to the false alarm rate. Put another way, the DLR compares how often the test result or risk

factor would occur in those with the diagnosis (i.e., sensitivity) versus its rate in those without the diagnosis (i.e., false alarm rate). If low birth weight was present in 3% of youths with ADHD but only 1% of those without ADHD, then the DLR attached to low birth weight would be 3.0 for ADHD. The DLR is the factor by which the odds of diagnosis change in Bayesian analysis. For clinical purposes, the conceptual status of low birth weight changes from an empirically identified "risk factor" to a variable contributing a specific weight to decision making about a particular individual case. EBM suggests that risk factors or tests producing DLRs of less than 2 are rarely worth adding to the evaluation process, whereas values around 5 are often helpful, and values greater than 10 frequently have decisive impact on an evaluation (Straus et al., 2011).

Clinical research agenda. Extensive developmental psychopathology research has focused on identifying risk and protective factors. However, these are primarily reported in terms of statistical significance and group-level effect sizes (Kraemer et al., 1999). The next step is to convert these findings into a metric amenable to idiographic assessment and decision making. The necessary statistics to generate DLRs for risk factors are simple. A chi-squared test comparing the presence or absence of the risk factor in those with or without the diagnosis is sufficient to test the validity of the risk factor (Kraemer, 1992). The next step, rarely taken in psychology to date, is to report the percentages: How common is the risk factor in those with the diagnosis versus without? Those constitute the numerator and denominator of the DLR.

4. Synthesize Broad Instruments into Revised Probability Estimates

Many clinics and practitioners use a broad assessment instrument as a standard element of their intake (e.g., Child Behavior Checklist, Behavior Assessment System for Children; Achenbach & Rescorla, 2001; Reynolds & Kamphaus, 2004). Broad instruments have a variety of strengths, including providing norm-referenced scores that compare the level of problems to what would be age- and gender-typical levels, as well as systematically assessing multiple components of emotional and behavior problems regardless of the particular referral question. This breadth prevents some cognitive heuristics that otherwise plague unstructured clinical assessments, such as concentrating only on one hypothesis, or "search satisficing" and stopping the evaluation as soon as one plausible diagnosis is identified (Croskerry, 2003; Spengler et al., 1995). The next step in an evidence-based assessment approach is to incorporate the test results and see how they raise or lower the posterior probability of the contending diagnoses. In the Bayesian EBM framework, the test score ranges have DLRs attached, and these get combined with the prior probability and risk factor DLRs to generate a revised probability estimate. It is worth noting that broad measures will not cover all possible conditions, despite their breadth. Problems that are rare in the general population may not have enough representation to generate their own "syndrome scale." This does not invalidate the use of broad measures in an EBA approach, but rather reminds us to be aware of the limits of content coverage and not unwittingly exclude clinical hypotheses outside of the scope of coverage.

Clinical research agenda. There have been a smattering of studies using Receiver Operating Characteristic (ROC) analyses to evaluate the diagnostic efficiency of broad instruments with regard to specific diagnoses such as ADHD (e.g., Chen, Faraone, Biederman, & Tsuang, 1994) and anxiety (e.g., Aschenbrand, Angelosante, & Kendall, 2005). The next step would be to calculate multilevel likelihood ratios attached to low, moderate, and high scores on the test (Guyatt & Rennie, 2002). The multilevel approach preserves more information from continuous measures, and it also is likely to be more generalizable and less sample dependent than approaches focused on picking the "optimal" cut scores (Kraemer, 1992). The approach can be simple yet still highly informative: Samples could be divided into thirds or quintiles on the Externalizing or Internalizing scale, and then the percentage of cases with the diagnosis compared to the percentage without the diagnosis in each score stratum to determine the diagnostic likelihood ratio (e.g., Youngstrom et al., 2004). As the research literature becomes more rich, then it would be possible for meta-analyses to test the generalizability of results and document moderating factors (Hasselbad & Hedges, 1995).

5. Add Narrow and Incremental Assessments to Clarify Diagnoses

At some clinics a common referral issue may not be adequately assessed by broad instruments. Pervasive developmental disorders, eating disorders, bipolar disorders, and other topics all may require the addition of more specialized measures or checklists (Mash & Hunsley, 2005). Again, a good survey of the common issues at a particular setting guides rational additions to the assessment battery. Some important issues may only be addressed by a single item or omitted entirely from broad assessment measures: The Achenbach instruments do not have scales for mania, eating

disorders, or autism, per se, for example. Psychological research has also made advances in terms of documenting incremental validity of combinations of tests (Johnston & Murray, 2003) as well as statistically testing what factors moderate the performance of tests (Cohen et al., 2003; Zumbo, 2007). The best candidates for addition to the assessment protocol will be tools that have demonstrated validity for the target diagnosis, and ideally have DLRs available so that the scores can be translated directly into a revised probability.

Clinical research agenda. Validating more narrow tests for diagnostic efficiency involves several steps. At early stages, studies performing receiver operating characteristic analyses would establish the discriminative validity of the assessment (McFall & Treat, 1999). Ideally the study design would follow the recommendations of the Standardized Reporting of Diagnostic tests guidelines (Bossuyt et al., 2003), and it would use clinically generalizable comparison groups to develop realistic estimates of performance (Youngstrom, Meyers, Youngstrom, Calabrese, & Findling, 2006a). Later steps in the research process could include comparing the ROC performance of multiple tests either in the same sample (using procedures developed by Hanley & McNeil, 1983), or meta-analytically (Hasselbad & Hedges, 1995). Logistic regression models, using diagnosis as the dependent variable, could test whether there is incremental value in combining different tests. Logistic regression also offers a flexible framework for testing potential moderators of assessment performance, such as gender, ethnicity, culture (Garb, 1998), or credibility of the informant (Youngstrom et al., 2011). EBM teaches us to ask, "Do these results apply to this patient?" (Straus et al., 2011). The psychometric tradition has developed powerful tools to answer the question of whether results generalize, versus the validity changing due to demographic or clinical characteristics (Borsboom, 2006). When appropriate samples are available, then generating multilevel likelihood ratios for the narrow instrument also would be crucial to facilitate clinical application.

6. Interpret Cross-Informant Data Patterns

A stock recommendation in clinical assessment of youths is to gather data from multiple informants, including parents, teachers, and direct observations, as well as self-report or performance measures from the youth. However, it is well-established that these different sources of information show only modest to moderate convergence, usually in the range of $r = .1$–$.4$ (Achenbach, McConaughy, & Howell, 1987). Additional data can actually degrade the quality of clinical decision making, especially when the new data have low validity for the criterion of interest or when suboptimal strategies are used to synthesize information. Context and diagnostic issue moderate the validity of data across informants (De Los Reyes & Kazdin, 2005). Self-report of attention problems, or teacher report of manic symptoms, are examples of information with validity that is significantly lower than could be gleaned by asking the same questions of other sources. Adding more tests to a battery always increases the time, cost, and complexity, but it does not always improve the output (Kraemer, 1992). Cross-informant data often add considerably to the time and expense of an assessment. The psychological assessment literature has developed to a point where we can decide when the additional assessment is worth the effort, and when it would be more efficient to forego. A related point is that we can anticipate common patterns of disagreement: Whoever initiates the referral will usually be the most worried party. Low cross-informant correlations and regression to the mean will combine so that the typical scenario often looks unimpressive in terms of agreement: If the average level of parent-reported problems has a T score of 70, the expected level of youth or teacher reported problems would be in the range of 54 to 56 (Achenbach & Rescorla, 2001; Youngstrom, Meyers, Youngstrom, Calabrese, & Findling, 2006b). Recognizing and thinking through common scenarios will help avoid misinterpreting patterns in the cross-informant data (Croskerry, 2003). When different informants have shown incremental validity, then integrating the different scores into a revised probability makes sense. Even when incremental validity for diagnostic purposes may be poor, there is still value in assessing cross-informant agreement with regard to motivation for treatment (Hunsley & Meyer, 2003).

Clinical research agenda. The ideas of cross-informant data and validity are well developed in psychological assessment and virtually unknown in the traditional EBM literature. ROC and logistic regression again provide an analytic framework for evaluating the diagnostic efficiency of each informant's perspective and testing whether there is significant incremental value added by combining different informants' perspectives.

7. Finalize Diagnoses by Adding Necessary Intensive Assessment Methods

One of the goals in sequencing the assessment steps is to try to set up a "fast and frugal" order that maximizes the information value of instruments already widely used (Gigerenzer & Goldstein, 1996) and that minimizes the additional time and expense used in the first wave of assessment for a case. Based on the initial findings,

many clinical hypotheses will be "ruled out." However, few of our assessment tools are sufficiently specific to a diagnostic issue or accurate enough to confirm a diagnosis on their own. After conducting the initial evaluation, clinicians will often find that the revised probability estimate falls in the middle "assessment zone," and additional assessment is needed to confirm or disconfirm the diagnosis. More intensive and expensive tests are justified for contending diagnoses at this stage: The prior steps have screened out low probability cases so that the more expensive methods are not being used indiscriminately (Kraemer, 1992). Reserving some procedures until there are documented risk factors and suggestive findings helps establish "medical necessity" for added assessment.

One good option would be to perform a structured or semistructured diagnostic interview, or at least the modules that are relevant to the diagnostic hypotheses for the particular case at hand. Structured interviews are more reliable and valid than unstructured clinical interviews, and they do a better job of detecting comorbid diagnoses if the full version is administered (Rettew, Lynch, Achenbach, Dumenci, & Ivanova, 2009). However, they are not a panacea: They do not have perfect validity themselves, and they can take more time than unstructured interviews (Kraemer, 1992). Also, none of them include all possible diagnoses, and any given protocol may omit at least one diagnosis that might be common at a particular setting. Until the most recent version, for example, the Kiddie Schedule for Affective Disorders and Schizophrenia (Kaufman et al., 1997) did not include a module for pervasive developmental disorders; and many interviews designed for use with youths omit bipolar disorder, eating disorders, nonsuicidal self-injury, or other conditions that have become a concern since the interviews were written or validated.

Of interest, structured approaches may be more popular with clients than with the practitioners, who cite concerns about damaging rapport as well as loss of professional autonomy as objections to routine use of more structured approaches (Suppiger et al., 2009). Structured approaches may put more administrative burden on the clinician as well as taking more time with the client (Ebesutani, Bernstein, Chorpita, & Weisz, 2012). By placing semistructured approaches at Step 7, I advocate a "combined" approach, where we consider the findings from our setting (e.g., base rates), any risk factors that might modify initial hypotheses, and the results from any checklists or rating scales *before* beginning an interview. Although Step 7 sounds late in the process, it actually falls in the first 5 to 15 min of working with an individual case. Equipped with the context and data from the prior steps, it becomes possible to decide whether to change interviews or augment with other modules or tests to cover gaps in the default interview. It also might be possible to omit modules from a semistructured interview based on revised probabilities falling below the "wait-test" threshold, although the time savings will be modest if the interview already was structured to "skip out" after a few negative responses to screening questions.

Other strategies that make sense to invoke at this stage include any other procedure that has shown incremental validity for the question of interest (Johnston & Murray, 2003) but might be too expensive or burdensome to use more generally. Essentially, this stage is a "selected or targeted" zone of assessment, analogous to selected, secondary interventions in the parlance of the International Institute of Medicine and of community mental health (Mechanic, 1989). Neurocognitive testing, daily mood charting, and soon various forms of brain imaging all might fit in this category.

Clinical research agenda. The field has been doing a good job of validating assessment strategies. The next step needed is to evaluate these tools embedded in assessment sequences tailored for distinct settings. Test consumers should not accept the developers' descriptions of test performance uncritically but rather think about how characteristics in the target and comparison group affect test performance (Bossuyt et al., 2003; Youngstrom et al., 2006a).

8. Refine Assessment for Case Formulation, Treatment Planning, and Goal Setting

There are a large number of general medical conditions and medication-related side effects that can masquerade as psychological issues. These often are measured in haphazard fashion, rather than via structured review of systems. Similarly, there are many potential treatment targets or outcome modifiers—such as personality or temperament traits, school adjustment, family functioning, parental education level—that also could be valuable to assess as part of case conceptualization and treatment selection. As we learn more about moderators of outcome, and factors that make people better matches for some treatments than others, organizing assessment to rapidly evaluate these relevant moderators will be an excellent opportunity to integrate research and practice. Assessing quality of life and functioning also is pivotal in establishing treatment goals beyond symptom reduction (Frisch, 1998).

Clinical research agenda. Much more needs to be done in terms of systematizing the evaluation of treatment moderators and also "Axis III" factors (American Psychiatric Association, 2000), such as medications and general medical conditions that have psychological

effects. Here, the initial research can move from descriptive studies to examining these variables as moderators of treatment response or predictors of optimal treatment match.

9. Measure Processes ("Quizzes, Homework, and Dashboards")

Once treatments are started, then the role of assessment changes from diagnosis to monitoring treatment progress, including mediators, process variables, and outcomes. Sometimes the intervention itself will generate products that can be used for progress checks. Examples would include behavior tracking charts, reward calendars, daily report cards, three-column and five-column charts from cognitive-behavioral therapy, and daily mood charts (Youngstrom, 2008). Many aspects of functional behavioral analysis fit well in this context, too (Vollmer & Northup, 1996). Activities completed outside of the therapy session are frequently described as "homework" to promote skill generalization. Extending the metaphor, skill assessments during sessions could be likened to "quizzes" to evaluate learning. All of these can be ratcheted toward enhancing outcome by tracking and plotting them systematically (Cone, 2001; Powsner & Tufte, 1994). Weight loss programs all measure weight repeatedly, and they have demonstrated added value of written records of food consumption and exercise on producing greater and more lasting change (Grilo, Masheb, Wilson, Gueorguieva, & White, 2011). Process measurement is much more elaborated in psychological assessment than in most of EBM, which has concentrated on diagnosis, treatment selection, and likelihood of help versus harm as the primary assessment activities (Straus et al., 2011). If the patient is failing to progress as anticipated, and especially if there are complications, we should also use this as an opportunity to reassess our case formulation and diagnoses.

Clinical research agenda. Much could be done looking at human factors that promote the uptake of some tracking methods over others. Does a smartphone application improve utilization compared to pencil and paper (e.g., Chambliss et al., 2011)? Does better utilization lead to better outcome or more durable effects? Augmentation or dismantling studies, adding or subtracting different elements of process tracking, can be embedded within other trials or routine care at clinics, helping to identify what forms of tracking are most helpful. Another promising line of work would be examining how to package these assessments into "dashboards" that provide a clear summary of progress easily interpreted by family and therapist alike (Few, 2006; Powsner & Tufte, 1994).

10. Chart Progress and Outcome ("Midterm and Final Exams")

Continuing with the education metaphor, outcome evaluation can be cast as the "final exam," measuring the amount of change over the course of treatment. There are several operational definitions of outcome, including loss of diagnosis, percentage reduction of symptoms on a severity measure, or more complex definitions of "clinically significant change" that combine information about the precision of the measure—such as the "reliable change index"—with comparisons to normative benchmarks based on distributions in clinical and nonclinical samples (Jacobson & Truax, 1991). All of these involve more lengthy and comprehensive evaluation than the "process" measures just described, and so these panels of assessment methods are used more episodically. In clinical practice, outcome evaluation is more likely to be informal, based on the view that it is obvious when people are improving, and the belief that clients and payers will not accept the additional assessment involved (Suppiger et al., 2009). Contrary to expectation, clients are likely to view thorough assessments positively (Suppiger et al., 2009), and payers are more likely to reimburse assessments that are clearly linked to treatment (Cashel, 2002). Services databases consistently show modest rates of improvement and great heterogeneity in outcomes for treatment as usual, with some cases improving markedly, and others actually deteriorating. Meehl and others have argued that the slow progress in psychological treatment is due in large part to our failure to measure outcomes and get corrective feedback about when our interventions help, are inert, or even harm (Christensen & Jacobson, 1994; Meehl, 1973).

Research about patterns of treatment response also indicates potential value in having a scheduled "midterm," where more intensive evaluation is done to quantify early response to treatment. Early response to intervention, both psychotherapy and pharmacological (Curry et al., 2011), often predicts long-term response (Howard, Moras, Brill, Martinovich, & Lutz, 1996). If a person does not show improvement over the first 4 to 8 weeks or sessions, then it makes sense to either augment or change the modality of treatment (Lambert, Hansen, & Finch, 2001). Careful assessment of early response is also crucial to monitoring side effects and potential treatment-emergent changes in mood or behavior that should trigger alterations in the treatment plan (Joseph, Youngstrom, & Soares, 2009). Outcome evaluation is another area where psychological assessment has developed more sophisticated models for evaluating individual change compared to the metrics commonly used in EBM. Number needed to treat (the number of people who would need exposure to the treatment for

one more case to have a good outcome), number needed to harm (the number of people who would need exposure to the treatment for one more case to experience harmful side effects or iatrogenic outcomes), and similar indices are all measures of probabilistic efficacy based on groups of cases and dichotomous outcomes (Guyatt & Rennie, 2002). Psychological assessment offers much in terms of benchmarking against typical ranges of functioning, looking at change on continuous measures, and considering the precision of measurement when evaluating individual outcomes.

Clinical research agenda. There are a variety of methods worth investigating, including trials examining whether the addition of assessment at the "midterm" or end of acute treatment changes engagement, adherence, and acute or long-term outcomes (e.g., Ogles, Melendez, Davis, & Lunnen, 2001). A second line of work could optimize instruments for outcome evaluation by demonstrating sensitivity to treatment effects, developing shorter versions that retain sufficient precision to guide individual treatment decisions, and establishing meaningful benchmarks for "clinically significant change" approaches.

11. Monitor Maintenance and Relapses

Many disorders of childhood and adolescence carry a high risk of relapse, such as mood disorders; others are associated with an elevated risk of developing later pathology, perhaps as forms of heterotypic continuity. Anxiety often augurs later depression (Mineka, Watson, & Clark, 1998), and ADHD often presages substance issues or conduct problems (Taurines et al., 2010). More could be done in terms of educating families around signs of relapse or cues of early onset of later problems. Creative work is being done with mood disorders, helping patients identify signs of "roughening" and changes in energy or behavior that might offer early warning of relapse (Sachs, 2004), and then planning ahead of time for strategies that can help restabilize mood or promote earlier intervention to minimize the effects of recurrence. Given what we know about the epidemiology of mental health problems and developmental changes through adolescence and early adulthood, a combination of general screening and brief, targeted evaluations of warning signs could accomplish much good. This aspect of assessment has not received much attention from either the EBM or psychological assessment traditions yet, and represents a major growth area.

Clinical research agenda. It would be intriguing to evaluate how customized assessment strategies might predict shorter lag to seeking treatment, increased utilization of prevention or early intervention services, or diversion from more acute and tertiary treatments. Similarly, it would be important to know whether brief, broad coverage measures might have a role in primary care or other settings as predictors of relapse or progression in youths who have previously benefitted from treatment. Advances in technology make a variety of "smart" applications feasible as methods for monitoring behavior for cues of relapse.

12. Solicit and Integrate Patient Preferences

The placement of the wait-test and treat-test thresholds is flexible in EBM (Straus et al., 2011) (see also Figure 2). Their location is supposed to be guided by the costs and benefits attached to the diagnosis or treatment, as well as patient preferences. For dichotomous outcomes, such as recovery or remission, there is a developed framework combining the number needed to treat with the number needed to harm, yielding a Likelihood of Help versus Harm that can be further adjusted based on patient preferences (Straus et al., 2011). There are other formal mathematical approaches to synthesizing costs, benefits, and assessment parameters to optimize decision thresholds (Kraemer, 1992; Swets, Dawes, & Monahan, 2000), too. The EBM approach is attractive because it is simple enough that it could be done in session with families, potentially working through several "what if..." scenarios together to help explore a range of options and guide consensual decisions.

There is a rich layer of additional information that could be added here, using surveys and interviews to solicit beliefs about causes of emotional and behavioral problems, differences in what is perceived as problematic, and attitudes toward help-seeking and different services. Beliefs about medication and therapy have great influence over treatment seeking and engagement (Yeh et al., 2005). The effects of culture on decisions to seek or continue treatment are likely to be as big or bigger than culture's moderating effects on the accuracy of assessments or intervention efficacy. This aspect of assessment is one of the most promising places to combine psychological assessment's sophistication about measuring beliefs, attitudes, and preferences with the mathematical framework and decision aids offered by EBM.

Clinical research agenda. Qualitative methods as well as quantitative interviews and surveys have much to add in terms of knowledge about patient preferences. There also is a great deal that could be done integrating preferences into the decision-making framework, adjusting the test score thresholds for screening programs at a policy level (Swets et al., 2000) or negotiating personalized decision making with individual cases (Straus et al.,

2011). The algorithms have been available for decades, but it is only recently that technology has made it convenient for families and practitioners to use the tools. Recent developments understanding the role of culture in service selection, stigma, and attitudes to treatment also provides more rich inputs into the decision-making process (Hinshaw & Cicchetti, 2000; Yeh et al., 2005). Although last in the "steps" listed here, understanding patient attitudes is something we could profitably weave through the entire assessment process.

DISCUSSION

When it convened more than a dozen years ago, the Psychological Assessment Work Group of the American Psychological Association concluded there was surprisingly little published data to document the value of conventional psychological assessment in terms of better outcomes (Eisman et al., 1998; Meyer et al., 1998). The situation has improved only modestly in subsequent years (Hunsley & Mash, 2007). Our failure to measure things that matter to families and for treatment still contributes to the slow progress of our interventions (Meehl, 1973; Nelson-Gray, 2003).

EBM lacks the psychometric sophistication that has characterized the best traditions of psychological assessment. Psychological assessment has developed a wide range of instruments, and psychometric models could provide sophisticated techniques for honing the analytical underpinnings of EBM (Borsboom, 2008). What EBM offers, though, is a pragmatic focus on understanding and helping the individual case. EBM ties assessment to clinical decision making with a directness and clarity that has been missing in much of psychological assessment. Integration is possible, keeping the psychometric and conceptual strengths of psychological assessment but incorporating them into the decision-making framework articulated in EBM. The fit is not seamless, but it is patient centered, clinically relevant, and compelling. Some of the looser connections will be promising areas of investigation in their own right. EBM has historically emphasized dichotomous outcomes (e.g., recovery, death), whereas psychology has focused more on continuous measures. It is possible to convert dimensional effect sizes, such as Cohen's *d* or a correlation coefficient, into other effect sizes such as risk ratios (Hasselbad & Hedges, 1995), making it possible to reexpress outcomes in metrics that fit within the EBM decision-making framework, but it also would be intriguing to develop parallel approaches that capitalize on the greater information intrinsic to continuous measures.

