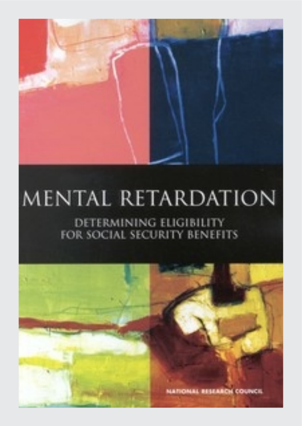

Mental Retardation: Determining Eligibility for Social Security Benefits

DETAILS

350 pages | 6 x 9 | PAPERBACK
ISBN 978-0-309-08323-2 | DOI 10.17226/10295

AUTHORS

Committee on Disability Determination for Mental Retardation, Daniel J. Reschly, Tracy G. Myers, and Christine R. Hartel, Editors, National Research Council

MENTAL RETARDATION

DETERMINING ELIGIBILITY FOR SOCIAL SECURITY BENEFITS

Committee on Disability Determination for Mental Retardation

Daniel J. Reschly, Tracy G. Myers, and
Christine R. Hartel, *Editors*

Board on Behavioral, Cognitive, and Sensory Sciences
Division of Behavioral and Social Sciences and Education
National Research Council

NATIONAL ACADEMY PRESS
Washington, D.C.

NATIONAL ACADEMY PRESS 2101 Constitution Avenue, N.W. Washington, D.C. 20418

NOTICE: The project that is the subject of this report was approved by the Governing Board of the National Research Council, whose members are drawn from the councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. The members of the committee responsible for the report were chosen for their special competences and with regard for appropriate balance.

The study was supported by Contract No. 0600-99-38803 between the National Academy of Sciences and the U.S. Social Security Administration. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the view of the organizations or agencies that provided support for this project.

Library of Congress Cataloging-in-Publication Data

Mental retardation : determining eligibility for social security
benefits / Daniel J. Reschly, Tracy G. Myers, and Christine R. Hartel,
editors ; Committee on Disability Determination for Mental Retardation,
Board on Behavioral, Cognitive, and Sensory Sciences, Division of
Behavioral and Social Sciences and Education, National Research Council.
p. ; cm.
Includes bibliographical references and index.
Supported by contract between the National Academy of Sciences and the
U.S. Social Security Administration Contract No. 0600-99-38803
ISBN 0-309-08323-0 (pbk.)
1. Mental retardation—Diagnosis. 2. Disability evaluation—United
States. 3. Social security—United States.
[DNLM: 1. United States. Social Security Administration. 2.
Disability Evaluation—United States. 3. Eligibility
Determination—standards—United States. 4. Mental
Retardation—diagnosis—United States. 5. Public Policy—United States.
6. Social Security—United States. WM 304 M549 2002] I. Reschly,
Daniel J. II. Myers, Tracy G. III. Hartel, Christine R., 1947- IV.
National Research Council (U.S.). Committee on Disability Determination
for Mental Retardation. V. United States. Social Security
Administration.
RC570.2 .M46 2002
616.85'884075—dc21
2002005467

Additional copies of this report are available from National Academy Press, 2101 Constitution Avenue, N.W., Washington, D.C. 20418

Call (800) 624-6242 or (202) 334-3313 (in the Washington metropolitan area)

This report is also available online at **http://www.nap.edu**

Printed in the United States of America

Suggested citation: National Research Council (2002) *Mental Retardation: Determining Eligibility for Social Security Benefits.* Committee on Disability Determination for Mental Retardation. Daniel J. Reschly, Tracy G. Myers, and Christine R. Hartel, editors. Division of Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.

Cover: Michele de la Menardiere. Detail from Colors #5. Acrylic on canvas.

THE NATIONAL ACADEMIES

National Academy of Sciences
National Academy of Engineering
Institute of Medicine
National Research Council

The **National Academy of Sciences** is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. Upon the authority of the charter granted to it by the Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Bruce M. Alberts is president of the National Academy of Sciences.

The **National Academy of Engineering** was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. Wm. A. Wulf is president of the National Academy of Engineering.

The **Institute of Medicine** was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, upon its own initiative, to identify issues of medical care, research, and education. Dr. Kenneth I. Shine is president of the Institute of Medicine.

The **National Research Council** was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both Academies and the Institute of Medicine. Dr. Bruce M. Alberts and Dr. Wm. A. Wulf are chairman and vice chairman, respectively, of the National Research Council.

COMMITTEE ON DISABILITY DETERMINATION FOR MENTAL RETARDATION

DANIEL J. RESCHLY (*Chair*), Department of Special Education, Vanderbilt University
HAROLYN M. E. BELCHER, Department of Pediatrics, The Johns Hopkins University School of Medicine
JAN BLACHER, School of Education, University of California at Riverside
BRUCE A. BRACKEN, School of Education, The College of William and Mary
ELISABETH M. DYKENS, Neuropsychiatric Institute, University of California, Los Angeles
THOMAS HEHIR, Graduate School of Education, Harvard University
JOHN W. JACOBSON, New York State Office of Mental Retardation and Developmental Disabilities, Albany
BETTE R. KELTNER, School of Nursing, Georgetown University
MARTY WYNGAARDEN KRAUSS, Heller Graduate School, Brandeis University
HANS C. PALMER, Economics Department, Pomona College
SHARON LANDESMAN RAMEY, Civitan International Research Center, University of Alabama, Birmingham
JAMES C. SHANTEAU, Department of Psychology, Kansas State University
WAYNE P. SILVERMAN, Department of Psychology, New York State Institute for Basic Research in Developmental Disabilities
GARY N. SIPERSTEIN, Center for Social Development and Education, University of Massachusetts, Boston
SARA S. SPARROW, Psychology Department, Yale University
KEITH F. WIDAMAN, Psychology Department, University of California, Davis

CHRISTINE R. HARTEL, *Study Director*
TRACY G. MYERS, *Senior Research Associate*
WENDY E. MINICH, *Senior Project Assistant*

BOARD ON BEHAVIORAL, COGNITIVE, AND SENSORY SCIENCES

Contents

Acknowledgments

This report is the result of two years of hard work on the part of a dedicated, diverse, talented, and energetic panel of experts, who were supported by an equally dedicated, talented, and energetic staff. The National Research Council (NRC), in response to a request from the U.S. Social Security Administration (SSA), organized the effort. This study examined the scientific base and current clinical practices in diagnosing mental retardation for disability benefit purposes. The primary focus is on individuals in the mild range of mental retardation, those with intelligence scores hovering around 70. The committee examined the scientific base regarding intelligence and its assessment, adaptive behavior and its assessment, the relationship between intelligence and adaptive behavior, and differential diagnosis. Members also recognized that the results of this work might have effects on SSA, disability benefit recipients, and other service systems that provide supports to this population. Consequently, the related literature was also examined.

The Committee on Disability Determination for Mental Retardation was aided by numerous individuals and organizations in the preparation of this report.

The Office of Disability in SSA sponsored the study. Its staff was very helpful in assisting committee members and staff to understand the disability system. Terrence Dunlop, our project manager, and David Barnes were helpful in explaining the purposes of the study from SSA's perspective. Cathy Lively and Cheryl Wrobel patiently answered members' questions about the workings of the disability determination system and provided supplemental information on the numbers of people applying for disability benefits.

The committee was assisted in its data collection efforts by a number of individuals who authored commissioned papers on various topics related to the study. These papers were very helpful in laying out issues that the committee had to tackle in its deliberations. Special thanks go to Susan Parish and David Braddock, whose analysis of the policy context of SSA's disability determination process forms the basis of the text in Chapter 2. The committee is also very grateful to Sharon Borthwick-Duffy for an excellent paper on the assessment of adaptive behavior; parts of it were used in Chapter 4. Thanks are also extended to Andrew Houtenville for his work on the economic experience of individuals with mental retardation; Jonathan Leland and Robyn Dawes for a thought-provoking paper on decision making among individuals with mental retardation; David Mank for his paper and presentation on employment and employability among individuals with mental retardation; and Bruce Shapiro for his paper on differential diagnosis. Jane Thompson conducted Monte Carlo simulations of the relationship between intelligence and adaptive behavior using widely available and well-standardized measures; these were crucial to the committee's deliberations and ultimately to its conclusions and recommendations.

The committee is indebted to the NRC staff for their work in

preparation of this report. Tracy Myers, Christine Hartel, and Wendy Minich shepherded the group through the task with efficiency and good humor. We were also fortunate to have an excellent editor, Christine McShane. Finally, I want to personally thank the committee for their intelligence and sense of purpose in approaching and completing the task.

This report has been reviewed in draft form by individuals chosen for their diverse perspectives and technical expertise, in accordance with procedures approved by the NRC's Report Review Committee. The purpose of this independent review is to provide candid and critical comments that will assist the institution in making the published report as sound as possible and to ensure that the report meets institutional standards for objectivity, evidence, and responsiveness to the study charge. The review comments and draft manuscript remain confidential to protect the integrity of the deliberative process.

We thank the following individuals for their participation in the review of this report: John M. Belmont, University of Kansas Medical Center; David Braddock, Coleman Institute for Cognitive Disabilities and Associate Vice President for Research, University of Colorado; Jacob Klerman, RAND, Santa Monica, CA; William MacLean, Department of Psychology, University of Wyoming; Kevin McGrew, Department of Educational Psychology, University of Minnesota; Iris Tan Mink, Department of Psychiatry and Behavioral Science, University of California, Los Angeles; Donald K. Routh, Department of Psychology, University of Miami; Frank Rusch, College of Education and Children's Research Center, University of Illinois at Urbana-Champaign; and Michael Rutter, Institute of Psychiatry, London, England.

Although the reviewers listed above have provided many constructive comments and suggestions, they were not asked to endorse the conclusions or recommendations nor did they see the final draft of the report before its release. The review of this report was overseen by Richard Wagner, Department of Psychology, Florida State University. Appointed by the National Research Council, he was responsible for making certain that an independent examination of this report was

carried out in accordance with institutional procedures and that all review comments were carefully considered. Responsibility for the final content of this report rests entirely with the authoring committee and the institution.

Daniel J. Reschly, *Chair*
Committee on Disability Determination
for Mental Retardation

MENTAL RETARDATION

Executive Summary

The U.S. Social Security Administration (SSA) asked the National Research Council to assess its disability determination process for mental retardation. The Committee on Disability Determination for Mental Retardation was formed to evaluate the existing determination process in the context of state-of-the-art scientific knowledge and clinical practice. The committee was also asked to suggest new procedures that may be necessary to ensure that SSA eligibility determinations are based on procedures and criteria that conform to best professional practices and to identify promising areas of research that may help to clarify unaddressed or incompletely answered questions.

Mental retardation is a condition characterized by significant deficits in intellectual capabilities and adaptive behavior. Its onset occurs during the developmental period, the period through age 21. The diagnosis can be particularly difficult to make in the mild range of retardation, in which adaptive behavior deficits are less easily quantified with commonly used instruments. Differences in content, standardization, and floor and ceiling effects on broadly used measures of adaptive behavior, as well as different conceptualizations of the nature of adaptive behavior, all contribute to this difficulty. Current estimates suggest that anywhere from 1 to 3 percent of people living in the United States will receive a diagnosis of mental retardation. These varying prevalence estimates reflect (1) differences in the way that mental retardation is defined, interpreted, and measured; (2) differences in the ways in which students are identified in urban and rural education systems; and (3) whether individuals or their families from varying cultural, racial, and ethnic backgrounds choose to apply for services. In addition, definitions of mental retardation vary, with SSA, the major professional organizations, and the World Health Organization all providing different definitions of the condition.

SSA provides income support and medical benefits to many individuals with mental retardation. Benefits are provided to adults unable to perform substantial gainful activity (i.e., work) because of mental retardation through the Disability Insurance (DI) program and the Supplemental Security Income (SSI) program. SSI benefits are provided as well to the families of children and adolescents who evidence "marked and severe" restrictions in functioning because of mental retardation. The determination decisions are made through state disability determination services, with payments coming from the federal agency and, in some jurisdictions, supplemented by state resources.

COMMITTEE CHARGE

Specifically, the committee has been asked to (a) examine the adequacy of the SSA definition of mental retardation, (b) comment on the current procedures for assessing intellectual capabilities and indicate how best to make that assessment consistent with current science and professional practice, (c) discuss the issue of adaptive behavior and its assessment consistent with current science and widespread professional practice, (d) provide advice on the most appropriate ways of combining data from intellectual and adaptive functioning in order to provide a complete profile of an individual's capabilities, and (e) clarify ways of better differentiating mental retardation from other conditions with which it shares signs and symptoms. The committee has also been asked to provide suggestions for research or evaluation that could clarify unaddressed or incompletely answered issues. This report answers these questions and makes specific recommendations as well.

ADEQUACY OF THE SSA DEFINITION

Although no single definition of mental retardation is universally accepted in all its details, intellectual deficits are central to all such definitions. The consensus of the major professional associations and health-related organizations is that mental retardation involves deficits in intellectual functioning and adaptive behavior and has its onset during the developmental period. Their definitions differ from each other on a number of critical specifics, including the nature and measurement of the deficits in adaptive behavior and the age of onset.

SSA's definition of mental retardation includes the criteria used by these authoritative sources. However, it diverges from the standard diagnostic nomenclature in some ways, including the nature and assessment of deficits in adaptive behavior, its basis for determining subaverage intellectual functioning, and the age of onset. In addition, the SSA definition goes beyond the criteria used by these other organizations in identifying individuals as having mental retardation if they have both deficits in intellectual functioning and comorbid medical or psychiatric disorders—a circumstance addressed by no other definition of mental retardation.

It is important to note that the differences between the SSA definition of mental retardation and those of the major professional and health-related organizations derive from the purpose for which it is used. The SSA definition is used not for diagnostic purposes, but rather for purposes of program eligibility. The SSA definition fulfills its purpose of identifying individuals with cognitive limitations who experience significant problems in their ability to perform work and may therefore be in need of governmental support. However, the committee recommends specific changes to the SSA definition, related to the criteria for intellectual functioning and adaptive behavior. The committee's proposed changes are detailed in the recommendations that follow.

INTELLIGENCE AND ITS ASSESSMENT

SSA and the major professional associations and health-related organizations all define mental retardation in terms of deficits in intellectual functioning that are approximately two standard deviations below the mean of the population on commonly used intelligence tests. In contrast to the position taken by the other organizations, which focuses on the composite intelligence score, SSA allows the use of partial or part scores in a determination of mental retardation. For instance, when either the verbal or performance IQ on Wechsler measures is at least two standard deviations below the mean, a person could be deemed to have met the intellectual functioning criterion for mental retardation. Although SSA says that it will accept part scores from any individually administered, standardized IQ test, it uses part scores from the Wechsler measures in its regulations as an example. The presumably unintended consequence of that example is that, in clinical practice, the Wechsler measures are used most frequently for disability determination.

In the committee's judgment, composite scores from intelligence tests should be used routinely in mental retardation diagnosis, except when the composite IQ score's validity is in doubt, in which case an appropriate part score may be used in its place. Significant and meaningful variation among an instrument's respective part scores may indicate evidence of compromised validity for one or more of them (for example, a low verbal scale score for an individual with a suspected speech disorder), which in turn would threaten the validity of the composite IQ score. In such situations, appropriate part scores may better represent the individual's true overall level of cognitive functioning.

Only part scores derived from scales that demonstrate high *g*-loadings—that is, ones that are better representations of general intellectual ability (e.g., crystallized, fluid measures of intelligence)—should be used in place of the composite IQ score when its validity is in doubt. Many intelligence tests access several facets of intelligence, but not all facets are equally important or predict life events equally well. Those intellectual facets that are heavily "*g*-saturated" provide the best sources for replacing the composite IQ score when its validity is questionable.

The committee makes the following recommendation related to intelligence and its assessment.

Recommendation: A client must have an intelligence test score that is two or more standard deviations (SD) below the mean (e.g., a score of 70 or below, if the mean = 100 and the standard deviation = 15).