Exploring the potential for synthesis reorganized my approach to assessment research, teaching, and supervision. Viewing assessment through an EBM tinted lens defines a set of clinical research topics that comprise a thematic program of investigation. The research designs and statistical methods are readily available and not complex. Adopting these methods need not add to the expense of the assessment process: Better decisions can be made by using the same tools but interpreting them differently. For example, we have found that there can be pronounced changes in clinical decisions about vignettes, with increased accuracy and consistency, and an elimination of a tendency to overdiagnose bipolar disorder, based on identical assessment data combined with brief training in the probability nomogram as a way of interpreting scores (Jenkins et al., 2011). The value of these methods is not limited to bipolar disorder, any more than it would be limited to any single area within medicine (Guyatt & Rennie, 2002). The hybridization of psychological assessment with EBM ideas produces ideas with vigor and clinical relevance to rejuvenate assessment and ultimately improve outcomes for families (Bauer, 2007).

REFERENCES

Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: Implication of cross-informant correlations for situational specificity. *Psychological Bulletin, 101*, 213–232. doi:10.1037/0033-2909.101.2.213

Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for the ASEBA School-Age Forms & Profiles*. Burlington: University of Vermont.

Algorta, G. P., Youngstrom, E. A., Phelps, J., Jenkins, M. M., Youngstrom, J. K., & Findling, R. L. (2012). An inexpensive family index of risk for mood issues improves identification of pediatric bipolar disorder. *Psychological Assessment*. Advance online publication. doi:10.1037/a0029225

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: Author.

American Psychological Association. (2005). Policy statement on evidence-based practice in psychology. Retrieved from http://www.apa.org/practice/resources/evidence/evidence-based-statement.pdf

Aschenbrand, S. G., Angelosante, A. G., & Kendall, P. C. (2005). Discriminant validity and clinical utility of the CBCL with anxiety-disordered youth. *Journal of Clinical Child and Adolescent Psychology, 34*, 735–746. doi:10.1207/s15374424jccp3404_15

Bauer, R. M. (2007). Evidence-based practice in psychology: implications for research and research training. *Journal of Clinical Psychology, 63*, 685–694. doi:10.1002/jclp.20374

Bayes, T., & Price, R. (1763). An essay towards solving a problem in the doctrine of chance. By the late Rev. Mr. Bayes, communicated by Mr. Price, in a letter to John Canton, M. A. and F. R. S. *Philosophical Transactions of the Royal Society of London, 53*, 370–418. doi:10.1098/rstl.1763.0053

Belter, R. W., & Piotrowski, C. (2001). Current status of doctoral-level training in psychological testing. *Journal of Clinical Psychology, 57*, 717–726.

Borsboom, D. (2006). The attack of the psychometricians. *Psychometrika, 71*, 425–440. doi:10.1007/s11336-006-1447-6

Borsboom, D. (2008). Psychometric perspectives on diagnostic systems. *Journal of Clinical Psychology, 64*, 1089–1108. doi:10.1002/jclp.20503

Bossuyt, P. M., Reitsma, J. B., Bruns, D. E., Gatsonis, C. A., Glasziou, P. P., Irwig, L. M., ... de Vet, H. C. W. (2003). Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD initiative. *British Medical Journal, 326*, 41–44. doi:10.1136/bmj.326.7379.41

Camara, W., Nathan, J., & Puente, A. (1998). *Psychological test usage in professional psychology: Report of the APA practice and science directorates* (p. 51). Washington, DC: American Psychological Association.

Carpenter-Song, E. (2009). Caught in the psychiatric net: meanings and experiences of ADHD, pediatric bipolar disorder and mental health treatment among a diverse group of families in the United States. *Culture, Medicine and Psychiatry, 33*, 61–85. doi:10.1007/s11013-008-9120-4

Cashel, M. L. (2002). Child and adolescent psychological assessment: Current clinical practices and the impact of managed care. *Professional Psychology: Research and Practice, 33*, 446–453. doi:10.1037//0735-7028.33.5.446

Chambliss, H. O., Huber, R. C., Finley, C. E., McDoniel, S. O., Kitzman-Ulrich, H., & Wilkinson, W. J. (2011). Computerized self-monitoring and technology-assisted feedback for weight loss with and without an enhanced behavioral component. *Patient Education and Counseling, 85*, 375–382. doi:10.1016/j.pec.2010.12.024

Chen, W. J., Faraone, S. V., Biederman, J., & Tsuang, M. T. (1994). Diagnostic accuracy of the Child Behavior Checklist scales for attention-deficit hyperactivity disorder: A receiver-operating characteristic analysis. *Journal of Consulting and Clinical Psychology, 62*, 1017–1025. doi:10.1037/0022-006X.62.5.1017

Childs, R. A., & Eyde, L. D. (2002). Assessment training in clinical psychology doctoral programs: what should we teach? What do we teach? *Journal of Personality Assessment, 78*, 130–144. doi:10.1207/S15327752JPA7801_08

Christensen, A., & Jacobson, N. S. (1994). Who (or what) can do psychotherapy: The status and challenge of nonprofessional therapies. *Psychological Science, 5*, 8–14. doi:10.1111/j.1467-9280.1994.tb00606.x

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Erlbaum.

Cone, J. D. (2001). *Evaluating outcomes: Empirical tools for effective practice*. Washington, DC: American Psychological Association.

Correll, C. U. (2008). Antipsychotic use in children and adolescents: Minimizing adverse effects to maximize outcomes. *Journal of the American Academy of Child & Adolescent Psychiatry, 47*, 9–20. doi:10.1097/chi.0b013e31815b5cb1

Croskerry, P. (2003). The importance of cognitive errors in diagnosis and strategies to minimize them. *Academic Medicine, 78*, 775–780. doi:10.1097/00001888-200308000-00003

Curry, J., Silva, S., Rohde, P., Ginsburg, G., Kratochvil, C., Simons, A., ... March, J. (2011). Recovery and recurrence following treatment for adolescent major depression. *Archives of General Psychiatry, 68*, 263–269. doi:10.1001/archgenpsychiatry.2010.150

Dawes, R. M., Faust, D., & Meehl, P. E. (1989). Clinical versus actuarial judgment. *Science, 243*, 1668–1674. doi:10.1126/science.2648573

De Los Reyes, A., & Kazdin, A. E. (2005). Informant discrepancies in the assessment of childhood psychopathology: A critical review, theoretical framework, and recommendations for further study. *Psychological Bulletin, 131*, 483–509. doi:10.1037/0033-2909.131.4.483

Ebesutani, C., Bernstein, A., Chorpita, B. F., & Weisz, J. R. (2012). A transportable assessment protocol for prescribing youth psychosocial treatments in real-world settings: Reducing assessment burden via self-report scales. *Psychological Assessment, 24*, 141–155. doi:10.1037/a0025176

Eisman, E. J., Dies, R. R., Finn, S. E., Eyde, L. D., Kay, G. G., Kubiszyn, T. W., ... Moreland, K. L. (1998). *Problems and limitations in the use of psychological assessment in contemporary health care delivery: Report of the Board of Professional Affairs Psychological Assessment Workgroup, Part II* (p. 22). Washington, DC: American Psychological Association.

Few, S. (2006). *Information dashboard design: The effective visual communication of data*. Cambridge, MA: O'Reilly Press.

Fletcher, J. M., Francis, D. J., Morris, R. D., & Lyon, G. R. (2005). Evidence-based assessment of learning disabilities in children and adolescents. *Journal of Clinical Child and Adolescent Psychology, 34*, 506–522. doi:10.1207/s15374424jccp3403_7

Frazier, T. W., & Youngstrom, E. A. (2006). Evidence-based assessment of attention-deficit/hyperactivity disorder: Using multiple sources of information. *Journal of the American Academy of Child & Adolescent Psychiatry, 45*, 614–620. doi:10.1097/01.chi.0000196597.09103.25

Frisch, M. B. (1998). Quality of life therapy and assessment in health care. *Clinical Psychology: Science and Practice, 5*, 19–40.

Garb, H. N. (1998). *Studying the clinician: Judgment research and psychological assessment*. Washington, DC: American Psychological Association.

Garland, A. F., Lewczyk-Boxmeyer, C. M., Gabayan, E. N., & Hawley, K. M. (2004). Multiple stakeholder agreement on desired outcomes for adolescents' mental health services. *Psychiatric Services, 55*, 671–676. doi:10.1176/appi.ps.55.6.671

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review, 103*, 650–669. doi:10.1037/0033-295X.103.4.650

Glutting, J. J., Youngstrom, E. A., Ward, T., Ward, S., & Hale, R. (1997). Incremental efficacy of WISC–III factor scores in predicting achievement: What do they tell us? *Psychological Assessment, 9*, 295–301. doi:10.1037/1040-3590.9.3.295

Gray, G. E. (2004). *Evidence-based psychiatry*. Washington, DC: American Psychiatric Publishing.

Grilo, C. M., Masheb, R. M., Wilson, G. T., Gueorguieva, R., & White, M. A. (2011). Cognitive-behavioral therapy, behavioral weight loss, and sequential treatment for obese patients with binge-eating disorder: A randomized controlled trial. *Journal of Consulting & Clinical Psychology, 79*, 675–685. doi:10.1037/a0025049

Guyatt, G. H., & Rennie, D. (Eds.). (2002). *Users' guides to the medical literature*. Chicago, IL: AMA Press.

Hanley, J. A., & McNeil, B. J. (1983). A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology, 148*, 839–843.

Harkness, A. R., & Lilienfeld, S. O. (1997). Individual differences science for treatment planning: Personality traits. *Psychological Assessment, 9*, 349–360.

Hasselbad, V., & Hedges, L. V. (1995). Meta-analysis of screening and diagnostic tests. *Psychological Bulletin, 117*, 167–178. doi:10.1037/0033-2909.117.1.167

Hayes, S. C., Nelson, R. O., & Jarrett, R. B. (1987). The treatment utility of assessment: A functional approach to evaluating assessment quality. *American Psychologist, 42*, 963–974.

Hinshaw, S. P., & Cicchetti, D. (2000). Stigma and mental disorder: Conceptions of illness, public attitudes, personal disclosure, and social policy. *Development & Psychopathology, 12*, 555–598. doi:10.1017/S0954579400004028

Hodgins, S., Faucher, B., Zarac, A., & Ellenbogen, M. (2002). Children of parents with bipolar disorder. A population at high risk for major affective disorders. *Child & Adolescent Psychiatric Clinics of North America, 11*, 533–553.

Howard, K. I., Moras, K., Brill, P. L., Martinovich, Z., & Lutz, W. (1996). Evaluation of psychotherapy: Efficacy, effectiveness, and

patient progress. *American Psychologist, 51*, 1059–1064. doi:10.1037/0003-066X.51.10.1059

Hummel, T. J. (1999). The usefulness of tests in clinical decisions. In J. W. Lichtenberg & R. K. Goodyear (Eds.), *Scientist-practitioner perspectives on test interpretation* (pp. 59–112). Boston, MA: Allyn and Bacon.

Hunsley, J., & Mash, E. J. (2005). Introduction to the special section on developing guidelines for the evidence-based assessment (EBA) of adult disorders. *Psychological Assessment, 17*, 251–255. doi:10.1037/1040-3590.17.3.251

Hunsley, J., & Mash, E. J. (2007). Evidence-based assessment. *Annual Review of Clinical Psychology, 3*, 29–51. doi:10.1146/annurev.clinpsy.3.022806.091419

Hunsley, J., & Mash, E. J. (Eds.). (2008). *A guide to assessments that work*. New York, NY: Oxford University Press.

Hunsley, J., & Meyer, G. J. (2003). The incremental validity of psychological testing and assessment: Conceptual, methodological, and statistical issues. *Psychological Assessment, 15*, 446–455.

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59*, 12–19. doi:10.1037/0022-006X.59.1.12

Jaeschke, R., Guyatt, G. H., & Sackett, D. L. (1994). Users' guides to the medical literature: III. How to use an article about a diagnostic test: B: What are the results and will they help me in caring for my patients? *Journal of the American Medical Association, 271*, 703–707.

Jenkins, M. M., Youngstrom, E. A., Washburn, J. J., & Youngstrom, J. K. (2011). Evidence-based strategies improve assessment of pediatric bipolar disorder by community practitioners. *Professional Psychology: Research and Practice, 42*, 121–129. doi:10.1037/a0022506

Jenkins, M. M., Youngstrom, E. A., Youngstrom, J. K., Feeny, N. C., & Findling, R. L. (2012). Generalizability of evidence-based assessment recommendations for pediatric bipolar disorder. *Psychological Assessment, 24*, 269–281. doi:10.1037/a0025775

Johnston, C., & Murray, C. (2003). Incremental validity in the psychological assessment of children and adolescents. *Psychological Assessment, 15*, 496–507.

Joseph, M., Youngstrom, E. A., & Soares, J. C. (2009). Antidepressant-coincident mania in children and adolescents treated with selective serotonin reuptake inhibitors. *Future Neurology, 4*, 87–102. doi:10.2217/14796708.4.1.87

Kaufman, J., Birmaher, B., Brent, D., Rao, U., Flynn, C., Moreci, P., ... Ryan, N. (1997). Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime version (K-SADS-PL): Initial reliability and validity data. *Journal of the American Academy of Child & Adolescent Psychiatry, 36*, 980–988. doi:10.1097/00004583-199707000-00021

Kelley, T. L. (1927). *Interpretation of educational measurements*. Yonkers, NY: World Books.

Kovacs, M. (1992). *Children's Depression Inventory Manual*. North Tonawanda, NY: Multi-Health Systems.

Kraemer, H. C. (1992). *Evaluating medical tests: Objective and quantitative guidelines*. Newbury Park, CA: Sage.

Kraemer, H. C., Kazdin, A. E., Offord, D. R., Kessler, R. C., Jensen, P. S., & Kupfer, D. J. (1999). Measuring the potency of risk factors for clinical or policy significance. *Psychological Methods, 4*, 257–271.

Krishnamurthy, R., VandeCreek, L., Kaslow, N. J., Tazeau, Y. N., Miville, M. L., Kerns, R., ... Benton, S. A. (2004). Achieving competency in psychological assessment: directions for education and training. *Journal of Clinical Psychology, 60*, 725–739. doi:10.1002/jclp.20010

Lambert, M. J., Hansen, N. B., & Finch, A. E. (2001). Patient-focused research: using patient outcome data to enhance treatment effects. *Journal of Consulting & Clinical Psychology, 69*, 159–172. doi:10.1037/0022-006X.69.2.159

MacKinnon, D. P., Fairchild, A. J., & Fritz, M. S. (2007). Mediation analysis. *Annual Review of Psychology, 58*, 593–614. doi:10.1146/annurev.psych.58.110405.085542

Mash, E. J., & Barkley, R. A. (Eds.). (2007). *Assessment in children and adolescents*. New York, NY: Guilford.

Mash, E. J., & Hunsley, J. (2005). Evidence-based assessment of child and adolescent disorders: Issues and challenges. *Journal of Clinical Child and Adolescent Psychology, 34*, 362–379. doi:10.1207/s15374424jccp3403_1

McFall, R. (1991). Manifesto for a science of clinical psychology. *The Clinical Psychologist, 44*, 75–88.

McFall, R. M., & Treat, T. A. (1999). Quantifying the information value of clinical assessment with signal detection theory. *Annual Review of Psychology, 50*, 215–241. doi:10.1146/annurev.psych.50.1.215

Mechanic, D. (1989). *Mental health and social policy*. Englewood Cliffs, NJ: Prentice-Hall.

Meehl, P. E. (1954). *Clinical versus statistical prediction: A theoretical analysis and a review of the evidence*. Minneapolis: University of Minnesota Press.

Meehl, P. (1973). Why I do not attend case conferences. In P. Meehl (Ed.), *Psychodiagnosis: Selected papers* (pp. 225–302). New York, NY: Norton.

Meehl, P. E. (1997). Credentialed persons, credentialed knowledge. *Clinical Psychology: Science and Practice, 4*, 91–98. doi:10.1111/j.1468-2850.1997.tb00103.x

Meehl, P. E., & Rosen, A. (1955). Antecedent probability and the efficiency of psychometric signs, patterns, or cutting scores. *Psychological Bulletin, 55*, 194–216.

Merenda, P. F. (2007a). Psychometrics and psychometricians in the 20th and 21st centuries: How it was in the 20th century and how it is now. *Perceptual & Motor Skills, 104*, 3–20. doi:10.2466/pms.104.1.3-20

Merenda, P. F. (2007b). Update on the decline in the education and training in psychological measurement and assessment. *Psychological Reports, 101*, 153–155. doi:10.2466/pr0.101.1.153-155

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist, 50*, 741–749. doi:10.1037/0003-066X.50.9.741

Meyer, G. J., Finn, S. E., Eyde, L. D., Kay, G. G., Kubiszyn, T. W., Moreland, K. L., ... Dies, R. R. (1998). *Benefits and costs of psychological assessment in health care delivery: Report of the Board of Professional Affairs Psychological Assessment Workgroup, Part I* (p. 90). Washington, DC: American Psychological Association.

Meyer, G. J., & Handler, L. (1997). The ability of the Rorschach to predict subsequent outcome: A meta-analysis of the Rorschach Prognostic Rating Scale. *Journal of Personality Assessment, 69*, 1–38. doi:10.1207/s15327752jpa6901_1

Mineka, S., Watson, D., & Clark, L. A. (1998). Comorbidity of anxiety and unipolar mood disorders. *Annual Review of Psychology, 49*, 377–412. doi:10.1146/annurev.psych.49.1.377

Nelson-Gray, R. O. (2003). Treatment utility of psychological assessment. *Psychological Assessment, 15*, 521–531.

Ogles, B. M., Melendez, G., Davis, D. C., & Lunnen, K. M. (2001). The Ohio Scales: Practical outcome assessment. *Journal of Child & Family Studies, 10*, 199–212.

Pepe, M. S. (2003). *The statistical evaluation of medical tests for classification and prediction*. New York, NY: Wiley.

Piotrowski, C. (1999). Assessment practices in the era of managed care: Current status and future directions. *Journal of Clinical Psychology, 55*, 787–796.

Powsner, S. M., & Tufte, E. R. (1994). Graphical summary of patient status. *The Lancet, 344*, 368–389. doi:10.1016/S0140-6736(94)91406-0

Ravens-Sieberer, U., & Bullinger, M. (1998). Assessing health-related quality of life in chronically ill children with the German KINDL: First psychometric and content analytic results. *Quality of Life Research, 7*, 399–407.

Rettew, D. C., Lynch, A. D., Achenbach, T. M., Dumenci, L., & Ivanova, M. Y. (2009). Meta-analyses of agreement between diagnoses made from clinical evaluations and standardized diagnostic interviews. *International Journal of Methods in Psychiatric Research, 18*, 169–184. doi:10.1002/mpr.289

Reynolds, C. R., & Kamphaus, R. (2004). *BASC–2 Behavior Assessment System for Children*. Circle Pines, MN: American Guidance Service.

Sachs, G. S. (2004). Strategies for improving treatment of bipolar disorder: Integration of measurement and management. *Acta Psychiatrica Scandinavica*, 7–17. doi:10.1111/j.1600-0447.2004.00409.x

Sattler, J. M. (2002). *Assessment of children: Behavioral and Clinical Applications* (4th ed.). La Mesa, CA: Publisher Inc.

Spengler, P. M., Strohmer, D. C., Dixon, D. N., & Shivy, V. A. (1995). A scientist–practitioner model of psychological assessment: Implications for training, practice and research. *The Counseling Psychologist, 23*, 506–534. doi:10.1177/0011000095233009

Spring, B. (2007). Evidence-based practice in clinical psychology: What it is, why it matters; What you need to know. *Journal of Clinical Psychology, 63*, 611–631.

Stedman, J. M., Hatch, J. P., & Schoenfeld, L. S. (2001). The current status of psychological assessment training in graduate and professional schools. *Journal of Personality Assessment, 77*, 398–407. doi:10.1207/S15327752JPA7703_02

Stedman, J. M., Hatch, J. P., Schoenfeld, L. S., & Keilin, W. G. (2005). The structure of internship training: Current patterns and implications for the future of clinical and counseling psychologists. *Professional Psychology: Research and Practice, 36*, 3–8. doi:10.1037/0735-7028.36.1.3

Straus, S. E., Glasziou, P., Richardson, W. S., & Haynes, R. B. (2011). *Evidence-based medicine: How to practice and teach EBM* (4th ed.). New York, NY: Churchill Livingstone.

Suppiger, A., In-Albon, T., Hendriksen, S., Hermann, E., Margraf, J., & Schneider, S. (2009). Acceptance of structured diagnostic interviews for mental disorders in clinical practice and research settings. *Behavior Therapy, 40*, 272–279. doi:S0005-7894(08)00088-9 [pii]

Swets, J. A., Dawes, R. M., & Monahan, J. (2000). Psychological science can improve diagnostic decisions. *Psychological Science in the Public Interest, 1*, 1–26. doi:10.1111/1529-1006.001

Taurines, R., Schmitt, J., Renner, T., Conner, A. C., Warnke, A., & Romanos, M. (2010). Developmental comorbidity in attention-deficit/hyperactivity disorder. *Attention Deficit and Hyperactivity Disorders, 2*, 267–289. doi:10.1007/s12402-010-0040-0

Vollmer, T. R., & Northup, J. (1996). Some implications of functional analysis for school psychology. *School Psychology Quarterly, 11*, 76–92.