- *Composite score is 70 or below:* **If the composite or total test score meets this criterion, then the individual has met the intellectual eligibility component.**[1]
- *Composite score is between 71 and 75:* **If the composite score is suspected to be an invalid indicator of the person's intellectual disability and falls in the range of 71-75, a part score of 70 or below can be used to satisfy the intellectual eligibility component.**
- *Composite score is 76 or above:* **No individual can be eligible on the intellectual criterion if the composite score is 76 or above, regardless of part scores.**[2]

---

[1] Discussion of the rare instance in which a composite IQ is 70 or below but is suspected to be invalid is in Chapters 3 and 5.

[2] Committee member Keith Widaman dissents from this part of the recommendation. Dr. Widaman believes that IQ part scores representing crystallized intelligence (*Gc*, similar to verbal IQ) and fluid intelligence (*Gf*, related to performance IQ) have clear discriminant validity and represent broad, general domains of intellectual functioning.

The committee recommends continuation of the criterion of presumptive eligibility for persons with IQs below 60.

The use of part scores, most often from the Wechsler measures, introduces an important consideration in the clinical use of intelligence measures for disability determination. Current scientific conceptions of intelligence focus primarily on fluid and crystallized abilities, with recognition that working or comprehensive memory is also important to overall intellectual functioning. Many intelligence tests are based on these distinctions. The Wechsler measures are also moving in this direction, with a focus on factor scores that are analogous to crystallized intelligence (e.g., verbal comprehension index), fluid intelligence (e.g., perceptual organization index), and working/comprehensive memory (e.g., working memory index). Consequently, the committee has recommended continued use of part scores in eligibility determination, but is advocating use of part scores that are consistent with current scientific thinking.

ADAPTIVE BEHAVIOR AND ITS ASSESSMENT

Deficits in adaptive behavior, together with deficits in intelligence, are also central to current definitions of mental retardation. SSA and the major professional organizations disagree on the nature and degree

---

Therefore, a score of 70 or below on either of these part scores from any standardized, individually administered intelligence test that reports such scores should be deemed sufficient to meet the listings for low general intellectual functioning regardless of the level of the composite score, providing that the part scores have adequate psychometric properties (e.g., high reliability, low standard error of measurement). Dr. Widaman notes that, without any clear justification, SSA currently accepts either a composite IQ score from any standardized, individually administered intelligence test or a verbal or performance IQ score, any one of which can be 70 or below. SSA does not stipulate that the composite IQ must be below a certain score for a part score to be used. Dr. Widaman's position provides a rationale for current SSA use of part scores, but it (a) aligns the acceptable part scores with the constructs of *Gc* and *Gf* used in contemporary theories of mental abilities and (b) argues that usable part scores for *Gc* and *Gf* should not be limited to those derived from any particular test instrument.

of adaptive behavior deficits that must be apparent before a diagnosis of mental retardation can be made. There are also differences in how these behavioral deficits are to be measured, with SSA, the American Association for Mental Retardation, and the *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition* suggesting that behavioral descriptions of adaptive functioning are sufficient, and Division 33 of the American Psychological Association and the World Health Organization requiring the use of standardized assessment measures. The committee's judgment is that the approach currently taken by SSA can result in inconsistent decision making, because different types and quality of information are used in making the determination.

The committee makes the following recommendations related to adaptive behavior and its assessment.

Recommendation: Standardized adaptive behavior instruments should be used to determine limitations in adaptive functioning. In general, the cutoff scores for adaptive behavior should be one standard deviation below the mean in two adaptive behavior areas or one and one-half standard deviations below the mean in one adaptive behavior area.

- **Adaptive behavior measures should be used whenever possible, but only when there is an instrument that matches the client's characteristics and when an appropriate third-party respondent is available.**
- **A client can be determined to have a significant limitation in adaptive functioning even with scores that do not meet the above criteria *if* there is compelling evidence of adaptive behavior deficits that significantly impair performance of expected behaviors.**

Recommendation: Revisions should be made in the adaptive behavior areas or domains emphasized in SSA regulations to conform to factor analytic results. The following areas by age should be adopted by SSA:

- **Infancy/early childhood (approximate ages birth to 4): motor/mobility, social, communication, daily living skills (self-help).**
- **Childhood (approximate ages 5 to 17): motor/mobility, social, communication/functional academic skills, daily living skills.**
- **Adolescence/adulthood (approximate ages 18 and older): motor/mobility, social, communication/practical cognitive skills, daily living skills, work skills/work-related behaviors.**

COMBINING IQ AND ADAPTIVE BEHAVIOR DATA

Interpreting the combination of data on intelligence and adaptive behavior functioning is at the crux of making mental retardation disability determinations, particularly for individuals whose intellectual functioning hovers around an IQ of 70. The committee recognizes that this is a very difficult task. Studies have estimated the relationship between IQ and adaptive behavior ranging from 0 (indicating no relationship), to almost +1 (indicating a perfect relationship). Data also suggest that the relationship between IQ and adaptive behavior varies significantly by age and levels of retardation, being strongest in the severe and moderate ranges and weakest in the mild range. There is a dearth of data on the relationship of IQ and adaptive behavior functioning at the mild level of retardation, affecting the group that is most difficult to assess for the SSA determination process. Consequently, in the committee's view, informed, trained judgments are needed in decisions about how to assess these core dimensions of mental retardation and how to interpret the findings that result.

The committee makes the following recommendations with respect to combining these two types of data to determine whether an individual has mental retardation for disability benefit purposes.

Recommendation: A diagnosis of mental retardation should be based on high-quality assessments of intellectual and adaptive functioning that meet the following criteria:

- **A broad variety of information on adaptive behavior and intelligence should be collected, including data on performance in different settings, from different sources, and using varying methods.**
- **Comprehensive, multifactored measures of intelligence and adaptive behavior should be used in mental retardation eligibility determination. Brief, unidimensional measures or short forms of comprehensive tests should not be used.**
- **The principle of convergent validity shall be applied in eligibility decisions about mental retardation. Information that is inconsistent with a diagnosis of mental retardation should be recognized, evaluated, and explained in the overall diagnostic decision.**
- **Assessments must be conducted by people with appropriate education and training for the kind of instrument used and the nature of the eligibility decision to be made. People conducting intellectual assessments must meet test publishers' requirements for Class C instruments.**
- **Measures of adaptive and intellectual functioning should be carefully selected and interpreted in order to minimize the negative effects of low validity, low reliability, floor and ceiling effects, and steep item gradients.**
- **The norms for measures of adaptive behavior and intellectual functioning must be suitably contemporary. Use of outdated norms or previous editions of recently restandardized measures is not acceptable. The norms for intellectual measures should be no older than 12 years because of the deterioration of normative standards over time.**
- **Decisions about eligibility for a diagnosis of mental retardation should be made by people with appropriate preparation in the areas of mental retardation and other disabilities and disorders, measurement of intellectual and adaptive functioning, knowledge of human development, and the influence of context on behavior.**

DIFFERENTIAL DIAGNOSIS

Distinguishing mental retardation from other disorders with similar behavioral and cognitive signs and symptoms is referred to as *differential diagnosis.* SSA has had particular difficulty differentiating mental retardation from other disorders in children and adolescents. Following recent legal and legislative action, including the Supreme Court decision in *Sullivan v. Zebley*, the Personal Responsibility and Work Opportunity Reconciliation Act of 1996, and the Balanced Budget Act of 1997, SSA inappropriately declassified some children who had been previously determined to have mental retardation for purposes of receiving disability benefits. Consequently, the agency asked the committee for advice on how best to differentiate conditions that share signs and symptoms with mental retardation. These conditions include autism and other pervasive developmental disorders, learning disability, and borderline intellectual functioning, as well as others.

The committee makes the following recommendations relevant to differential diagnosis.

Recommendation: Social Security Disability Determination Specialists may differentiate other conditions from mental retardation by using intelligence and adaptive behavior test criteria as outlined in the committee's recommendations. Data, including school test results, intelligence and adaptive behavior test results, and psychiatric and medical test results, from community-based agencies (such as schools, hospitals, or clinics) can be used to inform the determination of SSA eligibility for the diagnosis of mental retardation, but the diagnoses given by community agencies should not be used.

- **Social Security Disability Determination Specialists may differentiate individuals with borderline intellectual functioning and learning disability from those with mental retardation by reviewing cognitive and adaptive behavior test results and determining whether the individual meets diagnostic criteria for**

mental retardation as outlined in the committee's first two recommendations.

- **Social Security Disability Determination Specialists do not need to determine the presence or absence of mental retardation in individuals who are eligible for SSI due to other neurodevelopmental or psychiatric disabilities (e.g., autism, pervasive developmental disorder, attention deficit hyperactivity disorder, genetic syndromes, intrauterine exposure to alcohol or environmental toxins, sensory impairments, seizure disorders, or severe emotional-behavioral disorders).**
- **Objective data on intellectual and adaptive functioning to determine mental retardation should be collected for individuals with mild neurodevelopmental or psychiatric disabilities who might have impairments that are consistent with or functionally equivalent to mental retardation.**

SSA PROGRAMS IN CONTEXT

The committee frames its response to SSA's charge in terms of the specific context of the DI and SSI disability benefit programs. These programs provide needed support to children and their families and to adults with mental retardation. Individuals with mental retardation constitute the largest diagnostic category of children receiving SSI—26 percent. The total number of individuals receiving SSA benefits for mental retardation is more than 1 million. These programs not only provide benefits to a large number of people, but also serve as a critical gateway to a number of other federal and state benefits. This linkage is particularly important because of the difficulty that individuals with mental retardation have in securing and maintaining substantial gainful employment.

Thus, the accuracy of a diagnosis of mental retardation is important to both the client and SSA: false positives allow people into the system who should not be receiving benefits and thus strain its resources, but false negatives not only deny SSI and DI benefits to those who are legally entitled to receive them, but also keep them from obtaining other federal and state benefits. Since all applicants for SSI benefits must also pass a means test, SSA is necessarily dealing with people who are already on the economic margins of society.

Recently, SSA has developed a number of new programs designed to encourage adults with mental retardation to obtain their first job or to return to work. These programs are at an early stage of implementation, and their effectiveness has yet to be determined. The agency appears to be balancing its role as a safety net with an attempt to encourage recipients to work. In the committee's view, this balance is the right one to strike.

The committee makes the following recommendation to assist SSA and disability benefit recipients in achieving this balance.

Recommendation: The Social Security Administration should remove disincentives for people with mental retardation to seek employment by:

- **Considering individuals with mental retardation to be presumptively re-eligible for benefits throughout their lives, if they have previously received benefits, subsequently secured gainful employment, and then lost that employment.**
- **Encouraging the use of work incentive programs for people with mental retardation, with appropriate and necessary protections of each program's role as a safety net for income support.**
- **Permitting individuals with mental retardation to retain eligibility for Medicaid independent of their employment status.**

RESEARCH

Finally, in reviewing the literature on mental retardation and related areas, the committee has identified a number of promising research opportunities that would help to clarify unaddressed or incompletely answered questions. We are also aware that SSA is expecting to receive momentarily from research organizations like RAND the results of various analyses it has commissioned, which may contribute to changes in its policies on disability determination. In addition, the American Association on Mental Retardation has undertaken deliberations that may soon result in changes to its 1992 definition of mental retardation and its constructs. These and other activities highlight the active nature of this field in both research and practice. The committee's recommendation for needed research is intended for SSA and other federal agencies that work on education, vocational training, health and mental health, and disability-related issues.

Recommendation: Federal agencies, including the Social Security Administration, should fund studies to evaluate the accuracy of program eligibility decisions and foster research on adults with mental retardation, including their adaptive behavior. The research funding should include investigations of multimethod techniques for the assessment of job-related skills, social adaptation, health, and well-being. In addition, relevant epidemiological studies and research on the accuracy of the diagnosis of mild mental retardation are essential to further inform policy and decision making.

- **SSA should evaluate the consequences of implementing the committee's recommendations in the context of public policies and economic conditions, reporting findings to the public within five years.**
- **Since improved accuracy in eligibility determination depends more on improved measures of the key dimensions of mental retardation than on adjusting cutoff scores, the committee recommends research on improving measures, especially adaptive behavior assessment, and on methods to combine information on adaptive and intellectual functioning in making**

eligibility decisions based on a diagnosis of mental retardation.

- **SSA should make available for use by legitimate researchers tapes of Supplemental Security Income and Disability Insurance program utilization, comparable to public use tapes available for Medicaid program utilization.**
- **SSA should link its data on individual benefit awards to other agency data on health care and service costs for those same beneficiaries.**
- **SSA should examine data on eligibility determination procedures across its 10 districts, in order to discover if implementation of classification policies is consistent or varies regionally.**

Chapter 1

Introduction

Current estimates suggest that between 1 and 3 percent of people living in the United States will receive a diagnosis of mental retardation. This report assesses the process used by the U.S. Social Security Administration (SSA) to identify individuals with cognitive limitations who experience significant problems in their ability to perform work and may therefore be in need of governmental support. It evaluates the existing disability determination process in the context of current scientific knowledge and clinical practice. Mental retardation, a condition characterized by deficits in intellectual capabilities and adaptive behavior, can be particularly difficult to diagnose in the mild range of the disability.

SOCIAL SECURITY ADMINISTRATION DISABILITY PROGRAMS

SSA administers two disability programs that provide income and medical benefits to individuals who are either unable to work or to function as expected given their age because of disability. The Disability Insurance (DI) program, which operates under Title II of the Social Security Act, provides monetary payments to formerly employed individuals who have contributed to the Social Security trust fund through Social Security tax on earnings. Certain classes of dependents of insured individuals are also eligible for DI benefits. The Supplemental Security Income (SSI) program, which operates under Title XVI of the Social Security Act, provides payments to individuals (including children younger than 18 years of age) with a disability who have limited income and other resources. Such a person does not have to have been employed or a contributor to Social Security to be eligible for SSI benefits, although a means test is applied to both income and assets.

Disability is defined similarly for both the DI and SSI programs. For all persons applying under the DI program and for adults applying under the SSI program, disability is defined as "the inability to engage in any substantial gainful activity by reason of any medically determinable physical or mental impairment(s) which can be expected to result in death or which has lasted or can be expected to last for a continuous period of not less than 12 months" (Social Security Administration, 2002, p. 4). Standards are different for children younger than age 18. For this group, disability is defined as having a "medically determinable" physical or mental impairment, or combination of impairments that cause "marked and severe functional limitations." The impairment must be expected to lead to death or to be present for a continuous 12-month period. A medically determinable impairment is one resulting from anatomical, physiological, or psychological abnormalities that can be established by medical evidence that includes signs, symptoms, and laboratory findings; findings must go beyond the individual's subjective complaints.

Disability Determination Process

The application process for DI and SSI benefits can be a long and complicated one. Individuals (or their parents or guardians) may file for disability benefits by telephone, mail, or by visiting the nearest Social Security Office. The nearest Social Security Office can be identified by calling the toll free number (1-800-772-1213; 1-800-325-0778 for TTY calls). An individual can also call the toll free number and make arrangements to complete an application by telephone. SSA will then set up a time for the local Social Security Office to contact the individual and take the application over the telephone. SSA's website (www.ssa.gov) has helpful information about qualification criteria and procedures for filing disability benefits.

The in-person or telephone interview will be conducted by a claims representative of SSA. The claims representative will determine if the technical aspects of disability eligibility or entitlement are met. If the individual is eligible for benefits under the Title II program, or the Title XVI program, or both programs, the medical and vocational material is then forwarded to the state disability determination services (DDS) in the applicant's state. Medical information can include telephone numbers of doctors, hospitals, clinics, and institutions in which the person received treatment, as well as the dates of treatment; names of all medications currently being prescribed; medical records; and laboratory and test results. Vocational information can include summaries of previous work experiences and locations and a copy of a W-2 form.

Trained disability examiners and medical consultants, who will review all the pertinent information and make the determination of disability, staff the DDS. DDS staff may require additional information in making their determinations—for example, information directly from the applicant's treating clinician. DDS staff may require the applicant to undergo an examination conducted by SSA-hired consultative examiners.

DDS staff make the determination of disability based on the questions presented in Box 1-1.