Wagner, K. D., Hirschfeld, R., Findling, R. L., Emslie, G. J., Gracious, B., & Reed, M. (2006). Validation of the mood disorder questionnaire for bipolar disorders in adolescents. *Journal of Clinical Psychiatry, 67*, 827–830. doi:10.4088/JCP.v67n0518

Watkins, M. W. (2000). Cognitive profile analysis: A shared professional myth. *School Psychology Quarterly, 15*, 465–479. doi:10.1037/h0088802

Wood, J. M., Nezworski, M. T., & Stejskal, W. J. (1996). The comprehensive system for the Rorschach: A critical examination. *Psychological Science, 7*, 3–10. doi:10.1111/j.1467-9280.1996.tb00658.x

Yeh, M., Hough, R. L., Fakhry, F., McCabe, K. M., Lau, A. S., & Garland, A. F. (2005). Why bother with beliefs? Examining relationships between race/ethnicity, parental beliefs about causes of child problems, and mental health service use. *Journal Consulting and Clinical Psychology, 73*, 800–807. doi:10.1037/0022-006X.73.5.800

Youngstrom, E. A. (2007). Pediatric bipolar disorder. In E. J. Mash & R. A. Barkley (Eds.), *Assessment of childhood disorders* (4th ed., pp. 253–304). New York, NY: Guilford.

Youngstrom, E. A. (2008). Evidence-based strategies for the assessment of developmental psychopathology: Measuring prediction, prescription, and process. In D. J. Miklowitz, W. E. Craighead, & L. Craighead (Eds.), *Developmental psychopathology* (pp. 34–77). New York, NY: Wiley.

Youngstrom, E. A., & Duax, J. (2005). Evidence based assessment of pediatric bipolar disorder, Part 1: Base rate and family history. *Journal of the American Academy of Child & Adolescent Psychiatry, 44*, 712–717. doi:10.1097/01.chi.0000162581.87710.bd

Youngstrom, E. A., Findling, R. L., Calabrese, J. R., Gracious, B. L., Demeter, C., DelPorto Bedoya, D., & Price, M. (2004). Comparing the diagnostic accuracy of six potential screening instruments for bipolar disorder in youths aged 5 to 17 years. *Journal of the American Academy of Child & Adolescent Psychiatry, 43*, 847–858. doi:10.1097/01.chi.0000125091.35109.1e

Youngstrom, E. A., Findling, R. L., Danielson, C. K., & Calabrese, J. R. (2001). Discriminative validity of parent report of hypomanic and depressive symptoms on the General Behavior Inventory. *Psychological Assessment, 13*, 267–276.

Youngstrom, E. A., Freeman, A. J., & Jenkins, M. M. (2009). The assessment of children and adolescents with bipolar disorder. *Child and Adolescent Psychiatric Clinics of North America, 18*, 353–390. doi:10.1016/j.chc.2008.12.002

Youngstrom, E. A., Jenkins, M. M., Jensen-Doss, A., & Youngstrom, J. K. (2012). Evidence-based assessment strategies for pediatric bipolar disorder. *Israel Journal of Psychiatry & Related Sciences, 49*, 15–27.

Youngstrom, E. A., & Kogos Youngstrom, J. (2005). Evidence-based assessment of pediatric bipolar disorder, Part 2: Incorporating information from behavior checklists. *Journal of the American Academy of Child & Adolescent Psychiatry, 44*, 823–828. doi:10.1097/01.chi.0000164589.10200.a4

Youngstrom, E. A., Meyers, O. I., Youngstrom, J. K., Calabrese, J. R., & Findling, R. L. (2006a). Comparing the effects of sampling designs on the diagnostic accuracy of eight promising screening algorithms for pediatric bipolar disorder. *Biological Psychiatry, 60*, 1013–1019. doi:10.1016/j.biopsych.2006.06.023

Youngstrom, E. A., Meyers, O., Youngstrom, J. K., Calabrese, J. R., & Findling, R. L. (2006b). Diagnostic and measurement issues in the assessment of pediatric bipolar disorder: Implications for understanding mood disorder across the life cycle. *Development and Psychopathology, 18*, 989–1021. doi:10.1017/S0954579406060494

Youngstrom, E. A., Youngstrom, J. K., Freeman, A. J., De Los Reyes, A., Feeny, N. C., & Findling, R. L. (2011). Informants are not all equal: predictors and correlates of clinician judgments about caregiver and youth credibility. *Journal of Child and Adolescent Psychopharmacology, 21*, 407–415. doi:10.1089/cap.2011.0032

Zumbo, B. D. (2007). Three generations of DIF analyses: Considering where it has been, where it is now, and where it is going. *Language Assessment Quarterly, 4*, 223–233.

APPENDIX

Case Example

Referral Question: Tandi is a 10-year-old girl living with her biological parents and older sister who is coming for an outpatient evaluation because her mother is concerned about her increasing "mood swings." Tandi is in

regular education at a public school, taking accelerated classes. Her mother describes her as having been outgoing and cheerful as a child, but recently seems to have become more quiet, irritable, and crabby, sometimes snapping at her family, and recently slamming doors and throwing things. According to her mom, the paternal aunt has been diagnosed with bipolar disorder, and her mom has heard that this runs in families. She wants to know if Tandi has bipolar disorder.

Steps 1 & 2. Identify the Most Common Diagnoses and Presenting Problems in Our Setting, and Know the Base Rates of the Condition in Our Setting. The clinic where Tandi's family presented uses an electronic medical record, so it is possible to produce a report listing the most frequent diagnoses. The most common diagnosis is adjustment disorder (~60% of cases), followed by attention deficit/hyperactivity disorder (ADHD; 40%), oppositional defiant disorder (ODD; 35%), and major depressive disorder (30%, but lower in younger children and higher postpubertally). Posttraumatic stress disorder (PTSD), conduct disorder, and bipolar spectrum disorders are all diagnosed in roughly 10% of cases. The clinician has compared these rates with published rates from other outpatient settings and knows that the rank order seems plausible compared to external benchmarks. The somewhat higher rates of externalizing problems and lower rates of anxiety disorders reflect logical patterns in local referral sources. Based on this, bipolar disorder is worth assessing to address the referral question, but it is not a leading candidate. The clinic has stocked rating scales and assessment tools for all of the diagnoses that occur in 10% or more of cases, so the resources are available to explore bipolar disorder further if warranted.

Step 3. Evaluate the Relevant Risk and Moderating Factors *(Also Illustrating the Use of the Probability Nomogram).* Family history of bipolar disorder is a well-established risk factor, based on decades of research and multiple reviews. A clear diagnosis of bipolar in a first degree relative is associated with a diagnostic likelihood ratio (DLR) of 5.0, indicating a fivefold increase in the odds of the youth having a bipolar disorder (Youngstrom & Duax, 2005). A second-degree relative, such as the paternal aunt, will share on average half as many genes with the person being assessed, and thus confer half as much risk. The clinician asks the mother for more details about the aunt. Per mother's report, the aunt has been psychiatrically hospitalized twice and treated with lithium as well as an atypical antipsychotic—all details that support a bipolar diagnosis. Conceptually, the aunt's history is a "yellow flag" increasing the index of suspicion for bipolar disorder. The clinician asks the mother to complete the half-page Family Index of Risk for Mood (Algorta et al., 2012) as a way of gathering information about other relatives. The aunt is the closest relative clearly affected by mood disorder, although other relatives have histories of substance use or depression. The clinician uses the probability nomogram (Figure 1) to estimate how the family history changes the probability of a bipolar disorder for Tandi. The clinician begins by plotting the base rate of bipolar spectrum disorder at the clinic on the left hand line of the nomogram, placing a dot at the 10%. The aunt's history of bipolar disorder would have a DLR of 2.5 (or half of the 5.0 attached to a first degree relative having bipolar disorder). The 2.5 is plotted on the middle line of the nomogram. Connecting the dots and extending across the right hand line yields an estimate of ~24% for the new, "posterior" probability of bipolar disorder. If the clinician used an online calculator instead of the nomogram, then he or she would generate a probability of 22%, not very different. The FIRM score of 8 also has a DLR of 2.5; plugging that DLR into the nomogram would lead to a probability of ~22 to 24%. Note that the clinician does not treat the FIRM score and the aunt's diagnosis as separate pieces of information. Instead, the clinician either chooses to focus on the one that seems more valid or uses each separately to generate two probabilities that "bracket" Tandi's risk in a form of sensitivity analysis that examines how sensitive the estimates are to changes in the inputs. Here, both results are close together. Both also are above the clinician's wait-test threshold. More assessment is needed to decide whether bipolar is present or absent for Tandi.

Family history of bipolar disorder also increases the risk of depression, ADHD, and a variety of other conditions, typically with a DLR in the range of 1.5 to 3.0 based on a meta-analysis (Hodgins, Faucher, Zarac, & Ellenbogen, 2002). However, because it is Tandi's aunt, not a first-degree relative, the conferred risk would be half as high (falling in the 1.25 to 1.5 range). This is low enough that the clinician decides to concentrate on looking for more valid information rather than spending time combining these DLRs with the prior probabilities for the other diagnoses (Straus et al., 2011).

Step 4. Synthesize Broad Instruments into Revised Probability Estimates. Tandi's mother completed the Child Behavior Checklist (CBCL) as part of the core intake battery the clinic uses. The *T* scores are 63 for Externalizing, 67 for Internalizing, 70 for Anxious/Depressed, 67 for Withdrawn/Depressed, 51 for Attention Problems, 66 for Aggressive Behavior, and 53 for Rule Breaking. Impressionistically, the scores could be consistent with an adjustment disorder (which is still the leading hypothesis) or depression. The Externalizing scores look mild for ODD, and the low Attention Problems decreases suspicion of ADHD substantially. The low Rule Breaking

score also decreases the probability of conduct disorder, which already was uncommon at the clinic (base rate of 10%). The clinician considers conduct disorder "ruled out" unless there is new information that increases concern about it. Adjustment disorder, depression, ODD, ADHD, and bipolar are still the focus of assessment. The clinician does a PubMed search on "Child Behavior Checklist" AND "bipolar disorder" AND "sensitivity and specificity" and finds a paper that published DLRs for the CBCL Externalizing score compared to a semi-structured diagnostic criterion (Youngstrom et al., 2004). The *T* of 63 is actually in the low range for youths with bipolar disorder, and it is more than twice as likely for youth to score in this range if they do not have a bipolar diagnosis (DLR = 0.47). The clinician uses the probability of 24% (from Step 3) as the new starting point on the nomogram left hand line, and plots the DLR of 0.47 on the midline, producing a revised estimate of ~15%. If the clinician used a calculator instead for all of the steps, the probability estimate would be 12%. Using similar approaches, the clinician finds that the probability of depression is up to about 65%, ADHD is down to below 20%, and no information is readily available for predicting adjustment disorder with the CBCL. To this point, the clinician has neither added any extra assessment tools to the battery except the FIRM nor spent any additional time interviewing the family. The steps have made the list of hypotheses and the interpretation more systematic than would otherwise often be the case, and relying on base rates and published weights counteracts potential cognitive heuristics due to the family's description of the presenting problem.

Step 5. Add Narrow and Incremental Assessments to Clarify Diagnoses. Based on the current hypotheses and probability estimates, the clinician decides to add some mood rating scales evaluating both depressive and hypomanic/manic symptoms as well as gather a teacher report about Tandi's school functioning. The clinician opts for the Achenbach Teacher Report Form as a concise way of gathering data about attention problems (potentially ruling ADHD out if low, vs. indicating continued assessment if high) as well as the degree of pervasiveness of the aggressive behaviors (helpful for the ODD hypothesis). The literature suggests that the teacher report of mood symptoms is unlikely to be helpful for differential diagnosis but could be helpful for treatment planning.

The clinician has Tandi complete the Child Depression Inventory (CDI; Kovacs, 1992) and the Mood Disorder Questionnaire (MDQ; Wagner et al., 2006), which has the easiest reading level of the hypomania/mania rating scales having published data with youths (Youngstrom, 2007). The clinician asks the mom to complete the Parent General Behavior Inventory, which asks about both depressive and hypomanic symptoms (PGBI; Youngstrom, Findling, Danielson, & Calabrese, 2001). Because the mother is specifically concerned about the possibility of bipolar disorder, the clinician and mother agree to have her do the full-length version rather than one of the abbreviated ones, to provide the most comprehensive description even though there is no statistical advantage of the longer versus shorter versions. Mom's scores for Tandi on the PGBI are 16 on the Hypomanic/Biphasic Scale (28 items) and 39 on the Depression Scale (46 items). The Hypomanic/Biphasic score falls in the low range for bipolar disorder, with a DLR of .46. Using the nomogram, this reduces the probability of a bipolar disorder to ~ 7%. Tandi's scores come back moderately high on the CDI and below threshold on the MDQ. Using the sensitivity (38%) and specificity (74%) published by Wagner et al. (2006) yields a DLR of 0.84. This is close enough to 1.0 that the clinician could ignore it rather than feeding it into the nomogram or a calculator; impressionistically, it is revising the low probability of bipolar disorder to become slightly lower still. The scores on the CDI and PGBI Depression are both suggestive of depression, raising the probability to ~85%.

Step 6. Interpret Cross-Informant Data Patterns. The Teacher Report Form (TRF) comes back with all scores below a *T* of 60. Tandi's grades have been good (all 3 s and 4 s on a 4-point scale). The low score on Attention Problems from the teacher, combined with the other assessment data, reduces the probability of ADHD below 5%. The clinician considers it functionally ruled out, based on the probability and the absence of any "red flags" in the academic record. The low scores do not change the probability of a mood disorder. They slightly reduce the chances of ODD. Tandi's high self-report of depressive symptoms is consistent with her mom's report of internalizing concerns, suggesting that Tandi may be motivated for treatment working on internalizing issues.

Step 7. Finalize Diagnoses by Adding Necessary Intensive Assessment Methods. The clinician selects the depression module of the MINI as a brief, structured interview to formally cover the diagnostic criteria for major depression and dysthymic disorder, along with the ODD module. The clinician also asks about recent life events and potential stressors, looking for possible precipitants for an adjustment disorder. At this stage, the clinician also considers other rival hypotheses that could be consistent with the presentation. Before diagnosing depression, we are supposed to rule out the possibility of medication side effects or general medical conditions. The clinician explains the rationale for doing the interview and asks about medications, vitamins, or other

drugs that Tandi might be taking. Tandi has had regular pediatrician visits, and her health has been good. She is not taking any prescription medication, and to her mom's knowledge, neither her peer group nor her older sister's is using any illicit substances. The MINI results identify a sufficient number of symptoms and duration for a diagnosis of a major depressive episode, with impairment at home. The severity appears mild to moderate based on the rating scales as well as descriptions during the MINI and the clinician's observations of Tandi. Based on assessment findings, the clinician assigns a diagnosis of major depressive disorder, single episode, moderate severity. The ODD module does not pass threshold, and the clinician formulates the irritability as being a feature of the depression rather than a separate diagnostic issue.

Step 8. Refine Assessment for Treatment Planning and Goal Setting. Based on the information so far, depression seems to be a main concern. The CDI and CBCL Internalizing provide good baseline scores for severity of the problem. The clinician has charts indicating the number of points each measure needs to change to demonstrate improvement (Youngstrom, 2007), based on the reliable change index approach, as well as benchmarks for treatment targets for "clinically significant change" on those as primary outcome measures (Jacobson & Truax, 1991). The clinician supplements this with measures of quality of life to look at positive aspects of functioning (Frisch, 1998) and selects the KINDL as a brief, developmentally appropriate instrument with both parent- and youth-report forms available (Ravens-Sieberer & Bullinger, 1998). To help decide which therapeutic modality might be most helpful in reducing the depressive symptoms, the clinician considers Tandi's verbal ability educational level of the family, and cultural background, all of which suggest a good fit with cognitive behavioral or psychoeducational approaches. The clinician also decides to gather more information about family functioning to gauge the extent to which family dynamics and communication might helpful to address, perhaps indicating a greater emphasis on family-focused therapy.

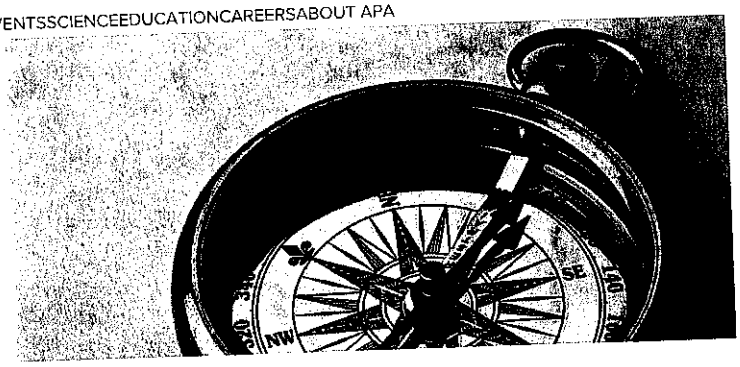
Step 12. Solicit and Integrate Patient Preferences. As noted in the article, it makes sense to do "Step 12" whenever in the assessment sequence it would be helpful in making decisions about assessment or treatment. The clinician presents the initial formulation to the family, discussing how changes in Tandi's mood can offer a parsimonious explanation for the clinical picture emerging from the testing. During the discussion, the clinician is able to directly address the mother's concern about possible bipolar disorder, stating that the probability of bipolar disorder is currently quite low, and pointing to specific findings establishing the basis for that judgment. The clinician and family discuss several different options for treatment, ranging from "wait and see," through individual therapy for Tandi (involving supportive discussion combined with problem-solving and coping skills coaching), or family therapy, and antidepressant medication. Because no one in the immediate family has taken an antidepressant before, the clinician talks through the risks and benefits, providing the number needed to treat and the number needed to harm estimates for each approach. The family decides to try an approach combining some family psychoeducation with individual therapy for Tandi, holding the medication in abeyance because her depressive symptoms are still only mild to moderate, and thus the potential benefit seems lower compared to the potential for side effects and the family's hesitation about using medication.

Step 9. Measure Processes ("Dashboards, Quizzes and Homework"). Tandi and her mother download a mood charting app onto the mother's smartphone, and they use this to track both of their moods on a daily basis. This feeds directly into the mood monitoring and problem-solving skills that the clinician works to teach Tandi in individual sessions. The clinician also uses a sticker chart with Tandi to track the number of times each week that she tries new problem solving skills.

Step 10. Chart Progress and Outcome. In addition to regularly reviewing the mood charting and "homework" sticker chart, the clinician has Tandi and her mom repeat the CDI and CBCL after six sessions to see if there is measurable improvement on the primary outcomes. The family completes these a third time, along with repeating the quality of life measures, as they approach the termination session. The updated scores are compared to the "clinical significance" benchmarks as well as the baseline scores. Discussing the benchmarks helps the mother to reduce her sense of perfectionism, and allays her concerns that Tandi's moodiness might be a sign of bipolar disorder, by giving her a better appreciation for the behaviors that fall within typical functioning for Tandi's age.

Step 11. Monitor Maintenance and Relapse. During the termination session, the clinician and family review progress, celebrate their success, and plan for the future. This includes a discussion about the possibility of relapse. The clinician decides that this is important to discuss given the high rate of relapse for depression, and the fact that both early onset of depression and family history of mood disorder are risk factors that

increase Tandi's chances of remission. The clinician frames the potential for relapse as a possibility but emphasizes that Tandi and the family have mastered the skills to beat mood issues. The group discusses what would be warning signs of depression starting to recur, and they also make a list of situations that might increase stress and risk for relapse (such as getting a bad grade, losing a friend, getting very sick, or if the family were to relocate . . .). The list is framed as a set of "reminders" to check in on everyone's mood and coping when dealing with stressful situations. The clinician and mother also discuss warning signs that might raise concern about bipolar disorder, as both the family history and early onset suggest that if Tandi develops future mood issues, they are more likely to follow a bipolar spectrum course over the long term, even though she did not show signs of bipolar illness during this initial episode.



Ethical Principles of Psychologists and Code of Conduct

Including 2010 and 2016 Amendments

9.01 Bases for Assessments

(a) Psychologists base the opinions contained in their recommendations, reports, and diagnostic or evaluative statements, including for testimony, on information and techniques sufficient to substantiate their findings. (See also Standard 2.04, Bases for Scientific and Professional Judgments (?item=5#204) .)

(b) Except as noted in 9.01c (#901c) , psychologists provide opinions of the psychological characteristics of individuals only after they have conducted an examination of the individuals adequate to support their statements or conclusions. When, despite reasonable efforts, such an examination is not practical, psychologists document the efforts they made and the result of those efforts, clarify the probable impact of their limited information on the reliability and validity of their opinions, and appropriately limit the nature and extent of their conclusions or recommendations. (See also Standards 2.01, Boundaries of Competence (?item=5#201) , and 9.06, Interpreting Assessment Results (#906) .)

(c) When psychologists conduct a record review or provide consultation or supervision and an individual examination is not warranted or necessary for the opinion, psychologists explain this and the sources of information on which they based their conclusions and recommendations.

9.02 Use of Assessments

(a) Psychologists administer, adapt, score, interpret, or use assessment techniques, interviews, tests, or instruments in a manner and for purposes that are appropriate in light of the research on or evidence of the usefulness and proper application of the techniques.

(b) Psychologists use assessment instruments whose validity and reliability have been established for use with members of the population tested. When such validity or reliability has not been established, psychologists describe the strengths and limitations of test results and interpretation.

(c) Psychologists use assessment methods that are appropriate to an individual's language preference and competence, unless the use of an alternative language is relevant to the assessment issues.

9.03 Informed Consent in Assessments

(a) Psychologists obtain informed consent for assessments, evaluations, or diagnostic services, as described in Standard 3.10, Informed Consent, except when (1) testing is mandated by law or governmental regulations; (2) informed consent is implied because testing is conducted as a routine educational, institutional, or organizational activity (e.g., when participants voluntarily agree to assessment when applying for a job); or (3) one purpose of the testing is to evaluate decisional capacity. Informed consent includes an explanation of the nature and purpose of the assessment, fees, involvement of third parties, and limits of confidentiality and sufficient opportunity for the client/patient to ask questions and receive answers.

(b) Psychologists inform persons with questionable capacity to consent or for whom testing is mandated by law or governmental regulations about the nature and purpose of the proposed assessment services, using language that is reasonably understandable to the person being assessed.

(c) Psychologists using the services of an interpreter obtain informed consent from the client/patient to use that interpreter, ensure that confidentiality of test results and test security are maintained, and include in their recommendations, reports, and diagnostic or evaluative statements, including forensic testimony, discussion of any limitations on the data obtained. (See also Standards 2.05, Delegation of Work to Others (?item=5#205) ; 4.01, Maintaining Confidentiality (?item=7#401) ; 9.01, Bases for Assessments (#901) ; 9.06, Interpreting Assessment Results (#906) ; and 9.07, Assessment by Unqualified Persons (#907) .)