The box outlines the five-step process for adults and the three-step process for children used to evaluate applicants for disability benefits. Beyond the stepwise determination process, there are several levels of appeal. There can be a hearing before an administrative law judge at SSA, review by SSA's Appeals Council, and ultimately review

Box 1-1
Sequential Evaluation Process for Disability Insurance and Social Security Income Benefits

Social Security regulations outline the sequential evaluation process used to evaluate applicants for disability benefits. Program rules further note that when an applicant is deemed not disabled at any one step, an evaluation under the subsequent steps is unnecessary.

For adults applying under the DI or SSI program, there is a five-step sequential evaluation process guided by the following questions:

- Step 1. Are you working? If you are working at the level of substantial gainful activity (defined as more than $740 per month in 2001), SSA will find that you are not disabled regardless of your medical condition or your age, education and work experience.
- Step 2. Is your impairment severe? A severe impairment(s) is defined as one that significantly limits physical or mental ability to do basic work activities. If your impairment(s) is severe, SSA proceeds to the next step.
- Step 3. Does your impairment(s) meet or medically equal a listed impairment? If so, and it meets the durational requirement, you will be found disabled. When an impairment(s) is not in the listings, SSA must decide whether it equals the severity of a listed condition. If the impairment(s) is severe but does not meet or equal a listing, SSA proceeds to the next step.
- Step 4. Can you perform past relevant work? If your impairment(s) prevents you from performing the physical and mental

by the federal courts. Clearly, both the stepwise determination process and the appeals process involve the use of objective and subjective indicators, as well as many judgments on the part of those making the determinations. Such a process inevitably results in errors, which can be of two types: they can incorrectly exclude someone who is legally entitled to benefits or can incorrectly result in benefits for someone who is not entitled to them. Since both types of errors are costly to the individual and to society, it is important to identify the sources of error and to minimize them. Some are inherent in the criteria for disability determination and the assessment instruments used to determine whether the client meets them.

demands of work that you have done in the past, SSA proceeds to the next step.

- Step 5. Can you do other types of work? SSA determines whether you can do work other than previous relevant work, considering your residual functional capacity, age, education, and past work. If you cannot, SSA will find you disabled.

For children and adolescents applying under the SSI program, there is a separate three-step sequential process. This process is guided by the following questions:

- Step 1. Are you working? If you are working at the level of substantial gainful activity (defined as more than $740 per month in 2001), SSA will find that you are not disabled regardless of your medical condition or your age, education, and work experience.
- Step 2. Do you have a medically determinable impairment(s) that is severe? Is you do not have a medically determinable impairment, or your impairment(s) is one that causes no more than minimal functional limitations, SSA will find that you do not have a severe impairment, and are, therefore, not disabled. If the impairment(s) is severe, SSA proceeds to the next step.
- Step 3. Does your impairment(s) meet, medically equal, or functionally equal the listings? If it does, and also meets the durational requirement, SSA will find you disabled. If it does not, SSA will find that you are not disabled.

Criteria for Mental Retardation

SSA disability determination for mental retardation requires that the individual have "significantly sub-average general intellectual functioning with deficits in adaptive functioning initially manifested during the developmental period; i.e., the evidence demonstrates or supports onset of the impairment before age 22" (Social Security Administration, 2002, p. 76). Children must also have significantly subaverage general intellectual functioning with deficits in adaptive behavior. Since they are children and under age 22, such findings will have manifested during the developmental period. The *Listing of Im-*

pairments, which specifies medical criteria and associated diagnoses, includes separate criteria for adults and for children and adolescents with mental retardation.

Listing 12.05 of Part A lays out criteria for mental retardation; it is closely paraphrased here. In order to be found eligible for benefits due to mental retardation, adults must be mentally retarded as defined above, and must meet one of four requirements:

- mental incapacity as evidenced by dependence upon others for personal needs (e.g., toileting, eating, dressing, etc.) and an inability to follow simple directions that is so severe that standardized measures of intellectual functioning cannot be administered;
- valid verbal IQ (VIQ), performance IQ (PIQ), or full-scale IQ (FSIQ) equal to 59 or less;
- valid VIQ, PIQ, or FSIQ between 60 and 70, and a separate physical or mental impairment that imposes an additional and significant limitation on work-related functioning; or
- valid VIQ, PIQ, or FSIQ between 60 and 70, along with at least two of the following: (a) marked restriction of activities of daily living, (b) marked difficulties maintaining social functioning, (c) deficiencies of concentration, persistence or pace that results in problems completing tasks in a timely manner, or (d) repeated episodes of decompensation.

Satisfaction of any one of these four criteria in an individual who has mental retardation meets the step 3 criterion of SSA's determination process; i.e., that the individual has a prima facie case of disability that results in an inability to work.

Separate determination criteria have been developed for children and adolescents, which recognize the different processes and effects that mental disorders have on their functioning. Determination criteria for children are further subdivided by age and associated developmental expectations. Criteria are provided for infants and toddlers (e.g., between ages 1 and 3) and three age groups of children and adolescents (e.g., ages 3 to 6, 6 to 12, and 12 to 18). These age criteria are designed to assess the severity of the disability's impact on the child's or adolescent's functioning, with benefits provided for conditions that cause "marked" restrictions, defined as "more than moderate but less than extreme." On standardized tests, a score that is "two standard deviations below the mean for the test" is evidence of a marked restriction. A score that is three standard deviations below the mean on a standardized test is evidence of an extreme limitation.

Medical criteria for evaluating children with mental retardation are described in *Listing 112.05.* Like the definition for adults, mental retardation in children for SSA disability purposes is characterized by significantly subaverage general intellectual functioning, with deficits in adaptive functioning. The *Listing,* again in paraphrase, includes six criteria for assessing severity of the condition:

- deficiencies in motor development, cognitive/communicative functioning, or social functioning for infants and toddlers; and for children and adolescents, deficiencies in at least two areas that include cognitive/communicative functioning, social functioning, personal functioning, or deficiencies in concentration, persistence, or pace that result in failure to complete tasks in a timely manner;
- a dependence on others for personal needs that is grossly in excess of age expectations, and an inability to follow directions that is so severe that standardized tests cannot be administered;
- valid VIQ, PIQ, or FSIQ of 59 or below;
- valid VIQ, PIQ, or FSIQ between 60 and 70 and a coexisting physical or other mental disorder that significantly impairs functioning;
- valid VIQ, PIQ, or FSIQ between 60 and 70 and, for infants and toddlers, the failure to attain development expectations for motor, cognitive/communication, and social functioning that is consistent with other children no more than two-thirds of their chronological age; for older children and adolescents, problems with cognitive/communicative function, social function, personal function, or problems with concentration, persistence, or pace; or

- failure of older infants and toddlers to attain motor, cognitive/communicative, and social milestones of children no more than two-thirds of their chronological age and another physical or other mental impairment that significantly impairs functioning; for older children and adolescents, problems with cognitive/communicative, social, or personal function or deficiencies in concentration, persistence, or pace that result in the failure to complete tasks in a timely manner and an additional physical or other mental impairment that significantly impairs functioning.

HISTORY OF MENTAL RETARDATION DEFINITIONS

The definition of mental retardation currently used by SSA differs from that used by other professional and health-related organizations. The concept of mental retardation, particularly a recognition that some portion of the population has cognitive deficits that significantly interfere with functioning, is an old one, although the ways in which this has been defined and measured have changed over time. Scheerenberger (1983) reports descriptions of the condition dating from 1500 B.C. in Egypt, in which disabilities of the mind and body due to brain damage were described.

Early definitions of the condition recognized differences in cognition that were associated with impaired functioning. In 1845, Esquirol (quoted in Scheerenberger, 1983) divided mental retardation into two primary categories based on performance on speech and language tasks. Seguin (1866) identified a severe defect in moral development as the primary characteristic of mental retardation.

According to Sheerenberger (1983), the major concepts common to current definitions of mental retardation were being used in the United States by 1900. These include onset during the developmental period (i.e., before age 22), intellectual deficits, and problems coping with the demands of everyday life. In its 1910 classification scheme, the progenitor of today's American Association on Mental Retardation (AAMR) (previously called the Association of Medical Officers of American Institutions for Idiotic and Feeble-minded Persons and the

American Association on Mental Deficiency) issued its first formal definition of mental retardation. AAMR defined persons with mental retardation as being feeble-minded, with development arrested at an early age or as evidenced by an inability to manage the demands of daily life or to keep up with peers (Committee on Classification of Feeble-Minded, 1910). Mental retardation was further divided into three levels of impairment: "idiots" had their development arrested at the level of a 2-year-old; "imbeciles" were at the developmental level of a 2- to 7-year-old; and "morons" were at the development of a 7- to 12-year-old.

Subsequent to the adoption of this definition, the field disagreed over whether mental retardation was a constitutional condition or one based on deficits in social competence (Biasini et al., 1999). Edgar Doll, for instance, proposed that mental retardation was a condition of genetic origin that resulted in social incompetence and arrested development (Doll, 1936a). He believed the condition was incurable. In contrast, Kuhlman (1920) proposed that the condition resulted from a subnormal rate of development, suggesting that it was a result of social functioning deficits rather than genetic conditions. Despite these differences in definition, however, they all focused on the inability to perform common behaviors, delays in social development, and low intelligence (Yepsen, 1941).

The 1959 AAMR definition was the first to integrate formally the measurement of intellectual capabilities and adaptive behavior functioning. This definition defined mental retardation as "subaverage general intellectual functioning which originates in the developmental period and is associated with impairment in adaptive behavior" (Jacobson, 1999). Subaverage intellectual functioning was defined as an IQ of 85 or less, with the developmental period extending only up to age 16. Deficits in adaptive behavior were a required part of the definition of the condition, even though there were no formal measures of the construct. AAMR recommended use of the Vineland Social Maturity Scale (Doll, 1953), with a subjective interpretation to be made by the evaluating clinician. A five-level classification scheme was also included for borderline (IQ 67-85), mild (IQ 50-66), moderate (IQ 33-49), severe (IQ 16-32), and profound (IQ <16) levels of retardation.

AAMR changed its definition in 1973, partly in response to concern about the inappropriate overidentification of minority students as mentally retarded. The new definition eliminated the classification of borderline retardation, and changed the upper criterion of scores on intelligence measures from 85 to 70 or below (Grossman, 1973). The result was a significant reduction in the numbers of children eligible for special school services and governmental supports. Levels of retardation were also redefined slightly.

AAMR's definition was revised again in 1977. This change suggested that IQs in the range of 70 to 75 might also be indicative of mental retardation if there were also significant deficits in adaptive behavior (Grossman, 1977). This change took into consideration the standard error of measurement on most tests of intelligence. In its most recent definition, adopted in 1992, AAMR has done away with the levels of retardation (American Association on Mental Retardation, 1992). The organization has also provided a list of 10 adaptive skill areas, with deficits in at least 2 of them required for a diagnosis of mental retardation. This current definition is discussed in more detail below.

The American Psychiatric Association, in its *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition* (DSM-IV), melded the 1977 and 1992 AAMR definitions, retaining the severity levels from 1972 and adopting a list of adaptive behavior areas similar to those chosen by AAMR in 1992. The DSM-IV also kept the upper limit of intelligence at equal to or less than 70.

It is important to note that the differences between the SSA definition of mental retardation and those of the major professional and health-related organizations derive from the purpose for which it is used. The SSA definition is used not for diagnostic purposes, but rather for purposes of program eligibility. The SSA definition fulfills its purpose of identifying individuals with cognitive limitations who experience significant problems in their ability to perform work and may therefore be in need of governmental support.

PREVALENCE OF MENTAL RETARDATION

In the General Population

There are widely disparate prevalence estimates of mental retardation in the U.S. population. Different studies report different rates depending on the definitions used, methods of diagnosis, and the particular population studied. For instance, the DSM-IV estimates the prevalence of mental retardation at 1 percent, although the basis for this number is not provided (American Psychiatric Association, 1994). This estimate is similar to that provided by other researchers (Hodapp & Dykens, 1996). Baroff (1991), using empirical sampling, estimates that 0.9 percent of the U.S. population can be presumed to have mental retardation. In a review of epidemiological studies, McClaren and Bryson (1987) report the prevalence of mental retardation at 1.25 percent, based on total population screening. Among school-age children, the U.S. Department of Education (1994) reports that prevalence estimates provided by different states in determining eligibility for special educational services ranged from 0.3 to 2.5 percent. In contrast, the U.S. surgeon general has estimated that some 7.5 million persons living in the United States have a diagnosis of mental retardation, representing almost 3 percent of the population.

The Centers for Disease Control and Prevention is conducting a longitudinal study called the Metropolitan Atlanta Developmental Disabilities Surveillance Program (MADDSP), which monitors the prevalence of developmental disabilities, including mental retardation, among children 3 to 10 years of age in the metropolitan Atlanta region (Boyle et al., 1996). The study used the definition of mental retardation listed in the *International Classification of Diseases, 9th Revision, Clinical Modification* (ICD-9) (World Health Organization, 1988), which includes severity ratings for mild, moderate, severe, and profound levels of retardation. Findings from the MADDSP indicate an overall prevalence of 8.7 per 1,000 children 3 to 10 years of age in Atlanta. Approximately two-thirds of all cases of retardation were of mild severity. Prevalence of mental retardation varied by age, ranging from 5.2/1,000 for children 3 to 4 years of age to 12.3/1,000 for children 9 to 10 years of age. Increases in prevalence were more likely to occur for children in the mild and moderate ranges of retardation, rather than in the severe and profound ranges.

These rates for children are similar to those reported in analyses of the National Health Interview Survey. Again, using the ICD-9 definition of mental retardation, Halfon and Newacheck (1999) reported unadjusted prevalence rates for mental retardation at 10.5/1,000 for children younger than 18. Analyses further indicated that the prevalence of mental retardation increased with age, ranging from 2.0/1,000 cases for children younger than 6, to 14.7/1,000 cases for children ages 6 to 12, and 15.7/1,000 cases for youth ages 12 to 17. The prevalence of mental retardation was also higher for males (13.0/1,000) than females (7.9/1,000).

Among different racial/ethnic groups, the prevalence of mental retardation was higher among black youth (16.2/1,000) than white (9.8/1,000), Hispanic (9.0/1,000) and other (6.4/1,000) youth. Prevalence rates are higher in some racial/ethnic groups partly because the responses to the National Health Interview Survey are provided by parents, who may have cultural reasons for concealing their child's cognitive disability. The correlation of low socioeconomic status and mental retardation is very high (see Chapter 2), and poverty rates are very high among black and Hispanic youth.

These prevalence estimates are vastly different, ranging from a low of 1 percent to a high of almost 3 percent. It is likely that the actual number of individuals with mental retardation ranges between 1 and 3 percent. This suggests that between 2.8 million and 7.5 million individuals could be diagnosed as having mental retardation.

In the Social Security System

Individuals with a diagnosis of mental retardation constitute a significant number of all recipients of SSA disability benefits. Data from SSA's *Annual Statistical Supplement* (2001a) indicate that as of December 2000, 567,151 persons with a diagnosis of mental retardation were receiving DI benefits, including 257,601 workers, 299,925 children age 18 or older, and 9,625 widows or widowers. Individuals classified as mentally retarded represented 10 percent of all workers with disabilities.

The SSI program has similarly high numbers of recipients with mental retardation. Among children receiving SSI benefits in December 2000, individuals with a mental retardation diagnosis constituted the largest diagnostic group at 32.8 percent (261,200 individuals). Consistent with other epidemiological data, the number of boys (162,230) outnumbered the number of girls (98,880) among children classified as having mental retardation (Social Security Administration, 2001b).

Data from December 2000 also highlight high numbers of adult recipients with mental retardation among SSI beneficiaries (Social Security Administration, 2001a). Among adult SSI recipients, individuals with mental retardation constituted 25.7 percent of all beneficiaries, representing the second largest diagnostic group of recipients ("all other mental disorders" was the largest at 34.4 percent). Males (52.5 percent) outnumbered females (47.5 percent) among SSI recipients with mental retardation.

ISSUES THAT PROMPTED THE CURRENT STUDY

The current study was designed to assess SSA's disability determination process for mental retardation. The committee was asked to examine new scientific opportunities and associated practice techniques to improve the current determination process. In addition, the committee has been asked to suggest new procedures to respond to these developments. Finally, this study will identify promising research opportunities that might help to clarify unaddressed or incompletely answered questions. SSA is most concerned about accurately diagnosing mental retardation among individuals in the mild range of retardation.