9.04 Release of Test Data

(a) The term *test data* refers to raw and scaled scores, client/patient responses to test questions or stimuli, and psychologists' notes and recordings concerning client/patient statements and behavior during an examination. Those portions of test materials that include client/patient responses are included in the definition of *test data*. Pursuant to a client/patient release, psychologists provide test data to the client/patient or other persons identified in the release. Psychologists may refrain from releasing test data to protect a client/patient or others from substantial harm or misuse or misrepresentation of the data or the test, recognizing that in many instances release of confidential information under these circumstances is regulated by law. (See also Standard 9.11, Maintaining Test Security (#911) .)

(b) In the absence of a client/patient release, psychologists provide test data only as required by law or court order.

9.05 Test Construction

Psychologists who develop tests and other assessment techniques use appropriate psychometric procedures and current scientific or professional knowledge for test design, standardization, validation, reduction or elimination of bias, and recommendations for use.

9.06 Interpreting Assessment Results

When interpreting assessment results, including automated interpretations, psychologists take into account the purpose of the assessment as well as the various test factors, test-taking abilities, and other characteristics of the person being assessed, such as situational, personal, linguistic, and cultural differences, that might affect psychologists' judgments or reduce the accuracy of their interpretations. They indicate any significant limitations of their interpretations. (See also Standards 2.01b and c, Boundaries of Competence (?item=5#201b) , and 3.01, Unfair Discrimination (?item=6#301) .)

9.07 Assessment by Unqualified Persons

Psychologists do not promote the use of psychological assessment techniques by unqualified persons, except when such use is conducted for training purposes with appropriate supervision. (See also Standard 2.05, Delegation of Work to Others (?item=5#205) .)

9.08 Obsolete Tests and Outdated Test Results

(a) Psychologists do not base their assessment or intervention decisions or recommendations on data or test results that are outdated for the current purpose.

(b) Psychologists do not base such decisions or recommendations on tests and measures that are obsolete and not useful for the curre purpose.

9.09 Test Scoring and Interpretation Services

(a) Psychologists who offer assessment or scoring services to other professionals accurately describe the purpose, norms, validity, reliability, and applications of the procedures and any special qualifications applicable to their use.

(b) Psychologists select scoring and interpretation services (including automated services) on the basis of evidence of the validity of the program and procedures as well as on other appropriate considerations. (See also Standard 2.01b and c, Boundaries of Competence (?item=5#201b) .)

(c) Psychologists retain responsibility for the appropriate application, interpretation, and use of assessment instruments, whether they score and interpret such tests themselves or use automated or other services.

9.10 Explaining Assessment Results

Regardless of whether the scoring and interpretation are done by psychologists, by employees or assistants, or by automated or other outside services, psychologists take reasonable steps to ensure that explanations of results are given to the individual or designated representative unless the nature of the relationship precludes provision of an explanation of results (such as in some organizational consulting, preemployment or security screenings, and forensic evaluations), and this fact has been clearly explained to the person being assessed in advance.

9.11 Maintaining Test Security

The term *test materials* refers to manuals, instruments, protocols, and test questions or stimuli and does not include *test data* as defined in Standard 9.04, Release of Test Data (#904) . Psychologists make reasonable efforts to maintain the integrity and security of test materials and other assessment techniques consistent with law and contractual obligations, and in a manner that permits adherence to this Ethics Code.

Section 10: Therapy

History and Effective Date

Amendments to the 2002 "Ethical Principles of Psychologists and Code of Conduct" in 2010 and 2016

American Psychological Association

Psychological Assessment

2019, Vol. 31, No. 1, 120–125
http://dx.doi.org/10.1037/pas0000650

1040-3590/19/$12.00

BRIEF REPORT

Examining False-Positive Rates of Wechsler Adult Intelligence Scale (WAIS-IV) Processing Speed-Based Embedded Validity Indicators Among Individuals With Schizophrenia Spectrum Disorders

David M. Glassmire, Mary Elizabeth Wood, Minhdan T. Ta, Dominique I. Kinney, and Stephen R. Nitch
Patton State Hospital, Patton, California

Recent research (Erdodi et al., 2017) indicated that certain Wechsler Adult Intelligence Scale (WAIS-IV) Processing Speed Index (PSI)–based indices may have some utility as embedded validity indicators (EVIs) among a diagnostically diverse sample of neuropsychology referrals. Individuals with schizophrenia spectrum disorders (SSD) are often evaluated in forensic contexts in which there is incentive to exaggerate deficits. Because individuals with SSD often have limitations in processing speed associated with their disorders, the current study sought to evaluate the false-positive (FP) rates of cutoffs identified by Erdodi et al. on WAIS-IV PSI-based EVIs among forensically committed psychiatric inpatients with SSD who had no known incentive to feign because of the nature of their legal commitments. In the current sample, the previously suggested cutoff scores on PSI-based EVIs resulted in FP rates ranging from 2% to 57% among schizophrenia spectrum patients, with unacceptable FP rates for most indices. In the current sample of SSD patients, WAIS-IV PSI-based EVIs that are calculated based on the relative performance between PSI subtests (as opposed to absolute performance on individual indices) demonstrated acceptable FP rates.

Public Significance Statement
This study showed that forensic psychiatric patients with schizophrenia spectrum disorders often perform poorly on tests that measure how quickly cognitive information can be processed, even when giving their best effort. Therefore, these tests are not good measures of effort for people with schizophrenia spectrum disorders.

Keywords: embedded validity indicators, schizophrenia spectrum disorders, malingering, performance validity

The American Academy of Clinical Neuropsychology Consensus Conference Statement on Neuropsychological Assessment of Effort, Response Bias, and Malingering indicated consensus among experts that psychometric indicators are the most valid approach to establishing performance validity (PV) during neuropsychological assessment and that both stand-alone and embedded validity indicators (EVIs) should be used as part of an assessment battery (Heilbronner et al., 2009). Some potential advantages to including EVIs are that they are time efficient (Jasinski, Berry, Shandera, & Clark, 2011; Mathias, Greve, Bianchini, Houston, & Crouch, 2002) and may be less susceptible than stand-alone measures to attorney coaching (Victor & Abeles, 2004; Youngjohn, 1995) or Internet exposure to information regarding PV assessment (Ruiz, Drake, Glass, Marcotte, & van Gorp, 2002). Many investigators have evaluated the utility of potential EVIs within objective cognitive tests such as the Wechsler Adult Intelligence Scale-Fourth Edition (WAIS-IV; Wechsler, 2008) and its predecessors. Although the most heavily researched of these indices are those based on the Digit Span subtest (Jasinski et al., 2011; Schroeder, Twumasi-Ankrah, Baade, & Marshall, 2012), measures of processing speed from various versions of the WAIS have also been investigated as possible EVIs.

The Processing Speed Index (PSI) on both the WAIS-III and WAIS-IV consists of two subtests: symbol search (SS) and (digit-

This article was published Online First August 13, 2018.

David M. Glassmire, Mary Elizabeth Wood, Minhdan T. Ta, Dominique I. Kinney, and Stephen R. Nitch, Department of Psychology, Patton State Hospital, Patton, California.

The statements and opinions in this article are those of the authors and do not constitute the official views or the official policy of Department of State Hospitals (DSH)-Patton, the California Department of Mental Health, or the State of California.

Correspondence concerning this article should be addressed to David M. Glassmire, Department of Psychology, Patton State Hospital, AX-234, 3102 Highland Avenue, Patton, CA 92369. E-mail: david.glassmire@dsh.ca.gov

symbol) coding (CD). These timed subtests require examinees to process and respond to visual information as quickly as possible. Therefore, an examinee's speed of cognitive processing strongly influences scores on both measures, potentially making them susceptible to false-positive (FP) errors when used to evaluate PV among examinees with clinical conditions that are associated with deficits in processing speed. Prior research using PSI measures as EVIs on the WAIS-R and WAIS-III generally indicates that very low cutoff scores need to be utilized to maintain adequate specificity (i.e., low FP rates) and that these low cutoff scores are associated with poor sensitivity to feigning or invalid performance (Trueblood, 1994; Etherton, Bianchini, Heinly, & Greve, 2006; Kim et al., 2010; Curtis, Greve, & Bianchini, 2009).

To date, only one published study (Erdodi et al., 2017) has evaluated the utility of WAIS-IV (Wechsler, 2008) PSI indices as EVIs. These authors found that a WAIS-IV PSI standard score cutoff of ≤79 had sensitivities ranging from .23 to .56 (depending on how criterion groups were formed), with specificities ranging from .92 to .98 (corresponding to FP rates of 2% to 8%) among a diagnostically heterogeneous mixed neuropsychology referral sample. The CD age-adjusted scaled score (CD_{AASS}) cutoff of ≤5 had sensitivities ranging from .04 to .28 and specificities ranging from .94 to 1.00. The SS Age-Adjusted Scaled Score (SS_{AASS}) cutoff of ≤6 had sensitivities ranging from .38 to .64 and specificities ranging from .88 to .90. Erdodi et al. (2017) found that two additional variables, calculated by examining relative performance across PSI subtests, showed promise as EVIs. The first variable was the absolute value of CD_{AASS} minus SS_{AASS} (CD–SS). It was hypothesized that because both subtests measure similar skills and were conormed on the same sample, scores among valid-responding examinees should be similar on these two subtests. The authors found that CD–SS ≥5 had sensitivities ranging from .08 to .12 and specificities ranging from .89 to .91. A second variable, the ratio of CD-to-SS raw scores (CD/SS) cutoff score of ≤1.41 or ≥3.57 had sensitivities ranging from .15 to .24 and specificities ranging from .87 to .93. Finally, Erdodi et al. found that a cutoff of ≥3 on a composite based on the sum of failures on the five aforementioned validity indicators at the most liberal cutoff scores had sensitivities ranging from .23 to .53 and specificities ranging from .89 to .94. Consistent with research on the WAIS-R and WAIS-III PSI-based EVIs, Erdodi et al.'s findings indicated that cutoffs could be identified with adequate specificity (i.e., low FP rates) but that such cutoff scores generally were associated with low sensitivity. Consequently, these authors concluded that PSI-based EVIs have some utility but should not be used in isolation because of their low sensitivity to feigning.

Individuals with schizophrenia spectrum disorders (SSDs) are often assessed within criminal forensic (Viljoen & Zapf, 2002) and disability evaluation (Piechowski, 2011) contexts in which the validity of obtained results is particularly important, given the potential incentive to exaggerate cognitive deficits in such contexts. Despite the frequency with which individuals with SSDs are assessed in forensic contexts, previous findings on PSI-based EVIs might not generalize to this population because the majority of previous studies investigated samples of mixed neuropsychology referrals or individuals with other diagnoses such as pain patients, traumatic brain injury patients, or patients with memory disorders. Although almost half of the sample of Erdodi et al. (2017; 47.3%) was characterized as having psychiatric diagnoses, the percentage of examinees with SSDs was not reported and the sample's overall FSIQ of 100.7 (SD = 15.4) was approximately 1⅓ SD higher than the schizophrenia sample from the WAIS-IV standardization groups (M = 79.8, SD = 6.3; Wechsler, 2009). Therefore, the cutoff scores identified by Erdodi et al. on WAIS-IV PSI-based EVIs might not generalize to evaluation settings characterized by high prevalence rates of SSD.

Individuals diagnosed with SSDs often display broad neurocognitive impairments, particularly in the areas of attention, working memory, processing speed, memory, and executive functioning (Heinrichs & Zakzanis, 1998; Mesholam-Gately, Giuliano, Goff, Faraone, & Seidman, 2009; Palmer, Dawes, & Heaton, 2009). The presence of these deficits raises concerns that individuals with SSDs might perform poorly on performance-based PV indices because of bona fide symptoms of the disorder (e.g., poor ability to sustain persistent task effort) rather than because of volitional attempts to provide intentionally invalid effort, thereby resulting in FP errors (Gorissen, Sanz, & Schmand, 2005). In fact, on the Dot-Counting Test (Boone, Lu, & Herzberg, 2002), a stand-alone performance validity test (PVT) that relies heavily on processing speed, the recommended cutoff score for SSD is one of the higher scores provided in the manual because a higher cutoff score was necessary to maintain an adequate FP rate among individuals with SSDs. Because of potential impact of bona fide cognitive impairments on WAIS-IV PSI-based EVIs among individuals with SSDs, normative data are needed for patients who represent the full spectrum of cognitive and intellectual levels typically seen within this group. The present study was conducted to determine the FP rates of WAIS-IV PSI-based EVIs among forensic psychiatric patients with SSDs who had no known incentive to perform poorly because of the nature of their legal commitments.

Method

Participants

This study received institutional review board approval from the California Committee for the Protection of Human Subjects. Data for the current study were taken from archival records of WAIS-IV administrations at a large, inpatient forensic hospital (N = 501). Given the nature of the study, data were excluded for patients with an ostensible reason to feign (i.e., removal of 199 records for individuals adjudicated incompetent to stand trial). As such, only data for patients who were committed to the hospital with legal commitments of not guilty by reason of insanity (NGRI), California's mentally disordered offender (MDO) statute, civil conservatorships (CC), or any combination thereof were included. Of import, patients adjudicated and committed as NGRI, MDO, and/or CC have incentive to appear their best psychologically because their forensic criteria for hospital release require them to demonstrate that their symptoms are controlled such that they no longer pose a danger to others because of their mental disorder (i.e., NGRI and MDO) or can live independently in the community (i.e., CC). Weinborn, Orr, Woods, Conover, and Feix (2003) noted that NGRI and CC patients are "more likely to minimize or deny cognitive impairment in order to increase the likelihood of discharge" (p. 982). Consequently, previous investigators have used NGRI and CC samples to investigate the specificity of feigning

measures (e.g., Weinborn et al., 2003; Glassmire, Jhawar, Burchett, & Tarescavage, 2017).

To be included in the current study, the primary diagnosis on the date of testing was required to be a *Diagnostic and Statistical Manual of Mental Disorders*, fourth edition (APA, 2000), revised, psychotic disorder diagnosis (i.e., diagnoses that would be classified as SSD under *Diagnostic and Statistical Manual of Mental Disorders*, fifth edition [APA, 2013]; $N = 220$). All diagnoses were rendered by an interdisciplinary treatment team that included a psychologist, a psychiatrist, a social worker, a rehabilitation therapist, and nursing staff. All patients were housed in an inpatient forensic facility in which they were observed on a 24-hr basis and their functioning and symptom presentation were well documented by nursing staff and clinical team members as part of the ongoing diagnostic and treatment process.

Several methods were utilized to identify and exclude patients for whom inadequate effort was suspected. First, patients were excluded if their performance on an at least one PVT was indicative of inadequate effort ($N = 5$).[1] Using cutoff scores identified by Glassmire, Toofanian Ross, Kinney, and Nitch (2016) for use of the Reliable Digit Span as an EVI in samples with SSDs, two additional patients were removed for earning a score of ≤4 on the Reliable Digit Span. In addition, data were included only for cases in which the test administration was explicitly characterized as valid in the report ($N = 140$), and an additional 20 records were excluded because the clinician did not explicitly opine that the patient was not feigning or malingering ($N = 120$). Although these cases likely represented scenarios in which the clinician believed the effort was valid and therefore did not require comment, we removed them to retain only cases that were explicitly opined to be valid by the clinician. In the small number of cases with discrepancies between clinician opinion and PVT, the discrepancies generally resulted from clinician opinions that standard PVT cutoffs resulted in FP classifications in the case at hand. Nonetheless, we removed all cases in which either the PVT or clinician opinion indicated a possibility of invalid effort. Although this method for classifying valid and invalid protocols resulted in the removal of many records, the intention was to increase certainty that only valid protocols were included. One final participant was excluded for a comorbid intellectual disability diagnosis, resulting in a final sample of data from 118 WAIS-IV administrations for forensic inpatients with SSDs. Most of the sample completed enough subtests to calculate WAIS-IV index scores (*ns* ranging from 110 to 115). Demographic information and WAIS-IV Index scores for the sample are presented in Table 1.

Table 1
Demographic Data and WAIS-IV Performances of Sample ($n = 118$)

Variables	*M* (*SD*)
Age (years)	41.0 (12.51)
Education (years)	11.3 (2.2)
Commitment length (days)	1776.4 (2232.8)
WAIS-IV indices	
Full-scale IQ	80.70 (14.97)
Verbal Comprehension Index	84.25 (15.35)
Perceptual Reasoning Index	86.66 (14.04)
Working Memory Index	83.20 (15.32)
Processing Speed Index	79.33 (14.99)
General Ability Index	83.90 (14.76)
	N / %
Legal commitment	70 / 59.3
Not guilty by reason of insanity	37 / 39.0
Mentally disordered offender	2 / 1.7
Civil conservatorship	9 / 7.6
Combination of above commitments	
Gender	93 / 78.8
Male	25 / 21.2
Female	
Ethnicity	8 / 6.7
Asian/Asian Pacific Islander	1 / .8
American Indian	44 / 37.3
African American	40 / 33.9
Caucasian	23 / 19.5
Hispanic/Latino	
Primary *DSM-IV* diagnosis	74 / 62.2
Schizophrenia (total)	43 / 58.1
Paranoid subtype	26 / 35.1
Disorganized subtype	5 / 6.8
Undifferentiated subtype	43 / 33.6
Schizoaffective disorder	4 / 3.4
Psychotic disorder NOS	

Note. *DSM-IV* = *Diagnostic and Statistical Manual of Mental Disorders*, fourth edition; WAIS-IV = Wechsler Adult Intelligence Scale, fourth edition: IQ = intelligence quotient; NOS = not otherwise specified. The median commitment length was 722.5 days. Percentages reported for schizophrenia subtypes are percentages of individuals with schizophrenia diagnoses who had each subtype. All other percentages are out of the total sample.

Measures

The WAIS-IV (Wechsler, 2008) is a comprehensive test of intellectual ability measured via an overall Full-Scale IQ (FSIQ) and four individual index scores: Verbal Comprehension Index, Perceptual Reasoning Index, Working Memory Index, and PSI. In addition, the General Ability Index (GAI) is calculated from the subtests that comprise the Verbal Comprehension Index and Perceptual Reasoning Index and provides a composite index that does not include the Working Memory Index or PSI subtests. Therefore, the GAI is not as heavily impacted by deficits in working memory or processing speed.

Analyses

Within the total sample, correlations were computed to determine whether the PSI, CD_{AASS}, SS_{AASS}, CD-SS, and CD/SS were significantly associated with the GAI. The GAI was used as the value to represent intellectual ability as opposed to FSIQ, given the criterion contamination between the PSI measures and FSIQ. To determine whether patients in various GAI bands performed dif-

[1] The criteria for PVT failure were taken from established cut scores published in the test manuals and available literature for use with this particular population (i.e., samples of individuals with schizophrenia spectrum illnesses). Criteria included the following: Dot-Counting Test E-Score ≥20; Rey 15-Item Combination Score <20 (Boone, Salazar, Lu, Warner-Chacon, & Razani, 2002); Test of Memory Malingering (Tombaugh, 1996) Trial 2 or Retention Trial Score <45).

ferently, WAIS-IV PSI-based EVIs were subjected to either multivariate or one-way analyses of variance (ANOVAs; GAI ≥80, GAI 70–79, and GAI ≤69). The specificity rates for a range of cutoff scores were calculated for each variable for the total sample and within each GAI band.

Results

The GAI significantly correlated with the PSI (r = .58), CD_{AASS} (r = .56), and SS_{AASS} (r = .54; all p's < .001) but with neither of the ancillary measures developed by Erdodi et al. (2017); significance values of .69 and .66 for the CD/SS and CD-SS measures, respectively). A multivariate ANOVA was computed with CD_{AASS} and SS_{AASS} as the dependent variables; the overall analysis was significant, $F(4, 214) = 11.79$, $p < .001$, $\eta_p^2 = .18$, as were the tests of between-subject effects for both the CD_{AASS} and SS_{AASS} measures (i.e., $F[2, 109] = 23.12$, $p < .001$, $\eta_p^2 = .30$, and $F[2, 109] = 19.76$, $p < .001$, $\eta_p^2 = .27$, respectively). Bonferroni-corrected post hoc testing revealed that patients scoring within the highest GAI band (i.e., GAI ≥80) earned significantly higher CD_{AASS} and SS_{AASS} relative to patients in either of the other two GAI bands (see Table 2), who did not differ from one another. One-way ANOVAs were also computed to test whether the remaining measures (i.e., PSI, CD/SS, and CD-SS) differed among GAI bands. The groups significantly differed on PSI, $F(2, 108) = 24.74$, $p < .001$, $\eta_p^2 = .31$, with patients in the highest GAI band scoring higher than patients in the remaining two groups, who again did not differ. Patients performed similarly on the two ancillary measures, regardless of GAI performance (i.e., p = .97 and .58 for the CD-SS and CD/SS comparisons, respectively). Given that scores differed on some of the measures by GAI, specificity rates were calculated for the total sample and within each GAI band (see Table 3). Within the overall sample, the following cutoff scores needed to be adopted to maintain a specificity of 90% or higher in the total sample (i.e., 10% or lower FP error rates): PSI ≤59, CD_{AASS} = 1, SS_{AASS} ≤ 2, and CD-SS ≥ 3. The CD/SS value ranges ≤1.41 or ≥3.57 recommended by Erdodi et al. (2017) resulted in a specificity of .92 in the total sample. It was notable that the indices that are based on absolute performance on a single index or subscale (PSI, CD_{AASS}, and SS_{AASS}) required adoption of very low cutoff scores to maintain FP rates below 10%.

Discussion

Because examinees with SSDs are frequently administered the WAIS-IV in various forensic contexts, it is imperative to have information regarding appropriate EVI cutoff scores for this population. The present study was conducted to determine whether the WAIS-IV PSI-based EVI cutoffs originally identified by Erdodi et al. (2017) among a diagnostically heterogeneous sample of neuropsychology referrals demonstrate acceptable FP rates in a forensic SSD sample, given the well-established link between SSD and deficits in processing speed.

Consistent with a priori hypotheses, our results suggest that EVIs based on absolute performance (i.e., PSI, CD_{AASS}, and SS_{AASS}) produce unacceptably high FP rates among individuals with SSDs. Indeed, between 49% (CD_{AASS} ≤ 5) and 57% (SS_{AASS} ≤6) of the total sample of SSD patients would be inappropriately classified as producing invalid effort when using cut scores on these indices recommended by Erdodi et al. (2017). It should be noted that Erdodi et al.'s sample was diagnostically diverse and had higher levels of intellectual functioning than usually seen in SSD samples. The very high FP rates in the present study are of particular concern, given the stringent exclusionary criteria used to exclude invalid WAIS-IV protocols because the stringent exclusionary criteria likely resulted in removal of some of the lowest functioning patients from the final sample. The high FP rates in our sample are not surprising because individuals with SSDs often have deficits in processing speed, which would be

Table 2
Average PSI Measure Scores Different GAI Bands

PSI measure	GAI ≥80 *M* (*SD*)	GAI 70–79 *M* (*SD*)	GAI ≤69 *M* (*SD*)	Total sample *M* (*SD*)
PSI	87.58 (13.00)$_a$	71.95 (12.59)$_b$	68.37 (11.11)$_b$	79.46 (15.09)
CD_{AASS}	7.61 (2.75)$_a$	4.28 (2.31)$_b$	4.06 (2.72)$_b$	5.95 (3.09)
SS_{AASS}	7.79 (2.49)$_a$	5.13 (2.64)$_b$	4.25 (1.81)$_b$	6.36 (2.86)
CD/SS	2.15 (.43)$_a$	2.02 (.77)$_a$	2.16 (.71)$_a$	2.10 (.61)
CD-SS	1.47 (1.12)$_a$	1.47 (1.20)$_a$	1.56 (1.75)$_a$	1.49 (1.24)

Note. Means in the same row with the same subscript did not differ at $p <$ .05 using Bonferroni post hoc pairwise comparisons. Total sample, n = 111; GAI ≥80, n = 57; GAI, 70–79, n = 38; GAI, ≤69, n = 16.

Table 3
Specificity Rates for PSI PV Measures Across GAI Bands

Cutoff score	Specificity (%) GAI ≤69	GAI 70–79	GAI ≥80	Total sample
PSI				
≤50	**.94**	**.97**	1.00	.98
≤59	.44	.87	1.00	**.93**
≤68	.44	.55	**.93**	.73
≤74	.31	.40	.81	.59
≤79	.19	.29	.70	.48
CD_{AASS}				
1	.69	**.92**	1.00	**.93**
≤2	.56	.76	.98	.87
≤3	.56	.61	**.93**	.77
≤4	.38	.42	.86	.64
≤5	.31	.26	.75	.51
≤6	.25	.16	.61	.40
SS_{AASS}				
1	**.94**	**.95**	1.00	.97
≤2	.81	.87	1.00	**.93**
≤3	.63	.82	.98	.88
≤4	.50	.55	**.93**	.74
≤5	.25	.37	.81	.58
≤6	.13	.26	.65	.43
CD/SS				
≤1.41/≥3.57	.87	**.92**	**.98**	**.92**
CD-SS				
1	.69	.58	.63	.63
≥2	.69	.84	.88	.84
≥3	.88	**.97**	**.93**	**.94**
≥4	**.94**	.97	.97	.97
≥5	.94	.97	1.00	.98
≥6	1.00	1.00	1.00	1.00

Note. Specificity is the percentage of participants scoring above the cutoff who were classified as having put forth valid effort using the cutoff. Numbers in bold indicate the highest cutoff scores that can be adopted while maintaining a specificity rate of 90% or higher. Total sample, n = 111; GAI ≥80, n = 57; GAI 70–79, n = 38; GAI ≤69, n = 16.

predicted to result in low scores on many PSI-based EVIs, even among individuals giving their best effort. Erdodi et al. recommended that their EVI cutoffs not be used as sole indicators of invalid performance because of high FP rates they found among individuals with moderate to severe traumatic brain injury. Our results indicate that similar caution should apply when evaluating individuals with SSD and that very low cutoff scores would need to be adopted to maintain acceptable FP rates with these diagnoses. In fact, on CD_{AASS}, a score of 1 would need to be adopted to maintain an FP rate below 10%. An AASS of 1 is the lowest possible score that can be attained on any WAIS-IV subtest, suggesting that this cutoff has little practical utility as an EVI.

In addition to unacceptably high FP rates associated with the cut scores on the absolute PSI-based indices evaluated by Erdodi et al. (2017), we found a significant association between these indices and overall intellectual ability, as measured by the GAI. Indeed, this association, as well as the relative reduction in specificity found among individuals with lower GAI performance, suggests that these indices disproportionally classify individuals who have both SSD and lower intellectual functioning as feigning.

The specificity estimates for ancillary measures developed by Erdodi et al. (2017), including CD-SS and CD/SS, were more promising. Indeed, these scores were not significantly correlated with the GAI in our sample, suggesting their utility for a range of ability levels among individuals with SSD. The conversion of PSI-based scores to a metric that uses relative as opposed to absolute performance effectively reduces systematic bias against individuals who perform more poorly than the normative sample. In other words, if a patient's performance is low on the CD_{AASS} and SS_{AASS}, the unconverted scores are likely to penalize (and potentially misclassify) the individual; however, comparison of relative performances on the two subtests, via looking at the absolute value of the difference and/or the ratio of the two scores, renders the reduction in (absolute) performance irrelevant. In our sample, the cut scores recommended by Erdodi et al. on these measures resulted in FP rates of 8% (CD/SS $\leq$1.41 or $\geq$3.57) and 2% (CD-SS $\geq$ 5). In fact, in our sample, a more liberal cut score of CD-SS $\geq$3 maintained adequate specificity, suggesting that individuals with SSDs have more uniformly poor processing speed than do other individuals presenting for neuropsychological evaluation. Given their low FP rates in our sample, we recommend that these indices based on relative performance on CD and SS be used as one part of a more comprehensive battery of PV measures. However, given the low sensitivity to feigning demonstrated by Erdodi et al. for these indices, scores in the valid range on these indices are inconclusive among examinees with SSDs. Therefore, these indices provide useful information only regarding PV when they fall in the invalid range.

The strength of this study is that it included SSD patients in a forensic setting and that very stringent exclusionary criteria were used to remove any patients who had any known incentive to perform poorly and/or who had indication of putting forth invalid effort. This increases confidence that any performances falling in the invalid range on an EVI are likely FP classifications. Moreover, the use of patients assessed in a forensic context increases generalizability to settings in which PV assessment is particularly important. A limitation of this study is that it does not provide data regarding the sensitivity of different cutoff scores on the WAIS-IV PSI-based EVIs. However, the focus of the present study was to investigate the FP rates of Erdodi et al. (2017) suggested cutoffs and to determine whether any cutoff scores can be adopted on these EVIs that have adequate specificity levels for use with forensic examinees with SSD. The present results suggest that very low cutoff scores must be adopted on the indices based on absolute performance (i.e., PSI, CD_{AASS}, and SS_{AASS}) to maintain adequate specificity in individuals with SSD. Erdodi et al. already demonstrated low sensitivity of these cutoffs in a more diagnostically heterogeneous sample of neuropsychology referrals. Therefore, the lack of additional sensitivity data for these measures did not detract from the main purpose of the study.

Because of the well-documented cognitive symptoms associated with SSD (Heinrichs et al., 1998; Mesholam-Gately et al., 2009; Palmer et al., 2009), it was hypothesized that the use of WAIS-IV PSI-based EVIs with SSD patients would result in unacceptable FP classifications. Overall, our results suggest that clinicians evaluating individuals with SSDs in a forensic context should not rely on WAIS-IV PSI-based EVIs that are derived from absolute scores because of unacceptable FP rates in this population. Scores falling in the invalid range on the two indices based on relative performance on CD and SS were associated with low FP rates and can therefore be used to signal the need for additional testing with PVTs that have demonstrated better utility for use with individuals with SSD. However, given the low sensitivity of these indices demonstrated by Erdodi et al. (2017), negative findings on these indices do not necessarily indicate valid performance. Therefore, overall, PSI-based EVIs provide limited utility in evaluating PV among examinees with SSDs.

References

American Psychiatric Association (APA). (2000). *Diagnostic and statistical manual of mental disorders: DSM-IV-TR*. Washington, DC: Author.

American Psychiatric Association (APA). (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.

Boone, K., Lu, P., & Herzberg, D. S. (2002). *The Dot Counting Test manual*. Los Angeles, CA: Western Psychological Services.

Boone, K. B., Salazar, X., Lu. P., Warner-Chacon. K., & Razani, J. (2002). The Rey 15-item recognition trial: A technique to enhance sensitivity of the Rey 15-item memorization test. *Journal of Clinical and Experimental Neuropsychology, 24,* 561–573. http://dx.doi.org/10.1076/jcen.24.5.561.1004

Curtis. K. L., Greve, K. W., & Bianchini, K. J. (2009). The Wechsler Adult Intelligence Scale-III and malingering in traumatic brain injury: Classification accuracy in known groups. *Clinical Neuropsychologist, 22,* 930–945. http://dx.doi.org/10.1080/13854040701563591

Erdodi, L. A., Abeare, C. A., Lichtenstein, J. D., Tyson, B. T., Kucharski, B., Zuccato, B. G., & Roth, R. M. (2017). Wechsler Adult Intelligence Scale-Fourth Edition (WAIS-IV) processing speed scores as measures of noncredible responding: The third generation of embedded performance validity indicators. *Psychological Assessment, 29,* 148–157. http://dx.doi.org/10.1037/pas0000319

Etherton, J. L., Bianchini, K. J., Heinly, M. T., & Greve, K. W. (2006). Pain, malingering, and performance on the WAIS-III Processing Speed Index. *Journal of Clinical and Experimental Neuropsychology, 28,* 1218–1237. http://dx.doi.org/10.1080/13803390500346595

Glassmire, D. M., Jhawar, A., Burchett, D., & Tarescavage, A. M. (2017). Evaluating item endorsement rates for the MMPI-2-RF F-r and Fp-r scales across ethnic, gender, and diagnostic groups with a forensic inpatient sample. *Psychological Assessment, 29,* 500–508. http://dx.doi.org/10.1037/pas0000366

Glassmire, D. M., Toofanian Ross, P., Kinney, D. I., & Nitch, S. R. (2016). Derivation and cross-validation of cutoff scores for patients with schizophrenia spectrum disorders on WAIS-IV Digit Span-based performance validity measures. *Assessment, 23,* 292–306. http://dx.doi.org/10.1177/1073191115587551

Gorissen, M., Sanz, J. C., & Schmand, B. (2005). Effort and cognition in schizophrenia patients. *Schizophrenia Research, 78,* 199–208. http://dx.doi.org/10.1016/j.schres.2005.02.016

Heilbronner, R. L., Sweet, J. J., Morgan, J. E., Larrabee, G. J., & Millis, S. R., & the Conference Participants. (2009). American Academy of Clinical Neuropsychology Consensus Conference Statement on the neuropsychological assessment of effort, response bias, and malingering. *Clinical Neuropsychologist, 23,* 1093–1129. http://dx.doi.org/10.1080/13854040903155063

Heinrichs, R. W., & Zakzanis, K. K. (1998). Neurocognitive deficit in schizophrenia: A quantitative review of the evidence. *Neuropsychology, 12,* 426–445. http://dx.doi.org/10.1037/0894-4105.12.3.426

Jasinski, L. J., Berry, D. T. R., Shandera, A. L., & Clark, J. A. (2011). Use of the Wechsler Adult Intelligence Scale Digit Span subtest for malingering detection: A meta-analytic review. *Journal of Clinical and Experimental Neuropsychology, 33,* 300–314. http://dx.doi.org/10.1080/13803395.2010.516743

Kim, N., Boone, K. B., Victor, T., Lu, P., Keatinge, C., & Mitchell, C. (2010). Sensitivity and specificity of a digit symbol recognition trial in the identification of response bias. *Archives of Clinical Neuropsychology, 25,* 420–428. http://dx.doi.org/10.1093/arclin/acq040

Mathias, C. W., Greve, K. W., Bianchini, K. J., Houston, R. J., & Crouch, J. A. (2002). Detecting malingered neurocognitive dysfunction using the reliable digit span in traumatic brain injury. *Assessment, 9,* 301–308. http://dx.doi.org/10.1177/1073191102009003009

Mesholam-Gately, R. I., Giuliano, A. J., Goff, K. P., Faraone, S. V., & Seidman, L. J. (2009). Neurocognition in first-episode schizophrenia: A meta-analytic review. *Neuropsychology, 23,* 315–336. http://dx.doi.org/10.1037/a0014708

Palmer, B. W., Dawes, S. E., & Heaton, R. K. (2009). What do we know about neuropsychological aspects of schizophrenia? *Neuropsychology Review, 19,* 365–384. http://dx.doi.org/10.1007/s11065-009-9109-y

Piechowski, L. D. (2011). *Evaluation of workplace disability.* New York, NY: Oxford University Press. http://dx.doi.org/10.1093/med:psych/9780195341096.001.0001

Ruiz, M. A., Drake, E. B., Glass, A., Marcotte, D., & van Gorp, W. G. (2002). Trying to beat the system: Misuse of the Internet to assist in avoiding the detection of psychological symptom dissimulation. *Professional Psychology, Research and Practice, 33,* 294–299. http://dx.doi.org/10.1037/0735-7028.33.3.294

Schroeder, R. W., Twumasi-Ankrah, P., Baade, L. E., & Marshall, P. S. (2012). Reliable Digit Span: A systematic review and cross-validation study. *Assessment, 19,* 21–30. http://dx.doi.org/10.1177/1073191111428764

Tombaugh, T. N. (1996). *Test of Memory Malingering.* Toronto, Ontario, Canada: Multi-Health Systems.

Trueblood, W. (1994). Qualitative and quantitative characteristics of malingered and other invalid WAIS-R and clinical memory data. *Journal of Clinical and Experimental Neuropsychology, 16,* 597–607. http://dx.doi.org/10.1080/01688639408402671

Victor, T. L., & Abeles, N. (2004). Coaching Clients to Take Psychological and Neuropsychological Tests: A Clash of Ethical Obligations. *Professional Psychology: Research and Practice, 35,* 373–379. http://dx.doi.org/10.1037/0735-7028.35.4.373

Viljoen, J. L., & Zapf, P. A. (2002). Fitness to stand trial evaluations: A comparison of referred and non-referred defendants. *International Journal of Forensic Mental Health, 1,* 127–238. http://dx.doi.org/10.1080/14999013.2002.10471168

Wechsler, D. (2008). *Wechsler Adult Intelligence Scale* (4th ed.). San Antonio, TX: NCS Pearson, Inc.

Wechsler, D. (2009). *WMS-IV technical and interpretive manual.* San Antonio, TX: NCS Pearson, Inc.

Weinborn, M., Orr, T., Woods, S. P., Conover, E., & Feix, J. (2003). A validation of the test of memory malingering in a forensic psychiatric setting. *Journal of Clinical and Experimental Neuropsychology, 25,* 979–990. http://dx.doi.org/10.1076/jcen.25.7.979.16481

Youngjohn, J. R. (1995). Confirmed attorney coaching prior to neuropsychological evaluation. *Assessment, 2,* 279–283. http://dx.doi.org/10.1177/1073191195002003007

Received May 1, 2018
Revision received June 14, 2018
Accepted July 12, 2018 ■

JIDR Journal of Intellectual Disability Research
Published on behalf of mencap and in association with IASSID

Journal of Intellectual Disability Research doi: 10.1111/jir.12340

VOLUME 60 PART 12 pp 1178–1188 DECEMBER 2016

Development and validation of the Learning Disabilities Needs Assessment Tool (LDNAT), a HoNOS-based needs assessment tool for use with people with intellectual disability

J. Painter,[1] L. Trevithick,[2] R. P. Hastings,[3,4] B. Ingham[5,6] & A. Roy[7]

1 *Faculty of Health and Wellbeing, Sheffield Hallam University, Sheffield, UK*
2 *Durham and Darlington Mental Health Services for Older People, Tees, Esk and Wear Valley NHS Foundation Trust, Durham, UK*
3 *CEDAR (Centre for Educational Development, Appraisal and Research), University of Warwick, Coventry, UK*
4 *Centre for Developmental Psychiatry and Psychology, Monash University Australia, Melbourne, Australia*
5 *Psychological Services, Northumberland, Tyne and Wear NHS Foundation Trust, Newcastle upon Tyne, UK*
6 *School of Psychology, Newcastle University, Newcastle upon Tyne, UK*
7 *Solihull Community Services, Faculty of Intellectual Disability Psychiatry, Royal College of Psychiatrists, Birmingham, UK*

Abstract

Background In meeting the needs of individuals with intellectual disabilities (ID) who access health services, a brief, holistic assessment of need is useful. This study outlines the development and testing of the Learning Disabilities Needs Assessment Tool (LDNAT), a tool intended for this purpose.

Method An existing mental health (MH) tool was extended by a multidisciplinary group of ID practitioners. Additional scales were drafted to capture needs across six ID treatment domains that the group identified. LDNAT ratings were analysed for the following: item redundancy, relevance, construct validity and internal consistency ($n = 1692$); test–retest reliability ($n = 27$); and concurrent validity ($n = 160$).

Results All LDNAT scales were deemed clinically relevant with little redundancy apparent. Principal component analysis indicated three components (developmental needs, challenging behaviour, MH and well-being). Internal consistency was good (Cronbach alpha 0.80). Individual item test–retest reliability was substantial-near perfect for 20 scales and slight-fair for three scales. Overall reliability was near perfect (intra-class correlation = 0.91). There were significant associations with five of six condition-specific measures, i.e. the Waisman Activities of Daily Living Scale (general ability/disability), Threshold Assessment Grid (risk), Behaviour Problems Inventory for Individuals with Intellectual Disabilities-Short Form (challenging behaviour) Social Communication Questionnaire (autism) and a bespoke physical health questionnaire. Additionally, the statistically significant correlations between these tools and the LDNAT components made sense clinically. There were no statistically significant correlations with the Psychiatric Assessment Schedules for Adults with Developmental Disabilities (a measure of MH symptoms in people with ID).

Conclusions The LDNAT had clinically utility when rating the needs of people with ID prior to condition-specific assessment(s). Analyses of internal and

Correspondence: Mr Jon Painter, Sheffield Hallam University, 38, Collegiate Crescent, Sheffield, S10 2BP, UK. (e-mail: J.Painter@shu.ac.uk).

external validity were promising. Further evaluation of its sensitivity to changes in needs is now required.

Keywords autistic spectrum disorder, challenging behaviour, HoNOS, mental health, needs assessment, screening

Background

A purely diagnostic approach to the provision of care/treatment intended to assist people with intellectual disabilities (ID) to lead meaningful lives has been recognised as limited (Xenitidis *et al.* 2000; Thompson *et al.* 2004; Snell *et al.* 2009). In a quest for more responsive, individualised interventions and services, there has been a shift towards a need-led approach (Parmenter & Riches 2002). In the UK at least, this has led to the boundary between ID and mental health (MH) services at times becoming blurred with both sets of practitioners addressing needs that would traditionally have fallen within the other's remit, depending on the primary presenting issue.

Thompson *et al.* (2009) classify service support needs into four types:

- Normative/objective (i.e. a professional's comparison of an individual's assessed needs against a notional standard for those particular circumstances).
- Felt (i.e. the individual's perception of their own needs).
- Expressed (i.e. usually a request for help from the individual).
- Comparative (i.e. the difference between the support an individual receives and the norm for their peers).

Within ID and MH services, a thorough individual assessment is generally accepted to be the cornerstone of effective treatment or support with an objective assessment of needs seen as integral to this process (Gamble & Brennan 2000; Snell *et al.* 2009). Several conceptual frameworks exist to describe needs and disability that help to provide structure to the clinical assessment processes. In particular, the socioecological model (Institute of Medicine 1991) and the biopsychosocial model (Engel 1977) have gained popularity, arguably because of their holistic nature and ability to accommodate most professional, service and service users' perspectives. Condition-specific assessment tools exist for more discrete domains such as MH problems (e.g. Moss *et al.* 1998) and challenging behaviour (e.g. Rojahn *et al.* 2001). However, a broader tool that encourages both ID and MH practitioners to consider how they might address needs across all aspects of the biopsychosocial model (prior to focusing in on their own sub-specialty) has merit in ensuring an individual's full range of needs are consistently considered.

Few MH screening tools address the full range of issues typically associated with ID, but some ID needs assessments do include MH problems. Examples of broad ID needs assessment tools include the Camberwell Assessment of Need for Adults with Developmental and Intellectual Disabilities (CANDID, Xenitidis *et al.* 2000); the Supports Intensity Scale (SIS, Thompson *et al.* 2004); and the Instrument for the Classification and Assessment of Support Needs (I-CAN, Riches *et al.* 2009). Fundamentally, these tools seek to capture normative needs with some also eliciting felt need and expressed need (including those articulated by the individual's carers). However, as there is no accepted 'gold standard' needs assessment tool for people with ID (Xenitidis *et al.* 2000), these existing tools will each have limitations. For example, the SIS was designed to elicit the support needs of people with ID, but it consists of over 80 scales, thus posing a significant time and cost burden on routine practice. Also, concerns have been raised about the degree of subjectivity it involves (Riches *et al.* 2009). The I-CAN is shorter but still takes between 30–60 min to complete and is yet to be validated for use in people with ID whose primary need is MH-related (Riches *et al.* 2009). The CANDID was adapted from the original version (designed for use in mainstream psychiatry) and is described as a screening rather than a diagnostic tool. Despite this, the CANDID takes around 30 min to compete, and hence, concerns have again been raised about the feasibility of its use in routine practice (Xenitidis *et al.* 2000). Overall therefore, given the shift towards more holistic, needs-led service provision, the list of brief needs assessment tools that are potentially suitable for routine practice

in both ID and MH settings remains somewhat modest.

The purpose of the present study was to extend an existing MH needs assessment tool to create and test a new needs assessment tool, the Learning Disability Needs Assessment Tool (LDNAT) that was suitable for use in either service as a broad assessment of need in people with ID prior to more detailed, condition-specific assessments. The tool therefore needed to be brief whilst adequately capturing the full range of needs that people with ID typically present with when accessing health services.

Method

Participants

The needs of 2063 individuals were recorded by specialist ID professionals from a range of disciplines across six National Health Service (NHS) services in England. The subsequent analyses focus on the 1692 cases with all the required data scales. Of those, 992 (54.5%) were male and the mean age was 41.7 years (range 18–90 years). Treatment setting information was available for 1466 cases, of which 84 (5.7%) were assessed in inpatient settings. Most individuals assessed (94.4%) (1540 of 1631) were White British, and 493 of 1170 (42.1%) cases with data available were recorded as living in a form of supported accommodation (i.e. community placements with varying levels of paid staff input). People with ID were being assessed by health services for a variety of health needs. Of 686 cases for whom their 'primary need' (i.e. main reason health service involvement) was recorded, most frequent were challenging behaviour (180, 26.2%), mental illness (110, 16.0%), autism spectrum disorder (ASD) (62, 9.0%) and problems with mobility and posture (55, 8.0%).