This report draws on recent research advances in the areas of cognition and learning, new information about the neural processes that occur during thinking and learning, and new knowledge about development of learned competencies, such as social skills and practical skills for everyday living. Furthermore, research from decision sciences has been examined to inform SSA's policy making. This study was prompted by a number of issues, including:

- the adequacy of SSA's definition of mental retardation,
- questions about intellectual assessment,
- issues raised by the assessment of adaptive behavior,
- thinking about how to combine data from intellectual and adaptive behavior assessments, particularly in the mild range of mental retardation, to accurately diagnose the condition, and
- issues related to distinguishing mental retardation from other diagnoses, particularly for children and adolescents.

Public Policy Implications

The diagnosis of mental retardation, as well as the receipt of benefits, has associated public policy implications. These policy issues relate to the context in which the program operates, as well as the impact of benefit receipt. Recipients get money to help with income maintenance, but they also get health care coverage through Medicaid. This health care coverage allows individuals with chronic medical conditions to receive needed treatment. In response to a number of issues, SSA changed the nature of its definitions of mental retardation and consequently the number of individuals receiving benefits. Any review of current practice has to consider that additional changes, while well meaning, may have negative effects on beneficiaries and the disability program itself. The committee has included an analysis of these issues in its assessment of the current determination system.

Adequacy of the SSA Definition

SSA's definition of mental retardation rests on subaverage intellectual functioning, either alone or in combination with other disabilities or functional impairments. The impairment must be present before the age of 22, although the diagnosis may be made at any time. This definition of mental retardation differs from that of several other organizations, including the American Psychiatric Association in its DSM-IV, the American Association of Mental Retardation, the World Health Organization in its *International Classification of Diseases and Related Health Problems, Tenth Revision* (ICD-10), and Division 33 (Mental Retardation and Developmental Disabilities) of the American Psychological Association.

DSM-IV defines mental retardation as significantly subaverage intellectual functioning (i.e., IQ no higher than approximately two standard deviations below the mean), which is accompanied by significant limitations in adaptive functioning in at least two of the following areas: communication, functional academic skills, health, home living, leisure, safety, self-care, self-direction, social/interpersonal skills, use of community resources, and work. Onset of these conditions must occur before age 18. The condition is further divided into four levels of retardation based on IQ, which include mild mental retardation (IQ between 50-55 and 70), moderate mental retardation (IQ between 35-40 and 50-55), severe mental retardation (IQ between 20-25 and 35-40), and profound mental retardation (IQ below 20 or 25) (American Psychiatric Association, 1994).

For AAMR, mental retardation is defined by substantial limitations in present functioning. It is defined as subaverage intellectual functioning that exists concurrently with deficits in two or more of the following adaptive skill areas: communication, self-care, home living, social skills, community use, self-direction, health and safety, functional academics, leisure, and work. The condition has to be manifested before age 18. AAMR further classifies mental retardation based on the nature and level of support needed by the individual, which can be intermittent, limited, extensive, or pervasive (American Association on Mental Retardation, 1992). These levels of support are not necessarily commensurate with the levels of retardation specified in the DSM-IV.

The World Health Organization (WHO), in its ICD-10, defines mental retardation as a "condition of arrested or incomplete development of the mind, which is especially characterized by impairment of skills manifested during the developmental period, which contribute to the overall level of intelligence, i.e., cognitive, language, motor, and social abilities." Guidelines suggest that an IQ of less than 70 is indicative of mental retardation. ICD-10 further suggests the use of culturally relevant, standardized measures of social maturity and adaptation for assessing functional abilities. No guidelines are provided, however, on cutoff scores for adaptive behavior deficits. ICD-10 allows for assigning a diagnosis of mental retardation when an individual has both mental retardation and a coexisting mental or physical disorder.

Division 33 of the American Psychological Association has defined mental retardation as "(a) significant limitations in general intellectual functioning; (b) significant limitations in adaptive functioning, which exist concurrently; and (c) onset of intellectual and adaptive limitations before the age of 22 years" (Jacobson & Mulick, 1996). Significant limitations for both intellectual and adaptive functioning are defined as two or more standard deviations below the population mean, using standardized assessment tools.

These definitions differ from those offered by SSA on a number of dimensions. The most significant differences are focused on the age of onset (e.g., 18 versus 22 years of age) and the nature of adaptive functioning deficits and how they are to be measured. WHO and Division 33 advocate the use of a standardized measure of adaptive behavior functioning, while AAMR and the American Psychiatric Association allow for descriptors of functional deficits across different domains. Even with this latter approach, however, the number of areas in which functional limitations must be present also differs. The American Psychiatric Association allows for deficits in at least 2 of their 11 areas, and AAMR specifies deficits in at least 2 of its 10 adaptive skill areas. In contrast, SSA does not specify the number of areas that need to be deficient, relying instead on "marked" restriction of activities of daily living and "marked" difficulties with social functioning. It is worth noting again that SSA uses its definition of mental retardation not for diagnostic purposes, but rather to determine legal eligibility for its benefit programs, in order to ensure that federal resources are used justly and correctly.

Intellectual Functioning and Its Assessment

SSA is similar to the other organizations in the level of intellectual impairment required to be present before a diagnosis of mental retardation can be assigned (i.e., IQ no higher than two standard deviations below the mean). For the other groups, however, that score has to be on the summary score attained on the intellectual functioning measure (e.g., equivalent to Wechsler FSIQ). SSA will also accept part scores from individually administered IQ tests, and specifically mentions Wechsler part scores as examples (e.g., VIQ, PIQ) in its regulations.

While SSA encourages the use of any standardized tests to determine intellectual and adaptive behavior functioning, it does not require these tests. It inadvertently gives preference to the Wechsler tests in its regulations by mentioning that the lowest of the overall summary score (FSIQ) and the two part scores (VIQ and PIQ) may be used in determining intellectual functioning. This not only cements a disparity among measures, without a solid empirical or policy basis, but also begs the question of whether one of these three scores provides the best relevant information. For instance, the FSIQ has higher reliability and validity coefficients than the two part scores. SSA needs to know if current practice and science support a policy of adjudicating on the basis of the lowest of multiple IQs; i.e., FSIQ, VIQ, or PIQ.

SSA further seeks to determine if its cutoff scores of 59 or less and 60 through 70 are also consistent with the current scientific literature on diagnosing mental retardation. The stringent upper limit fails to take into consideration the standard of error of measurement characteristic of all IQ tests. These basic assessment issues are further compounded when tests are administered to a culturally and linguistically diverse population. In some cases, instruments may not be available in a person's native language, or norming procedures may make the instrument inappropriate for use with some culturally and linguistically defined subpopulations.

It is important to know whether the major instruments in the field, such as the Wechsler scales and Stanford-Binet Test of Intelligence, adequately assess intelligence in a given case. If they do not, clinically acceptable and programmatically workable alternative instruments should be explored. This may entail identifying other instruments (including nonverbal intelligence assessment instruments as well as instruments available in languages other than English) that have sufficient reliability and validity to adequately diagnose mental retardation. Of course, any additional instruments identified should have the potential for wide use in clinical practice settings.

A number of research areas have produced reliable findings that are relevant and ready for implementation in practice. Advances in the assessment of developmental functioning have expanded the examination of intelligence from a dependence on verbal and performance intelligence scores to a broader view that incorporates measures of process as well as product. Multiple components that comprise intellectual functioning can now be more easily separated, for example, attentional processes, computational processes, problem-solving skills, and performance processes.

In the area of developmental assessment, standardized preschool measures of competence (Bayley, 1993) are required to assess multiple domains of functioning. These include fine motor, gross motor, cognitive, communication, and social skills. Impairment judgments based only on verbal and performance IQs may not reflect current intelligence testing practices for preschool children. The committee was charged with determining if other instruments better assess young children's intellectual functioning.