Measures

The Learning Disability Needs Assessment Tool

An established MH tool – the Mental Health Clustering Tool (MHCT) (Self *et al.* 2008) was developed from the Health of the Nation Outcome Scales (HoNOS; Wing *et al.* 1998) to uniformly identify and rate individual need to assist with the costing, planning, commissioning and most importantly the delivery of appropriate MH treatment in the UK.

In a desire for more integrated MH and ID services that facilitated non-stigmatising, needs-led access (as advocated by Snell *et al.* 2009), a multidisciplinary group of 70 ID practitioners reviewed the MHCT in a workshop format for its applicability. Initially, they identified six broad treatment domains that they felt were important when planning support/treatment for people with ID (i.e. general ability/disability severity, risks, MH, challenging behaviour, ASD and physical health). By mapping each MHCT scale in turn to these domains, they were able to identify omissions. Through a similar consensus approach to that of Wing *et al.* (1998), these omissions were then translated into more specific, clearly defined descriptions of need (e.g. social communication and interaction difficulties, non-accidental self-injury associated with cognitive impairment, communication and problems with understanding). Finally, with reference to the HoNOS-LD (Roy *et al.* 2002), a 5-point scale for each need was developed by small working groups. A short description of the resulting LDNAT scales together with their origins is listed in Table 1.

The first 24-item version of the tool (including both expressive and receptive communication scales) was piloted with more than 2000 individuals by qualified ID staff from 18 NHS organisations with qualitative and quantitative data used to refine the new LDNAT scales. Clinical feedback suggested broad consensual content validity however, the expressive and receptive communication scales, were highly correlated and hence were collapsed into a single item. Following this and other minor refinements, the LDNAT was judged to capture the full range of needs from the six domains identified at the original multidisciplinary workshop and be a seamless extension to the MH version of the tool.

Measures used to validate the Learning Disability Needs Assessment Tool

For each of the six treatment domains, candidate condition-specific measures were identified from a brief literature review. The final choice for each domain was taken by a small multidisciplinary group of ID practitioners and was based on criteria including brevity, simplicity, psychometric quality

Table 1 Item titles, derivation and summary scoring statistics for LDNAT (n = 1692)

Item and title		Derivation	Mean score	Standard deviation
1	Overactive, aggressive, disruptive or agitated behaviour	Original HoNOS scale	1.12	1.09
2	Non-accidental self-injury	Original HoNOS scale	0.30	0.72
3	Problem drinking or drug taking	Original HoNOS scale	0.11	0.46
4	Cognitive problems	Original HoNOS scale	2.00	1.10
5	Physical illness or disability problems	Original HoNOS scale	1.38	1.29
6	Hallucinations or delusions	Original HoNOS scale	0.21	0.64
7	Depressed mood	Original HoNOS scale	0.63	0.86
8	Other mental and behavioural problems (choose from: A phobic; B anxiety; C obsessive–compulsive; D mental strain/tension; E dissociative; F somatoform; G eating; H sleep; I sexual; J other).	Original HoNOS scale	1.24	1.25
9	Relationships	Original HoNOS scale	1.18	1.14
10	Activities of daily living	Original HoNOS scale	1.75	1.28
11	Living conditions	Original HoNOS scale	0.46	0.83
12	Occupation and activities	Original HoNOS scale	0.80	1.02
13	Strong unreasonable beliefs	Original HoNOS scale	0.32	0.83
14	Non-accidental self-injury (associated with cognitive impairment)	Original MHCT scale	0.33	0.75
15	Physical problems with eating and drinking	Adapted from HoNOS-LD	0.47	0.98
16	Agitated behaviour/expansive mood	Adapted from HoNOS-LD	1.82	1.35
17	Repeat self-harm	Original MHCT scale	0.62	1.04
18	Safeguarding other children and vulnerable dependent adults	Original MHCT scale	0.62	1.13
19	Engagement	Original MHCT scale	1.23	1.33
20	Vulnerability	Original MHCT scale	2.40	1.15
21	Social communication difficulties	Original MHCT scale	1.51	1.35
22	Communication problems*	New LDNAT item	1.55	1.31
23	Seizures	Adapted from HoNOS-LD	0.52	1.02
LDNAT total		Adapted from HoNOS-LD	0.98	1.24

*Initially piloted as two separate (expressive and receptive) communication items, but combined because of redundancy.
HoNOS, Health of the Nation Outcome Scales; HoNOS-LD, Health of the Nation Outcome Scales for people with Learning Disabilities; LDNAT, Learning Disabilities Needs Assessment Tool; MHCT, Mental Health Clustering Tool.

and cost. The first two of these criteria were deemed particularly important given the tools would be completed independently by informants who would receive no training. The final list was as follows.

The Waisman Activities of Daily Living Scale (Maenner *et al.* 2013) was used to assess general ability/disability. Raters record whether an individual can complete various activities of daily living independently (score 2), with help (score 1) or not at all (score 0). The tool consists of 17 activities ranging from basic skills (e.g. drinking from a cup) to more advanced tasks (e.g. simple home repairs and budgeting). The tool was validated on people with a broad range of ID diagnoses including autism and Fragile-X syndrome.

The Threshold Assessment Grid (TAG) (Slade *et al.* 2000) was selected to provide an overall risk rating. It was originally developed and validated through a series of workshops and a Delphi consultation as a means of prioritising access to mainstream MH services. Seven scales are each rated on 4-point or 5-point scales to give an overall rating of illness severity. However, the first four scales (i.e. intentional self-harm, unintentional self-harm, risk from others and risk to others) were deemed by

clinicians to adequately capture risks to/from people with ID.

The Psychiatric Assessment Schedules for Adults with Developmental Disabilities checklist (PAS-ADD checklist) (Moss *et al.* 1998) was used to rate the severity of MH problems. The tool consists of 24 scales written using lay terms to allow non-professionals to identify MH problems in people with ID. Originally developed as a screening tool, it includes three different scoring triggers for a fuller MH assessment. Scales include irritability, loss of appetite and strange unshakeable beliefs. Scales are rated on a 4-point scale, which combines intensity and frequency, and is based on the previous 4 weeks but specifically excludes long-standing issues.

The Behaviour Problems Inventory for Individuals with Intellectual Disabilities-Short Form (Rojahn *et al.* 2012a,2012b; Mascitelli *et al.* 2015) was selected to rate challenging behaviours. This shortened version captures self-injurious behaviours (e.g. head hitting), aggressive/destructive behaviours (e.g. verbal aggression) and stereotyped behaviours (e.g. rocking/repetitive body movements) and is based on a longer (52-item) original version. The frequency rating for each of the 30 scales was used to provide an overall total score.

The Social Communication Questionnaire, (Rutter *et al.* 2003) was selected to provide a rating of the severity of ASD symptoms. Valid for both children and adults (Brooks & Benson 2013), it consists of 40 'yes/no' questions intended to capture the key features of ASD for example: 'Does he/she have interests that pre-occupy him/her and might seem odd to other people (e.g. traffic lights, drainpipes or timetables)?'

No single suitable physical health measure could be identified, and so, a bespoke questionnaire was created by the authors (available on request). It consisted of 12 yes/no questions (e.g. 'Is the person blind/visually impaired?'), three rating scale questions (e.g. 'How good is the person's health in general? Very good/ good/ fair/ bad/ very bad/ don't know') and two that ask for height and weight. The yes/no and scaled questions were used to create a total score representing the overall level of physical disability. Although yet to be fully validated, it was based on the POMONA study (Haverman *et al.* 2011), and a brief investigation of its internal consistency yielded acceptable results in the present sample (Cronbach alpha = 0.73).

Procedure

Six NHS services in England used the LDNAT between 01/07/2014 and 31/08/2015 to systematically record the needs of their users following routine assessment. Qualified staff from a range of disciplines attended a one-day training event before cascading this information to staff in their own organisations. LDNAT ratings were then recorded as part of their routine assessments. Participating NHS services sourced the data required for the study from their case records before submission via a standardised, encrypted data set to the lead organisation for collation and central analysis. The study received NHS approval for the purposes of an NHS service evaluation project.

A subset of these services was able to consider the nature of their users in greater detail. For each routine referral to these services, the LDNAT assessor was contacted to identify an independent rater who knew the person well enough to complete the six additional assessment measures. Typically, this was the referrer, the GP or a family/carer. These individuals were contacted by telephone, and if they were willing to provide this more detailed level of referral information, the six validation questionnaires were posted out for completion within 2 weeks. This exercise resulted in 160 cases that had the six independently rated questionnaires in addition to their LDNAT ratings. This convenience sample did not differ significantly from the full data set in terms of their demographics other than having a higher proportion of people assessed in inpatient settings (21% vs. 5.7% in the main data set). These additional ratings were then included in the electronic data submission for analysis.

Data analysis

Statistical validation of the LDNAT involved several different analyses. Potentially redundant scales were assessed using correlations between scales. Scales with limited application to the population were assessed with reference to the percentage of cases scoring zero (i.e. scoring as having no problems on the item). Construct validity was addressed through principal components analysis (PCA) to investigate the structure of the LDNAT. Internal consistency was assessed using Cronbach's alpha. Test–retest reliability was assessed using the records from individuals who, (typically because of a transfer between teams) had second LDNAT assessments

completed within 30 days of the first (total $n = 27$). Finally, concurrent validity was assessed by examining associations between the six additional questionnaires and the LDNAT scores.

Results

Once duplicate entries, repeat assessments and assessments of children were removed from the data set, mean scores for each item were calculated. Table 1 shows summary statistics for each item and the LDNAT total score. The range for all scales was identical (0–4), indicating that at least one person had been rated as having 'no needs' and one as having a 'severe need' in each area of the LDNAT.

Item redundancy

Correlations were examined to assess for possible redundant scales. Because of their closely related clinical interpretations, close attention was paid to scales 2, 14 and 17 (differing forms of self-injurious behaviour). There were weak correlations between items 2 and 14 ($r(1952) = 0.370$, $P = <0.001$); items 14 and 17 ($r(1944) = 0.353$, $P < 0.001$); and a weak-moderate correlation between items 2 and 17 ($r = (1948) = 0.452$, $P < 0.001$). Given that the shared variance represented by these correlation values was low, each item was retained because of its potential clinical utility. In general, raters were using these scales differently (to assess different needs).

Item relevance

Items with potentially limited application for people accessing ID services were assessed by identifying items with a high incidence of 'zero' scores (indicating no problem). Items with high percentages of 'zero' scores were items 2 ($n = 1395$, 82.4%), 3 ($n = 1578$, 93.3%), 6 ($n = 1494$, 88.3%) and 13 ($n = 1430$, 84.5%). The mean frequency of 'zeros' for the remaining scales was 46.1%. All of these scales were retained in the LDNAT for clinical assessment purposes and were initially retained also for PCA.

Principal components analysis

A PCA was conducted on the 21 scales using orthogonal rotation (varimax). Scales 3 and 6 were excluded from this analysis as earlier exploratory tests revealed they did not significantly load onto any component, and it had already been established that these two scales had high levels of zero scores in the population. The Kaiser–Meyer–Olkin (KMO) measure verified the overall sampling adequacy for the analysis, KMO = 0.83, and all KMO values for individual scales were above 0.73.

An initial analysis (PCA with varimax rotation) was performed to obtain eigenvalues for each component in the data. Five components had eigenvalues over Kaiser's criterion of 1 and together explained 54.7% of total variance. The point of inflexion on the scree plot indicated the retention of three components that together explained 43.0% of total variance. Table 2 shows the component loadings after rotation as well as Cronbach's alpha values for each of the three potential LDNAT sub-scale scores identified through the PCA.

Learning Disability Needs Assessment Tool internal consistency

The internal consistency of the LDNAT total score was evaluated using Cronbach's alpha. The alpha value was 0.80 indicating good internal reliability of the tool, according to Nunally's rule of thumb (Nunnally & Bernstein 1994). Corrected item-total correlations were examined with only one very low value for item 3 – 'alcohol and drug problems' $r = 0.05$. However, removing this item did not improve the internal consistency of the total score, so this item was retained for its potential clinical value.

Test–retest reliability

Test–retest reliability was assessed by calculating intra-class correlation (ICC) coefficients. Individuals were included in this analysis if they had two LDNATs completed within 30 days of each other. Table 3 shows the ICCs and confidence intervals for individual LDNAT scales and for each independent component derived from the PCA. Using Landis and Koch's (1977) thresholds, these values suggest substantial or near perfect agreement for all but three scales (hallucination/delusions, living conditions and self-injurious behaviour) that show slight or fair agreement over time.

Table 2 Summary of PCA varimax rotated component loadings. $N = 1692$ Component loadings above 0.40 appear in bold

	LDNAT item	Component 1 Developmental Needs	Component 2 Challenging Behaviour	Component 3 Mental Health and well-being
22	Communication problems	**0.80**	0.20	−0.12
10	Activities of daily living	**0.77**	−0.05	0.22
4	Cognitive problems	**0.70**	0.13	−0.09
15	Physical problems with eating and drinking	**0.60**	−0.24	0.09
5	Physical illness or disability problems	**0.60**	−0.32	0.10
20	Vulnerability	**0.56**	0.30	0.06
21	Social communication difficulties	**0.56**	0.40	0.06
23	Seizures	**0.41**	0.03	−0.09
16	Agitated behaviour/expansive mood	0.01	**0.79**	0.05
1	Overactive, aggressive, disruptive or agitated behaviour	0.15	**0.69**	0.27
17	Repeat self-harm	−0.04	**0.58**	0.07
19	Engagement	0.14	**0.51**	0.23
18	Safeguarding other children and vulnerable dependent adults	−0.04	**0.49**	0.11
14	Non-accidental self-injury (associated with cognitive impairment)	0.35	**0.45**	0.02
8	Other mental and behavioural problems	0.05	**0.45**	**0.40**
2	Non-accidental self-injury	−0.03	**0.44**	0.33
12	Occupation and activities	0.11	−0.01	**0.68**
7	Depressed mood	−0.16	0.11	**0.62**
9	Relationships	0.13	0.35	**0.61**
11	Living conditions	0.07	0.11	**0.56**
13	Strong unreasonable beliefs	−0.10	0.25	**0.52**
	Eigenvalues	3.47	3.21	2.35
	% of variance	16.51	15.33	11.2
	Cronbach's alpha value	0.79	0.76	0.68

LDNAT, Learning Disabilities Needs Assessment Tool; PCA, principal components analysis.

Learning Disability Needs Assessment Tool concurrent validity

There were 160 cases from three NHS services who had scores for the six additional measures recorded by independent assessors. Correlations between the sub-scales and totals for these six measures and the three LDNAT component scores and the total LDNAT score were investigated. These results are summarised in Table 4.

The Waisman Activities of Daily Living Scale total (used as a measure of general disability) was negatively correlated with LDNAT total score and also the developmental needs component. The TAG total score (a measure of risk) was significantly associated with the LDNAT total as were the three TAG sub-scales. More specifically, the TAG needs and disabilities sub-scale was associated with the LDNAT developmental needs component, whilst the TAG safety and risk sub-scales both had significant correlations with the LDNAT challenging behaviour component. There were no statistically significant associations between any of the PAS-ADD checklist sub-scales and the LDNAT total score or the LDNAT component totals. The LDNAT's relationship with the Behaviour Problems Inventory for Individuals with Intellectual Disabilities-Short Form (the selected independent measure of challenging behaviour) was as expected, with the total frequency score associated with the LDNAT total score. They were also strongly correlated with the LDNAT challenging behaviour component. The

Table 3 Intra-class correlation coefficients and 95% confidence intervals for cases (n-27) with two completed LDNAT assessments within 30 days

Item	ICC	95%CI
Overactive, aggressive, disruptive or agitated behaviour	0.75	0.52, 0.88
Non-accidental self-injury	0.61	0.32, 0.80
Problem drinking or drug taking	0.95	0.89, 0.98
Cognitive problems	0.85	0.70, 0.93
Physical illness or disability problems	0.89	0.77, 0.95
Hallucinations or delusions	0.00	−0.37, 0.37
Depressed mood	0.84	0.65, 0.93
Other mental and behavioural problems	0.68	0.41, 0.84
Relationships	0.82	0.65, 0.91
Activities of daily living	0.87	0.73, 0.94
Living conditions	0.17	−0.22, 0.52
Occupation and activities	0.78	0.57, 0.89
Strong unreasonable beliefs	0.72	0.48, 0.86
Non-accidental self-injury (associated with cognitive impairment)	0.33	−0.07, 0.63
Physical problems with eating and drinking	0.72	0.47, 0.86
Agitated behaviour/expansive mood	0.96	0.91, 0.98
Repeat Self-Harm	0.92	0.83, 0.96
Safeguarding other children and vulnerable dependent adults	0.89	0.78, 0.95
Engagement	0.81	0.62, 0.91
Vulnerability	0.79	0.60, 0.90
Social communication difficulties	0.94	0.87, 0.97
Communication problems	0.98	0.95, 0.99
Seizures	0.74	0.51, 0.88
Developmental needs component	0.95	0.89, 0.98
Challenging behaviour component	0.93	0.85, 0.97
Mental health and well-being component	0.88	0.76, 0.94
LDNAT total	0.91	0.82, 0.96

CI, confidence interval; ICC, intra-class correlation; LDNAT, Learning Disabilities Needs Assessment Tool.

Social Communication Questionnaire-total score (used to measure autism symptoms) was strongly correlated with the LDNAT total, the developmental needs component and to a lesser but still significant extent to the challenging behaviour component. Finally, the total score for the physical health tool (calculated from 15/17 questions) had a strong association with the LDNAT developmental needs component.

Discussion

This study outlines the development and validation of a needs assessment tool (the LDNAT) for use in ID health services as a precursor to more detailed, condition-specific assessment. Whilst the catalyst for its development stemmed from a move away from traditional commissioning arrangements, it is important to stress that, as with similar MH developments, this has been a clinically led project. The primary aim for these project staff was to encourage a holistic assessment of the needs of people with ID regardless of whether they enter a specialist MH or ID service.

The resulting 23-item LDNAT showed good overall internal consistency for the total score and three potential sub-scale scores. Corrected item-total correlations were all acceptable with the exception of scale 3 (drug and alcohol problems) that was retained because of its clinical value. The original MH tool already contained two scales concerned with differing types of self-injury; however, neither was felt to adequately capture the self-injurious behaviour traditionally associated with cognitive impairment, and hence, a third item was created. Analysis confirmed the shared variance between these three scales to be small, confirming that they were being used to capture different clinical constructs. In contrast, redundancy between two new scales concerning communication was identified during piloting, resulting in their replacement by a single, over-arching communication item.

Principal component analysis using a component loading threshold of 0.4 identified three components: developmental needs, challenging behaviour and MH and well-being. At this loading threshold, item 8 (a pick list item of 'other mental and behavioural problems') loaded onto the second and third components (presumably because of the plurality of mental and behavioural issues it was designed to capture). The item related to seizures did not load significantly onto any component. It must be stressed that these three components are merely a preliminary exploration of the LDNAT's structure used to facilitate validation. Further research (particularly of the heterogeneous MH and well-being) component is required before they could be used clinically.

Intra-class correlation coefficients indicated a high level of test–retest reliability for scales, total LDNAT scores and the three sub-scale/component scores. ICCs for only three scales fell below an acceptable level in terms of test–retest reliability.

Table 4 Correlations between LDNAT component and total scores with sub-scale and total scores for the six independently rated questionnaires

LDNAT component/ total score		W-ADL	TAG			PAS-ADD			BPI	SCQ	
		Total	Total score	Safety sub-scale	Risk sub-scale	Needs and disabilities	Organic	Affective and neurotic	Psychotic	Total frequency score	Total	Physical health Total
Developmental needs component	Pearson correlation	−0.668**	0.258**	0.206*	0.088	0.277**	0.067	−0.132	−0.01	0.260**	0.401**	0.497**
	N	150	147	155	153	154	149	139	147	98	141	104
Challenging behaviours component	Pearson correlation	0.02	0.532**	0.457**	0.446**	0.446**	−0.058	−0.059	0.131	0.472**	0.198*	−0.196*
	N	150	147	155	153	154	149	139	147	98	141	104
Mental health and well-being component	Pearson correlation	−0.011	0.412**	0.296**	0.324**	0.420**	−0.073	0.019	0.042	0.224*	0.152	−0.201*
	N	151	148	156	154	155	150	140	148	98	142	105
LDNAT total score	Pearson correlation	−0.298**	0.531**	0.431**	0.392**	0.495**	−0.006	−0.078	0.093	0.437**	0.342**	0.084
	N	150	147	155	153	154	149	139	147	98	141	104

*$P < 0.05$
**$P < 0.01$
BPI, Behaviour Problems Inventory; LDNAT, Learning Disabilities Needs Assessment Tool; PAS-ADD, Psychiatric Assessment Schedules for Adults with Developmental Disabilities; SCQ, Social Communication Questionnaire; TAG, Threshold Assessment Grid; W-ADL, Waisman Activities of Daily Living Scale.

Finally, convergent validity was demonstrated by comparing the tool's performance to that of more specific measures for each of the six needs domains identified by professionals as important in care/treatment planning. The statistical associations with five of the six validation measures all had clinical face validity. However, the LDNAT had no statistically significant associations with the PAS-ADD checklist (a screening tool for MH problems in people with ID). One possible reason for this is the heterogeneous nature of the LDNAT's MH and well-being component. Alternatively, the PAS-ADD checklist's scoring guidance specifically excludes issues that have 'always been like this', whilst eight of the 23 LDNAT scales addressing MH problems encourage raters to consider historical issues and behaviours that remain relevant to the current plan of care. It is then possible that individuals had long-standing MH problems rather than new episodes or acute deteriorations that the PAS-ADD checklist was not designed to capture. Finally, data suggest that the PAS-ADD checklist may have some psychometric weaknesses (as muted by Hatton & Taylor 2008).

Although the reliability and validity data for the LDNAT are encouraging, there are a number of limitations with this study that should be borne in mind when interpreting the results.

The naturalistic nature of the study has both strengths and weaknesses. Training was cascaded, albeit with a standardised structure, which may have led to variable accuracy of ratings. Conversely, however, this approach did not foster any unrealistic expectations, i.e. that already stretched services can create and sustain the burden of gold-standard research conditions in routine practice. Also, the sample was neither randomly selected nor stratified (e.g. on the basis of IQ) and hence may or may not be representative of the entire population of people with ID accessing health services.

Second, the inability to identify the names of raters from the much smaller data set ($n = 27$) used for test–retest analysis means that this analysis was in effect a combination of inter-rater reliability and stability over time. Anecdotally, however, these second ratings were usually a consequence of movement between services and hence mainly undertaken by a different practitioner, making the ICCs all the more encouraging.

The choice of additional measures was heavily influenced by their suitability for untrained informant completion. Whilst independence from LDNAT raters was seen as crucial, as with the study of Xenitidis *et al.* (2000), this inevitably led to trade-offs between the utility of tools and their validity. For example, the PAS-ADD was selected as an appropriate measure selected despite known potential psychometric weaknesses (Hatton & Taylor 2008). Equally, the creation and use of a largely untested physical health assessment was deemed more favourable than using any of the alternative tools identified.

Overall therefore, in the context of ever-blurring service boundaries, the LDNAT has the potential to support the brief but holistic assessment of a wide range of needs associated with both MH and ID by staff in either setting. The tool now needs to be subjected to further testing with data from new samples. Also, evaluation of its potential to monitor changes in need in the same way that MHCT data can monitor outcomes in MH services (Speak *et al.* 2015) is required. In this way, a second application of the tool could serve to reduce the perceived burden its use places on practitioners and services through adding value to service functions.

Conflict of Interest

The authors report no conflicts of interest.

References

Brooks W. T. & Benson B. A. (2013) The validity of the social communication questionnaire in adults with intellectual disability. *Research in Autism Spectrum Disorders* **7**(2), 247–55. DOI: 10.1016/j.rasd.2012.10.002.

Engel G. (1977) The need for a new medical model: a challenge for biomedicine. *Science* **196**, 129–96.

Gamble C. & Brennan G. (2000) Assessment: a rationale and glossary of tools. In: *Working with Serious Mental Illness: A Manual for Clinical Practice* (eds C. Gamble & G. Brennan), pp. 67–84. Balliere Tindall, London.

Hatton C. & Taylor J. L. (2008) Factor structure of the PAS-ADD checklist with adults with intellectual disabilities. *Journal of Intellectual and Developmental Disability* **33**, 330–6.

Haverman M., Perry J., Salvador-Carulla L., Walsh P., Kerr M., Van Schrojenstein Lantman-de Valk H. *et al.* (2011) Ageing and health status in adults with intellectual

disabilities: results of the European POMONA II study. *Journal of Intellectual Disability* **3**, 49–60.

Institute of Medicine (1991) A model for disability and disability prevention. In: *Disability in America: Toward a National Agenda for Prevention*, pp. 76–108. Available at: http://www.nap.edu/read/1579/chapter/5#80. The National Academies Press, Washington, DC.

Landis J. R. & Koch G. G. (1977) The measurement of observer agreement for categorical data. *Biometrics* **33**, 159–74.

Maenner M. J., Smith L. E., Hong J., Makuch R., Greenberg J. S. & Mailick M. R. (2013) Evaluation of an activities of daily living scale for adolescents and adults with developmental disabilities. *Disability and Health Journal* **6**, 8–17. Available at: http://www.sciencedirect.com/science/article/pii/S1936657412001185.

Mascitelli A. N., Rojahn J., Moore L., Hastings R. P. & Christian-Jones C. (2015) The behaviour problems inventory-short form: reliability and factorial validity in adults with intellectual disabilities. *Journal of applied research in intellectual disabilities : JARID* **28**, 561–71. Available at: http://www.ncbi.nlm.nih.gov/pubmed/25758786 [Accessed February 24, 2016].

Moss S., Prosser H., Costello H., Simpson N., Patel P., Rowe S. *et al.* (1998) Reliability and validity of the PAS-ADD checklist for detecting psychiatric disorders in adults with intellectual disability. *Journal of Intellectual Disability Research* **42**, 173–83.

Nunnally J. C. & Bernstein I. H. (1994) *Psychometric Theory*, 3rd edn. McGraw-Hill, New York.

Parmenter T. R. & Riches V. C. (2002) Pathways 6 Conference 2002. In *Assessment and Classification of Support Needs*. Centre for Developmental Disability Studies, Sydney University.

Riches V. C., Parmenter T. R., Llewellyn G., Hindmarsh G. & Chan J. (2009) I-CAN: a new instrument to classify support needs for people with disability: Part I. *Journal of Applied Research in Intellectual Disabilities* **22**, 326–39. Available at: http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2009-09049-002&site=ehost-live\nvriches@med.usyd.edu.au.

Rojahn J., Matson J. L., Lott D., Esbensen A. J. & Smalls Y. (2001) The behavior problems inventory: an instrument for the assessment of self-injury, stereotyped behavior, and aggression/destruction in individuals with developmental disabilities. *Journal of Autism and Developmental Disorders* **31**, 577–88.

Rojahn J., Rowe E. W., Sharber A. C., Hastings R. P., Matson J. L., Didden R. *et al.* (2012a) The Behavior Problems Inventory-Short Form for individuals with intellectual disabilities: Part I: development and provisional clinical reference data. *Journal of Intellectual Disability Research* **56**, 527–45.

Rojahn J., Rowe E. W., Sharber A. C., Hastings R. P., Matson J. L., Didden R. *et al.* (2012b) The Behavior Problems Inventory-Short Form for individuals with intellectual disabilities: Part II: Reliability and validity. *Journal of Intellectual Disability Research* **56**, 546–5.

Roy A., Matthews H., Clifford P., Fowler V. & Martin D. (2002) Health of the Nation Outcome Scales for People with learning disabilities (HoNOS–LD). *The British Journal of Psychiatry* **180**, 61–6.

Rutter M., Bailey A. & Lord C. (2003) The Social Communication Questionnaire: Manual.

Self R., Rigby A., Leggett C. & Paxtorn (2008) Clinical Decision Support Tool: a rational needs-based approach to making clinical decisions. *Journal of Mental Health* **17**, 33–48.

Slade M., Powell R., Rosen A. & Strathdee G. (2000) Threshold Assessment Grid (TAG): the development of a valid and brief scale to assess the severity of mental illness. *Social Psychiatry and Psychiatric Epidemiology* **35**, 78–85.

Snell M., Luckasson R., Borthwick-Duffy S., Bradley V., Buntinx W. H. E., Coulter D. L. *et al.* (2009) Characteristics and needs of people with intellectual disability who have higher IQs. *Intellectual and Developmental Disabilities* **47**, 220–33.

Speak B., Hay P. & Muncer S. (2015) HoNOS – their utility for payment by results in mental health. *International Journal of Health Care Quality Assurance* **28**, 115–128.

Thompson J., Bradley V. J., Buntinx W. H. E., Schalock R. L., Shogren K. A., Snell M. E. *et al.* (2009) Conceptualising supports and the support needs of people with intellectual disability. *Intellectual and Developmental Disabilities* **47**, 135–46.

Thompson J. R. Bryant B., Campbell E. M., Craig E. M., Hughes C., Rotholtz D. A., *et al.* (2004) *Supports Intensity Scale*. American Association of Mental Retardation, Washington.

Wing J., Beevor A. S., Curtis R. H., Park S. B. G., Hadden S. & Burns A. (1998) Health of the Nation Outcome Scales (HoNOS). *British Journal of Psychiatry* **172**, 11–18.

Xenitidis K., Thornicroft G., Leese M., Slade M., Fotiadou M., Philp H. *et al.* (2000) Reliability and validity of the CANDID-a needs assessment instrument for adults with learning disabilities and mental health problems. *The British Journal of Psychiatry* **176**, 473–8.

Accepted 25 August 2016

Clinical Psychologist, Vol. 13, No. 3, November 2009, pp. 102–110.

The Australian Psychological Society Ltd

Analysis of suicidal behaviour in Israeli veterans and terror victims with post-traumatic stress disorder by using the computerised Gottschalk–Gleser scales

SHARON GALOR[1] & UWE HENTSCHEL[2]

[1]*Department of Psychology, Innsbruck University, Innsbruck, Austria and* [2]*Department of Psychology, Leiden University, Leiden, The Netherlands*

Abstract

The primary objective of this study was to identify the vulnerability factors for suicide attempts in an Israeli sample, with the help of the Gottschalk–Gleser content analysis scales. The respondents were divided into four groups: suicide attempters; controls; post-traumatic stress disorder and depressed patients who did not report suicidal behaviour; and suicide ideators. The significant results represent conscious and unconscious psychological states, which suicide attempters have in common and can be seen as potential suicide risk factors. The main recurring risk-related themes are hopelessness, sickness, deterrents, frustrated dependency strivings, total anxiety and total depression.

Keywords: *death, dying, Gottschalk–Gleser scales, measurement/individual differences, population issues, post-traumatic stress disorder, psychological disorders, psychological testing, stress, suicide risk factors*

The aim of the present study was to examine suicidal behaviour among Israeli veterans and terror victims with post-traumatic stress disorder (PTSD) and depression. Every year 400 Israeli subjects commit suicide and each year there are approximately 4000 unsuccessful attempts (Neizer, 2008). According to the Ministry of Health (2005), every 10–14 days one soldier in active duty commits suicide.

According to the *Diagnostic and statistical manual of mental disorders* (*DSM-IV*) (American Psychiatric Association, 2000), the main symptoms of PTSD include intrusion, avoidance and hyper-arousal. Individuals with PTSD often also suffer from depression (Southwick, Yehuda, & Giller, 1991). Depression is mainly characterised by helplessness, hopelessness, sadness and unworthiness (*DSM-IV*). Depression makes the emptiness of one's life become self-evident. Life becomes meaningless and filled with despair.

Unbearable psychological pain, frustrated psychological needs, anger, hopelessness and helplessness were linked by Shneidman (1985) to suicide. According to the study by Barak and Miron (2005), Israeli suicidal individuals were more self-focused, revealed cognitive constrictions and reported more psychological pain than non-suicidal subjects.

Re-experiencing symptoms of PTSD was named as a predictor of suicidal ideations among Vietnam combat veterans with chronic PTSD (Bell & Nye, 2007). PTSD patients reported more anger, impulsivity, lower social support, psychological distress and hopelessness, which increase the risk of suicide attempts (Bell & Nye, 2007).

Israeli combat soldiers who committed suicide had a higher sense of duty, excessive motivation to excel, fewer referrals to psychological evaluations and scored higher on autonomy than the non-suicidal group (Bodner, Ben-Artzi, & Kaplan, 2006). Suicide among non-combat soldiers was linked in the same study to personality weakness.

Terrorism has also been linked to suicidal behaviour. Shortly after the occurrence of a terrorist attack, suicide rates in the general population decline. Allegedly, the shocking reports of the death of innocents (Salib, 2003) and the national grief influence suicidal behaviour (Hawton & van Heeringen, 2000).

Correspondence: Uwe Hentschel, University of Leiden, Department of Psychology, Wassenaarseweg 52, 2333 AK Leiden, The Netherlands.
E-mail: hentsche@fsw.leidenuniv.nl

ISSN 1328-4207 print/ISSN 1742-9552 online
Published by Taylor & Francis
DOI: 10.1080/13284200903353072

The following aspects will portray for the present study what sets the Israeli respondents apart from others. Besides the compulsory army service, Israeli citizens have been exposed to several wars and to continuous terrorist attacks. The extensive media coverage of each terrorist attack exposes every viewer indirectly to the horrifying sights of physical as well as psychological traumata. Stressful life events that the Israeli veterans experience throughout life, contribute significantly to PTSD symptomatology, even more than the combat exposure (Solomon, Zur-Naah, Horsh, Zerach, & Keinan, 2008).

The negative stereotypes that still persist in Israel often lead to shame and unwillingness to seek psychological help (Tal, Roe, & Corrigan, 2007). Self-disclosure is imperative for the clinician. Low self-disclosure limits the ability to diagnose and to apply preventive strategies. Self-disclosure was found to be low among suicidal Israeli adolescents (Horesh & Apter, 2006). Low self-disclosure can predict the lethality and seriousness of the suicide attempt (Apter, Horesh, Gothelf, Graffi, & Lepkifker, 2001). Low self-disclosure is a more important risk factor for near lethal suicide attempts than psychiatric illness and the presence of mental pain (Levi et al., 2008).

This study will use content analysis to identify the psychological states that may increase suicidal risk and which are not necessarily also explicitly reported by the patient. Content analysis is based on the hypothesis that words that people choose can contain information about their psychological states (Viney, 1983). Analysing verbal communication can help to interpret the way a person experiences him/herself and the surroundings on conscious and unconscious levels. The use of content analysis thus unveils important information, which otherwise may not have been revealed.

The Gottschalk–Gleser scales (GG scales) (Gottschalk & Gleser, 1969) measure a variety of psychological dimensions from the language naturally used by the respondent. The computerisation of the scales (Psychiatric Content Analysis and Diagnosis [PCAD] 2000) (Gottschalk & Bechtel, 1998–2003) enables the interviewer to examine, almost effortlessly, 33 scales in a very short time. The main scales in the PCAD 2000 are anxiety, hostility, personal disorganisation or schizophrenia, depression, cognitive and intellectual impairment, hope, hopelessness, human relations, self-accusation, somatic concerns, achievement strivings, psychomotor retardation, support, dependency strivings and frustrated dependency strivings, deterrents, health–sickness and quality of life (Gottschalk & Bechtel, 1982, 1998–2003).

Suicide researchers have yet to discover precise diagnostic criteria that can adequately detect high-risk suicidal individuals at an early stage. The aim of this study was therefore to find a solution to the problem of low self-disclosure and identify potential suicide risk factors with the help of the GG scales. This will be done by comparing the distinctions between the following groups: (a) the control group and suicide attempters (this comparison provides us with an overview of all probable suicide risk factors; the detection of these risk factors in individuals from the general population should be perceived as anomalous and alarming; (b) individuals with PTSD and depression without suicidal behaviour and suicide attempters (a comparison between suicidal and non-suicidal PTSD and depressed individuals enables us to distinguish potential suicide risk factors, the presence of which enhances the risk of a suicide attempt in individuals with this psychopathology); and (c) suicide ideators and suicide attempters (this comparison enables us to determine the potential risk factors that enhance the risk of an escalation of suicidal ideations towards a suicide attempt).

According to the hypothesis, the suicide attempters will have higher scores on the anxiety, depression, hopelessness and deterrents scales in comparison to the other groups. It is hypothesised that the GG scales will identify and exhibit the psychological states that enhance the risk of a suicide attempt. These results will thus help to improve the clinician's diagnosis and to implement preventive measures on time.

Method

Sample

All respondents were Israeli subjects who had mastered the Hebrew language and who had lived at least 5 years in Israel. It should be mentioned that the Israeli population is essentially a melting pot of many cultures. The collective Israeli identity is firmly shaped by the compulsory army service and living with the daily threats. One could expect that after 5 years of integrating with the Israeli society, speaking the language and being a veteran, the Israeli mentality becomes strongly present in the immigrant. It is impossible to control for cultural differences within the Israeli population. The respondents represent the sociodemographic profile in Israel.

According to the Ministry of Health (2008), the rates of suicide attempts in 2001–2004 among Ethiopian immigrants aged 15–44 years was higher than among immigrants from the former Soviet Union. The suicide rate among the two groups is similar for the age range 45–65 years. Apter et al. (2008) found no difference in suicidal behaviour

regarding ethnicity and immigrant status. One should take into consideration the following additional arguments: no Ethiopians immigrants took part in this research; furthermore, acculturation stress, which might lead in extreme cases to psychiatric disorders, decreases after the third year of immigration (Tartakovsky, 2007), and our criterion was a minimum of 5 years.

All respondents had a minimum of 9 years of schooling. The age of the respondents ranged between 25 and 51 years. All experienced combat and/or terrorism. No sexual and/or physical abuse or rape victims participated in this study. The traumatic events were of a collective nature (i.e., missiles attacks, witnessing atrocities, etc.). The traumatic background and negative life events were controlled by reading the dossiers and a life events questionnaire. The questionnaire included general life events and events that are directly or indirectly linked to combat and terrorism.

The exclusion criteria consisted of mental fragility (identified by one's therapist as psychologically not capable of participating in the research) or schizophrenia or manic state; all cluster A personality disorders, borderline personality disorders, antisocial personality disorders, bipolar depression, Tourette syndromes, drug/alcohol addictions, low IQ and attention-deficit–hyperactivity disorder. The purpose of the criteria was to prevent any possible additional burden to respondents, who were clearly psychologically not capable to carry out the tasks. The control group included respondents who were not in the middle of an intake procedure, not in therapy and not diagnosed with the studied psychopathologies.

In total there were 153 respondents. The control group ($n=55$); PTSD and depressed respondents, who had not reported suicidal behaviour ($n=47$); the suicide ideation group, which ranged from mild to very severe and occurred in the month prior to data gathering ($n=31$); and attempters, who had attempted suicide up to 4 years prior to the study ($n=20$).

"Mild" suicidal ideation refers to low intensity (<5, [1 = lowest intensity–10 = highest intensity]) and frequency (i.e., a fleeting thought maximum twice a week). Suicidal ideations that are "present more than 50% of the time" describe ideations that have progressed into actions, for example, preparing a suicide plan. The ideations occur daily and are high in frequency and intensity (5–7), [1–10]). "Very severe suicide ideations" refers to a very high risk of a suicide attempt. The individual has already started to carry the suicide plan out. The intensity (8–10) and frequency (majority of the day) are very high.

The attempters and ideators had also been diagnosed with PTSD and/or depression. Due to the high comorbidity and similarities between PTSD and depression they were put into the same group.

Instruments

The diagnoses (*DSM-IV*) were initially determined and indicated by a psychiatrist. Patients who were diagnosed as being dysthymic or having major depression were included. All respondents were additionally assessed by the first author for current PTSD and depressive symptoms. PTSD was assessed with the help of the PTSD Checklist-Civilian version (PCL-C) (Weathers, Huska, & Keane, 1991), a self-report instrument with a 5-point Likert-scale, which has 17 items defined by diagnostic criteria B, C, D for PTSD in *DSM-IV*. The score range is 17–85. A total score >50 implies active PTSD (Berlant, 2004), which was applied as a cut-off. It provided a diagnostic sensitivity of .82 and specificity of .83 with a kappa coefficient of .64 (McKnight, McFall, & Kivilahan, 2001).

Depressive symptoms were initially assessed with the Hamilton scale (Hamilton, 1960). The Hamilton scale for depression is an observer-rated scale that is used to assess the presence and the severity of depression. It is typically administered by way of a semi-structured interview, in which 17 symptoms relating to depression are evaluated. The scoring was as follows: 10–13, mild depression; 14–17, moderate; 17–22, severe; and >23, very severe depression. It is a highly reliable and valid instrument. From 10 comparable studies, Kaslow et al. (2000) concluded a standard error of measurement of 3.07, and a test–retest reliability after 24 hr of .98.

The Depressive Experiences Questionnaire (DEQ) (Blatt, D'Afflitti, & Quinlan, 1976) enabled the identification of dependent depression and self-critical depression. It is a self-report questionnaire with a 7-point scale that contains 66 items. The DEQ demonstrates a high Cronbach alpha ($>.75$) and test–retest reliability (12 months) of .74 (Zuroff, Quinlan, & Blatt, 1990).

The GG scales are used as measurements of psychological change, evaluation of therapy and as a preliminary diagnostic tool. They were used in neuropsychopharmacological and psychosomatic studies as well as in various psychiatric sample groups, such as patients suffering from schizophrenia, anxiety disorders, alcohol/drug addictions and depression. Throughout the years more scales were added and were thoroughly tested. The computer version of the scales includes all the new additions.

Many studies were carried out in the United States, Canada, Germany and Australia, where the reliability of the scales was tested. The reliability for each scale was $\geq.80$ (Gottschalk, 1995; Viney 1983). The program's reliability increases with the number

of words in the sample. To our knowledge the GG scales were not previously used on PTSD and suicidal respondents. The influence of the interviewer on the respondents' free speech sample was limited by giving the standard neutral instruction with the use of a tape recorder. The speech samples were translated from Hebrew into English, which was controlled by another Israeli psychologist and scored by the computer program.

Additionally, with the help of a self-constructed semi-structured interview, the respondents were also asked about suicidal ideation and past suicide attempts. The interview began with general questions estimating the manner in which the respondents experienced life, for example, "how would you rate your ability to endure emotional pain (*low*, *average*, *good*, *very good*, *excellent*). The remaining questions focused on the intensity of thoughts, frequency, duration and preoccupation with death and suicide. The respondents were also asked about the way in which they coped with suicidal ideations and whether this was communicated.

Procedure

Respondents for this study were found in various locations in Israel: the hospital Hillel Yaffe (general psychiatric ward), Hadera's regional mental health clinic, Tirat Acarmel mental health centre and Tirat Acarmel regional mental health clinic. The "normal group" was recruited by advertisements placed in public places. The ethic committees of both hospitals approved the study and each respondent signed an informed consent.

The first author read the dossiers of each potential respondent in order to control for the diagnoses and the exclusion criteria. Patients who were still in therapy were approached only after permission was given by their therapist. A letter was sent to the respondents. They were told that the purpose of the study was to gather information about the thoughts, emotions and behaviours that they experience in their daily lives. Suicidal behaviour was mentioned in between a list of topics, so it would not appear to be an essential topic. The respondents were informed that participation was voluntary and that the data would remain anonymous and confidential. The letter also described the procedure. All respondents were reimbursed for their transportation expenses and only the control group received a small financial reward.

The one-to-one appointments with each respondent started with a short introduction. The respondent was asked to tell a 5-min story of their free choice about an influential event in his or her life. Unlike other projective or structured psychodiagnostic instruments, which confront the respondents with suggestive material that can influence the respondent's answers, the GG scales are based only on the free speech sample of the respondent. The story was audiotaped and then the PCL-C, DEQ and the Hamilton scale were administered. The semi-structured interview was carried out towards the end of the session and then a debriefing took place about the respondents' impressions and questions about the study.

After collection the data were transferred to a PC and analysed using SPSS (SPSS, Chicago, IL, USA). The translated stories were scored on the PCAD 2000. Mean differences between the four groups were calculated using one-way ANOVAs.

Results

Because sociodemographic variables may influence suicidal behaviour, partial correlations with controls for age, gender, partner status, combat/non-combat, exposure to terrorism and level of education were run for the diagnoses and suicidal behaviour, and no significant deviations were found.

Table I presents the variables gender, education level and the age groups of the respondents. Table I shows that the majority of the respondents were male and that the largest age group was 25–30 years. The distribution of high-school and university education was not statistically different.

The top six common negative life events reported by all respondents were death of a dear person (67.1%), serious illness in the family (50.3%), stress in the family (41.3%), own serious illness (34.2%), unemployment (29.7%), and a serious car accident (29%).

Out of the 33 GG scales the ANOVAs yielded 15 (Bonferroni corrected) significant results. Each significant F-score also produced one to three significant post hoc scores. The ANOVA results per GG scale are presented in Table II, which also lists the post hoc results, and effect-sizes are reported in parentheses.

Table I. Subject details

	Controls	PTSD/ MDD	Suicide ideators	Suicide attempters
$N = 153$	55	47	31	20
Female	19	22	9	9
Male	36	25	22	11
25–30 years	37	31	19	6
31–40 years	12	10	4	5
41–51 years	6	6	8	9
University	26	20	16	9
Some kind of higher education[a]	9	6	0	1
High school	20	21	15	10

Notes. MDD = major depressive disorder; PTSD = post-traumatic stress disorder.

[a]Complete and incomplete higher education, not of a university level.

Table II. Significant ANOVAs and post hoc results per group comparison and GG scale

GG scales	*Controls*		PTSD/depression without suicidal behaviour		Suicide ideators		Suicide attempters			Suicide attempters ($n = 20$) vs. controls ($n = 55$)	Suicide attempters ($n = 20$) vs. PTSD/depression without suicidal behaviour ($n = 47$)	Suicide attempters ($n = 20$) vs. suicide ideators ($n = 31$)
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*F*/(*df*)	*p* (ES)	*p* (ES)	*p* (ES)
Deterrents	−1.75	1.38	−2.33	1.60	−1.90	1.60	−3.57	1.77	7.08/(3;149)			** (−.99)
Achievement strivings	1.39	2.36	.40	2.50	.72	2.50	−.59	2.25	4.01/(3;149)	** (−.86)		
Health/sickness	.57	2.43	.03	2.77	.04	2.77	−1.69	2.66	3.78/(3;149)	** (−.89)		
Sickness	1.81	1.67	2.50	1.82	2.10	1.82	3.94	2.10	6.85/(3;149)	*** (1.12)		** (.91)
Self accusation	.25	.64	.43	.68	.68	.68	.75	.96	5.41/(3;149)	** (.61)		
Hopelessness	.64	.12	.68	.12	.64	.15	.79	.12	7.42/(3;149)	*** (.53)	** (.42)	** (.46)
Hope	.27	.87	.02	.73	.16	.73	−.35	.48	5.04/(3;149)	** (−.88)		
Human relations	−.24	1.29	−.47	1.31	−.45	1.31	−1.35	1.30	3.64/(3;149)	** (−.86)		
Frustrated dependency	.42	.91	.55	1.08	.74	1.08	1.25	1.41	7.97/(3;149)	*** (.70)	*** (.56)	** (.41)
Total depression	6.88	1.47	7.37	1.59	7.20	1.59	9.00	1.91	8.83/(3;149)	*** (1.24)	** (.93)	** (1.04)
Separation depression	.69	.24	.74	.20	.66	.20	.89	.32	4.34/(3;149)	** (.71)		** (.85)
Ambivalent hostility	.74	.09	.75	.09	.80	.09	.83	.12	5.41/(3;149)	** (.73)		
Total anxiety	1.85	.68	2.23	.71	2.23	.71	2.84	.71	8.69/(3;149)	*** (1.42)		
Separation anxiety	.69	.24	.76	.23	.66	.23	.90	.33	4.45/(3;149)	** (.73)		** (.87)
Mutilation anxiety	.76	.28	.89	.31	.89	.31	1.05	.43	3.78/(3;149)	** (.80)		

Notes. ES = effect size (all values. > .40 are to be considered as high, but they are sometimes different from the ANOVA results); GG scales, Gottschalk–Gleser scales; PTSD, post-traumatic stress disorder.

A Bonferroni correction was used for the post hoc multiple group comparisons.

$^{**}p < .017$, $^{***}p < .001$.

Table II lists the significant results of the GG scales in all group combinations. More significant scores were found for the comparisons between the attempters and the controls. Frustrated dependency, Total depression and Hopelessness yielded significant results in all three groups combinations.

Significant and probable suicide risk factors for each group comparison

Attempters and control group. All listed GG scales apart from the Deterrents scale had significant effects. It is apparent that the attempters group suffered from many psychological symptoms that are not prominently present among the "healthy" controls. High scores on these scales should alert the clinician to the seriousness of the suicidal behaviour.

Attempters and PTSD and depressive patients not reporting suicidal behaviour. The scales scores that were significantly higher were Hopelessness, Frustrated dependency and Total depression. The attempters had the highest score on the Total depression scale. The severity of these variables in an individual who suffers from PTSD and depression, increases the risk of suicidal behaviour.

Suicide attempters and ideators. The scales that yielded significant results were Deterrents, Sickness, Hopelessness, Frustrated dependency, Total depression, Separation depression and Separation anxiety. The highest score was on the Total depression scale. An individual with suicidal ideations who presents psychological states that are prominent in past suicide attempters, is likely to be at a very high risk of attempting suicide.

Discussion

The results of the group comparisons provide a deeper insight into the various psychological states that enhance the risk of suicide attempt.

Suicide attempters in comparison to controls

The comparison of suicide attempters to controls yielded the most suicidal risk factors: the attempters communicated a higher ambivalent hostility. Other people are perceived as threatening to the self. They are criticising, disappointing, abandoning and depriving. The self is also threatened with death from non-human and non-living objects. Suicide attempters exhibit aggressive behaviour towards others, as well as towards themselves (Plutchik, van Praag, Conte, & Picard, 1989). According to the present results, however, suicide attempters do not seem to express more aggression. Perhaps the suicide attempters are not struggling with their own rage, as much as they perceive themselves as victims of other's rage and aggressiveness. A suicide attempt may be their way of escaping the hurt and suffering that is inflicted on them by others.

Suicide attempters expressed lower Hope and Achievement strivings and higher Hopelessness and Frustrated dependency strivings than the control group. Suicide attempters have poor expectations of the future. They lack confidence and the belief that they are able to conquer the hardships in their lives, and therefore their motivation to strive for better circumstances is poor. The discrepancy between reality and their aspirations leads to aspiration strain (Zhang & Lester, 2008).

The suicide attempters also communicated higher levels of total anxiety than the control group. The attempters express anxiety of desertion, abandonment and loss of support or relationships. Ironically, they are also anxious about the occurrence of an event that may injure or physically damage them. The high total anxiety score implies that some elements of the other subscales are also present such as anxiety of death, inadequacy, exposure of shortcomings, criticism, shame, guilt and humiliation. Attempters perceive the self as weak and vulnerable. The source of their anxiety may be real or imaginary, but the anxiety felt is existent. A suicide attempt is thus perceived by them as an escape route towards safety and peacefulness.

The Total depression scale includes a broad range of many subscales, among which are also Psychomotor Retardation, Somatic Concerns, Death and Mutilation Depression, and Hostility Outward. On the whole scale there was a significant difference between the groups, but the score of each subscale separately was insignificant. The remaining subscales of the Total depression scale, namely Hopelessness, Separation Depression and Self-Accusation yielded significant differences between the groups. Attempters are highly depressed. They are preoccupied with and afraid of being abandoned, losing support and love. The Self-accusation scale includes guilt and shame depression. A significant score suggests that the attempters constantly think and fear the threat of being humiliated, morally disapproved of and criticised, and that their deficiencies would be exposed.

The attempters also ponder issues of sickness and health. They focus on their pain and suffering. The suicide attempters have poorer human relations than the controls. Their degree of interest and their capacity for obtaining positive, mutually helpful or satisfying relations is poor. Furthermore, the attempters seem to have a low level of altruism. Their empathy, caring, sincerity and compassion abilities are limited, which undoubtedly reduces their ability to relate to others.

Suicide attempters in comparison to PTSD and depressive patients not reporting suicidal behaviour

This group comparison defines the psychological states that differentiated the suicidal from the non-suicidal PTSD and depressed patients. The presence of these psychological states could be considered as a potential suicide risk factor.

As expected, total depression is higher among the attempters. It is noteworthy, however, that this group combination produced non-significant subscale results for the depression scale and for the entire anxiety scale. The attempters communicated higher levels of hopelessness and depression than the non-suicidal respondents. The constant re-experiencing of the traumatic event and the continuous presence of anxiety causes a sense of inescapability and lack of control. The reduced level of normal functioning leads to high frustration from having to depend on others for support and approval. There are also references in the speech samples of oral frustrations.

A failure to prevent death and killing of others in combat correlates with suicide attempts (Fontana, Rosenheck, & Brett, 1992). Roles in combat that are perceived as connected with high personal responsibility increase suicidal behaviour. In the present study, however, results for the Self-accusation scale were not significant. Hendin and Hass (1991) claimed that self-blame and anger are important affective states that are linked to suicide attempts. The hostility scales did not generate significant results. According to the present results it is probable that the combination of Hopelessness, Total depression and Frustrated dependency increase the attractiveness of death and the idea of committing suicide.

An additional astonishing point is the insignificant difference of the scores on the cognitive impairment scale. This suggests that the level of cognitive functioning of individuals with PTSD and depression is not much better than that of the attempters. Cognitive ability is important for the generation of alternative solutions, but from the present results it can be concluded that cognitive impairment is not the straw that breaks the camel's back. Rather, the presence of the aforementioned variables makes the difference as to whether an individual with PTSD and depression will attempt suicide or not.

Suicide attempters in comparison to the ideators group

The presence of these psychological states, which are more typical for suicide attempters than suicide ideators, in an individual with suicidal ideations suggests that the individual is at a high risk of attempting suicide. Hence, these psychological states intensify and increase the escalation of ideations to a suicide attempt:

The results show that suicide attempters experience anxiety and depression more intensely than suicide ideators. The total depression score is especially high. The main subscale in both dimensions is the theme of separation. Suicide attempters are afraid of and are preoccupied with loss. They communicate anguish, a feeling of being neglected and isolated by others, but at the same time they hate themselves for feeling this way. Being alone, lonely and abandoned increases their depression. It appears that the attempters yearn for intimacy or at least mourn its loss. They detest the fact that they are in need of support and help from others. There seems to be an inner conflict, which tears them apart, and in most likelihood increases their sense of hopelessness. Hopelessness is a catalyst to suicidal behaviour (Beck, Steer, Kovacs, & Garrison, 1985). Suicide risk increases as hopelessness reduces the will to live.

Deterrents increase the sense of hopelessness, which suicide attempters conveyed in their speech samples. Suicide attempters also displayed preoccupation with illness. They feel mentally and physically broken. The idea of achieving relief or a better future seems an inconceivable illusion.

According to van Egmond and Diekstra (1989) aggressive, agitated and violent behaviour is more common among ideators than among attempters. The results of the current study do not support their assertion. High hostility thus cannot be perceived as a variable that increases the risk of a suicide attempt. The high dependency on others for support and protection limits the ability to cope with the situation. Anxiety and low hope also hinder successful coping and thus magnify the perception of suicide as a good and reasonable solution.

Evaluation of the current study

In retrospect, the choice to use the GG scales on this fragile and challenging population was valuable. The ability to tap into the conscious and unconscious psychological states of the respondents has a great potential for the improvement of diagnostic abilities and the development of appropriate therapy. The freedom that this method provides gives an illusion of control, which lowers the respondents' defences and eliminates any need to conform, alter or produce socially desirable answers. It also gives the respondents the feeling that someone is listening to what they would like to tell, a feeling that usually improves communication.

In comparison to other projective tests, the application of the objective, systematic and computerised scales to a speech sample reduces the raters' bias. The results of the computerised version are also not influenced by the experimenter's level of

experience with the scoring system. The subdivisions of the scales produce a specific and detailed picture of the problematic domains, thus potentially increasing the reliability of the diagnosis.

Because Israeli subjects perceive disclosure of psychological problems as a sign of weakness, the use of the GG scales on this sample enabled the clinician to explore deep psychological problems without being dependent on the respondents' openness, self-knowledge or their willingness to comply. The respondents' perception of overexposure is thus limited. Consequentially, the use of the GG scales reduces the obstacle of poor self-disclosure, especially among individuals with suicidal tendencies and thus increases the ability of the clinician to derive a more accurate diagnosis at an early stage.

Crisis hotline and online therapy are currently alternative methods of providing psychological help. Seriously suicidal individuals who reach out to telephone crisis services report significant decrease in the level of suicidality at the end of the conversation (Gould, Kalafat, Harrismunfa, & Kleinman, 2007). Application of the GG scales could be useful for these therapy formats. The ability to analyse free speech samples of a caller will give the therapist the quickest and most reliable diagnosis of the individual within minutes. This option is not possible with other structured or projective tests. This method could therefore increase the effectiveness of the therapy.

The analysis of suicide notes of suicide attempters and completers is often used to better understand the psychological factors that led to the attempts (i.e., Barak & Miron, 2005). Why wait for a suicide note when one can analyse the free speech sample with the help of PCAD 2000 and prevent suicidal ideations from escalating to a suicide attempt? Naturally, more research should take place, but this could be an innovative tool for suicide prevention.

There are, however, a few disadvantages in this method. The personality and gender of the interviewer have been seen as influential on the content of speech. Men, for example, have the tendency to restrain their hostility expression in the presence of a female interviewer (Gottschalk & Gleser, 1969). It is possible that if the interviewer had been male, the hostility results would have been much higher. Naturally, this speculation can be tested by repeating the study with a male experimenter. Another limitation of the Gottschalk-Gleser method is that it ignores non-verbal data. Many non-verbal signs such as tempo of speech, intonation, facial movements and pauses can add or change the way a message is received.

This study focused on suicidal behaviour within specific Israeli groups. Population characteristics need to be taken into consideration when evaluating these results. For a cross-validation, future research could use similar experimental groups or groups with a different cultural background and then compare the results. Similar findings with different groups would indicate that the results are due to the suicidal behaviour and not due to cultural influences.

Another restriction of the present study could be seen in the relatively small size of the sample. Furthermore, this study was retrospective by nature. The suicide attempts happened up to 4 years prior to the study, thus the current suicide risk may not be acute.

The hypothesis that differences between the groups exist was supported. For all intents and purposes, even more psychological factors than imagined in the hypothesis yielded significant results. As far as we know, some of these psychological factors were not found previously with the help of structured tests. This study gave a new and fresh look at the psychological states that increase the risk of suicide behaviour among individuals from the general population, with PTSD and depression, suicidal ideations and attempters.

Diagnosticians should be especially alert to the psychological states that produced significant differences between all groups (sickness, hopelessness, frustrated dependency, total depression and total anxiety), but group-specific differences should not be ignored either. Applying the knowledge of these results, especially at the diagnostic phase, may help clinicians to take preventative measures – on time – before suicidal ideations turn into suicide attempts.

Acknowledgement

We thank Robert Bechtel for having run all protocols and producing all scores. We would like to dedicate this paper to the late Louis Gottschalk (M.D., Ph.D.) Professor em. at UC Irvine, who in person and by the Gottschalk-Gleser method of analyzing speech samples has inspired this research.

References

American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: Author.

Apter, A., Horesh, N., Gothelf, D., Graffi, H., & Lepkifker, E. (2001). Relationship between self-disclosure and serious suicidal behavior. *Comprehensive Psychiatry, 42*, 70–75.

Apter, A., King, R. A., Bleich, A., Fluch, A., Kottler, M., & Kron, S. (2008). Fatal and non-fatal suicidal behavior in Israeli adolescent male ideators. *Archives of Suicide Research, 12*, 20–29.

Barak, A., & Miron, O. (2005). Writing characteristics of suicidal people on the internet: A psychological investigation of emerging social environment. *Suicide and Life Threatening Behavior, 35*, 507–524.

Beck, A. T., Steer, R. A., Kovacs, M., & Garrison, B. (1985). Hopelessness and eventual suicide: A 10 years prospective study of patients hospitalized with suicide ideations. *American Journal of Psychiatry, 142*, 559–563.

Bell, J. B., & Nye, E. C. (2007). Specific symptoms predict suicidal ideation in Vietnam combat veterans with chronic PTSD. *Military Medicine, 172*, 1144–1147.

Berlant, J. L. (2004). Prospective open-label study of add-on and monotherapy topiramate in civilians with chronic nonhallucinatory posttraumatic stress disorder. *BMC Psychiatry, 4*, 24.

Blatt, S. J., D'Affliti, J. P., & Quinlan, D. M. (1976). *Depressive experiences questionnaire.* New Haven, CT: Yale University Press.

Bodner, E., Ben-Artzi, E., & Kaplan, Z. (2006). Soldiers who kill themselves: The contribution of dispositional and situational factors. *Archives of Suicide Research, 10*, 29–43.

Fontana, A., Rosenheck, R. A., & Brett, E. A. (1992). War zone trauma and posttraumatic stress disorder symptomatology. *Journal of Nervous and Mental Disease, 180*, 748–755.

Gottschalk, L. A. (1995). *Content analysis of verbal behaviour: New findings and clinical applications.* Hillsdale, NJ: Lawrence Erlbaum.

Gottschalk, L. A., & Bechtel, R. J. (1982). Manual of uses and applications of the Gottschalk-Gleser verbal behaviour scales. *Research Communications in Psychology, Psychiatry and Behavior, 7*, 273–327.

Gottschalk, L. A., & Bechtel, R. J. (1998–2003). *PCAD 2000: Psychiatric Content Analysis and Diagnosis.* Corona del Mar, CA: GB software.

Gottschalk, L. A., & Gleser, G. C. (1969). *The measurement of psychological states through the content analysis of verbal behavior.* Berkley, CA: University of California Press.

Gould, M. S., Kalafat, J., Harrismunfa, J. L., & Kleinman, M. (2007). An evaluation of crisis hotline outcome: Part 2 suicidal calls. *Suicide and Life Threatening Behaviour, 37*, 338–352.

Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery and Psychiatry, 23*, 56–62.

Hawton, K., & van Heeringen, K. (2000). *The international handbook of suicide and attempted suicide.* Chichester: John Wiley & Sons.

Hendin, H., & Hass, A. P. (1991). Suicide and guilt as manifestations of PTSD in Vietnam combat veterans. *American Journal of Psychiatry, 148*, 586–591.

Horesh, N., & Apter, A. (2006). Self disclosure, depression, anxiety and suicidal behavior in adolescent psychiatric inpatients. *Crisis, 27*, 66–71.

Kaslow, N. J., Thompson, M. P., Meadows, L. A., Chance, S., Puett, R., Hollins, L., et al. (2000). Computer assessment of depression: Automating the Hamilton depression rating scale. *Drug Information, 34*, 145–156.

Levi, Y., Horesh, N., Fischel, T., Treves, I., Or, E., & Apter, A. (2008). Mental pain and its communication in medically serious suicide attempters: An impossible situation. *Journal of Affective Disorders, 111*, 244–250.

McKnight, K. M., McFall, M. E., & Kivilahan, D. R. (2001). Trauma symptoms of PTSD and associated problems among incarcerated veterans. *Psychiatric Services, 52*, 959–964.

Ministry of Health. (2005). *Statistical annual report 2005 of the Ministry of Health of Israel.* Retrieved 5 January 2008, from http://www.health.gov.il

Ministry of Health. (2008). *Statistical annual report 2008 of the Ministry of Health of Israel.* Retrieved 5 June 2009, from http://www.health.gov.il

Neizer, T. (2008, April 16). Can the media reduce the suicide rate? Retrieved 22 October 2009, from http://www.ynet.co.il (an online version of *Yediot Hacharonot*).

Plutchik, R., van Praag, H. M., Conte, H. R., & Picard, S. (1989). Correlates of suicide and violence risk: The suicide measure. *Comprehensive Psychiatry, 30*, 296–302.

Salib, E. (2003). Effects of 11 September 2001 on suicide and homicide in England and Wales. *British Journal of Psychiatry, 183*, 207–212.

Shneidman, E. S. (1985). *Definition of suicide.* New York: John Wiley and Sons.

Solomon, Z., Zur-Naah, S., Horsh, D., Zerach, G., & Keinan, G. (2008). The contribution of stressful life events through the life cycle to combat induced psychopathology. *Journal of Traumatic Stress, 21*, 318–325.

Southwick, S. M., Yehuda, R., & Giller, E. L. Jr. (1991). Characterization of depression in war-related posttraumatic stress disorder. *American Journal of Psychiatry, 148*, 179–183.

Tal, A., Roe, D., & Corrigan, P. W. (2007). Mental illness stigma in the Israeli context: Deliberations and suggestions. *International Journal of Social Psychiatry, 53*, 547–563.

Tartakovsky, E. (2007). A longitudinal study of acculturative stress and homesickness: High school adolescents immigrating from Russia and Ukraine to Israel without parents. *Social Psychiatry and Psychiatric Epidemiology, 42*, 485–494.

Van Egmond, M., & Diekstra, R. F. W. (1989). The predictability of suicidal behavior: The results of a meta-analysis of published studies. In: R. F. W. Diekstra, R. A. Maris, S. Platt, A. Schmidtke, & G. Sonneck (Eds.), *Suicide and its prevention* (pp. 37–61). Leiden: E. J. Brill.

Viney, L. L. (1983). The assessment of psychological states through content analysis of verbal communications. *Psychological Bulletin, 94*, 542–563.

Weathers, F., Huska, J., & Keane, T. (1991). *The PTSD Checklist-civilian version.* Boston, MA: National Centre for PTSD.

Zhang, J., & Lester, D. (2008). Psychological tensions found in suicidal notes: A test for strain theory of suicide. *Archives of Suicide Research, 12*, 67–73.

Zuroff, D. C., Quinlan, D. M., & Blatt, S. J. (1990). Psychometric properties of the Depressive experiences questionnaire. *Journal of Personality Assessment, 55*, 65–72.