Adaptive Behavior Functioning and Its Assessment

For individuals with an intelligence score greater than 59, SSA requires documentation of deficits in adaptive behavior and functioning in order for a classification of mental retardation to be made, as long as no other serious medical condition is present. This may include the results of standardized tests (e.g., Scales of Independent-

Behavior-Revised, Vineland Adaptive Behavior Scales) or descriptions from parents, teachers, or treating clinicians. The use of either descriptive evidence or standardized test results seems to reflect variations in practice throughout the field of mental retardation. AAMR and the American Psychiatric Association both allow for behavioral descriptors of adaptive behavior in order to diagnose mental retardation. Division 33 of the American Psychological Association and WHO, however, suggest that results from standardized psychological assessments should be used (Jacobson & Mulick, 1996).

SSA has asked the committee for advice on how best to assess adaptive behavior for eligibility determination and award of benefits. With its current practice, SSA may frequently receive different kinds of information from different sources. The lack of standardization in the assessment of adaptive behavior may lead to a number of difficulties. Currently, claims may be adjudicated on the basis of different kinds of information, and trained lay examiners have the responsibility of combining data from different sources to try and sift out any evidence of deficits in adaptive behavior functioning, working in concert with medical consultants. The result may be inconsistent decision making and a time-consuming determination process.

Other important issues are how well major current measures of communicative, social, personal, motor, and community living skills identify and quantify deficits in adaptive behavior, and how well they meet current standards of reliability and validity. The committee explored alternative approaches that are clinically acceptable and programmatically workable, as well as measures that are currently used but are not satisfactory.

Combining IQ and Adaptive Functioning Data

Describing the nature of behavioral deficits becomes most needed—and most problematic—for individuals whose IQ is close to 70. Currently, SSA combines standardized IQ data with varying kinds of information on adaptive functioning. For adults, disability examiners consider adjustment in occupational and social settings; self-care is also a focus. For children, adjustment, including meeting developmental expectations, is a focus. There are, however, no guidelines about how to consistently combine these two kinds of information.

SSA is interested in the unique contribution each type of measure makes to the analysis of the adjustment of the individual to his or her world. A subset of this issue is the particular contribution of each to the evaluation and diagnosis of mental retardation in borderline situations, i.e., in cases in which the obtained IQ hovers near 70.

Current SSA practices allow a wide role for the qualitative evaluation of performance. Advances in the study of adaptive functioning have provided a more differentiated view of individual social and personal competence. Reviewing current thinking on the multiple dimensions of adaptive functioning could produce new models for such assessment or improve the utility of older assessments by identifying or updating appropriate cutoff scores for disability, or it may point to desirable and justified alterations of functional areas that are the focus of review in the eligibility determination process. The committee has reviewed the practice of allowing qualitative assessments of adaptive functioning, as well as instruments that provide standardized evaluations of adaptive functioning.

Differential Diagnosis

The issue of better differentiating mental retardation from other disorders that may have similar behavioral and cognitive manifestations—called differential diagnosis—is considered last. SSA has had particular difficulty distinguishing mental retardation from other disorders in children and adolescents. The *Sullivan v. Zebley* decision in 1990, in which the U.S. Supreme Court relaxed the criteria whereby children became eligible for SSI benefits; changes in determination criteria secondary to the Personal Responsibility and Work Opportunity Reconciliation Act of 1996; and the Balanced Budget Act of 1997 have all resulted in criticisms of SSA's attempts to distinguish mental retardation from other cognitive disabilities.

A number of conditions, such as autism, learning disabilities, borderline intellectual functioning, and some organic mental disorders, like traumatic brain injury, are associated with features that overlap with those seen in mental retardation. In addition, some genetic and behavioral disorders share features with mental retardation.

The committee has reviewed these other diagnoses, evaluating the signs and symptoms that are similar to those exhibited by individuals diagnosed with mental retardation, and is providing SSA with suggestions for better distinguishing these cognitive, genetic, and behavioral disorders from mental retardation.

Additional Research Needs

Finally, the process of evaluating scientific evidence generally reviews an area in great detail. The committee summarizes here its finding with respect to additional research that might improve the assessment and diagnosis of mental retardation. It is important to know what research needs to be conducted so that individuals with mental retardation can be better identified and can therefore have access to more appropriate services from education, health, and social service agencies. This question is designed to address the long-term needs of SSA and disability benefit recipients.

STUDY APPROACH

The committee and staff cast a wide net in examining the literature on mental retardation and its assessment. This approach was designed to gather information from a wide range of sources and assess the strengths and weaknesses of various pieces of evidence, with a goal of finding convergence of information from descriptive and inferential data and theoretical and conceptual frameworks.

Literature searches were conducted in peer-reviewed journals; technical manuals on intelligence and adaptive behavior measures were reviewed; papers were commissioned from experts on a number of topics central to the committee's work; and feedback was solicited from professional practice, advocacy, and other relevant groups. Members also reviewed technical and policy literature from SSA and other government agencies to get a better sense of the disability programs and benefits provided to individuals with mental retardation. To better understand the practical and policy implications of proposed recommendations on benefit receipt, the committee conducted statistical procedures called Monte Carlo simulations to examine the consequences of altering the criteria for scores on intelligence and adaptive behavior measures. In all of its review work, the committee focused in particular on the area of mild mental retardation, which is most problematic.

ORGANIZATION AND SCOPE OF THE REPORT

This report is focused on specifying criteria for the determination of mental retardation for SSI/DI eligibility purposes. It examines the contextual issues affecting SSA disability benefit programs, with Committee members recognizing that any evaluation of the current determination process for mental retardation is likely to have public policy effects. These effects are discussed in Chapter 2. SSA's charge to the committee posed several questions. The first, do current IQ tests adequately reflect widely accepted concepts of intelligence, is discussed in Chapter 3. The second asks how adaptive functioning is best defined and assessed; the committee's detailed response is in Chapter 4. SSA also asked about the relationship between measures of intelligence and adaptive behavior, which is covered in Chapter 5. Chapter 6, on differential diagnosis, explains how the conditions that share signs and symptoms with mental retardation are best distinguished from it. Suggestions for additional research that might shed light on any unaddressed or incompletely resolved issues in the field of mental retardation are mentioned throughout the text and are summarized in Chapter 5.

In focusing on specifying criteria for the determination of mental retardation for SSI/DI eligibility purposes, this report speaks to the intellectual and adaptive behavior criteria that should be used in making these determinations. The committee's findings, conclusions, and recommendations address initial eligibility determinations, that is, individuals who are first applying for disability benefits. For this reason, the committee has reviewed information that applies to all classes of potential beneficiaries, including children and adolescents and adults.

The committee has not explicitly addressed eligibility redetermination, the process of periodically recertifying eligibility for SSA benefits. The time interval for conducting redeterminations varies according to the medical condition. In general, they are scheduled every 7 years for individuals with conditions unlikely to change; every 3 years for conditions amenable to improvement; and as soon as 18 months for conditions likely to improve in the near future. Redeterminations for mental retardation are conducted every 7 years. Committee members view the standards for intellectual functioning and adaptive behavior assessment outlined in this report as applicable to the redetermination process as well.

Chapter 2

The Policy Context

This chapter examines the pivotal role of the two benefits programs of the Social Security Administration (SSA) in the system of supports for children and adults with mental retardation: Supplemental Security Income (SSI) and Disability Insurance (DI). This information is essential to set the context of the committee's conclusions and recommendations. Specifically, changes in the criteria for eligibility determination and the processes of enrollment for benefits affect not only the applicants, but also their families and the service systems in which they are embedded.

The chapter begins with a review of the prevalence of mental retardation in the SSI and DI programs. These data illustrate that chil-

---

This chapter is drawn largely from a paper written by Susan L. Parish, Ph.D., University of Wisconsin, and David Braddock, Ph.D., University of Colorado, for the committee.

dren and adults with mental retardation are a large group of recipients, particularly in the SSI program. Prevalence data on the severity of mental retardation, from mild to severe, in this program (and in the DI program) are not available. We then note that these programs serve as a gateway to other critical federal and state benefits—most notably, Medicaid and Medicare. The linkage between SSI and DI benefits and other service programs underscores the importance of such eligibility for people with mental retardation and also raises the political, social, and economic stakes regarding any changes in eligibility criteria and determination procedures. We also discuss the enduring problem of unemployment and underemployment of adults with mental retardation, a fundamental issue undergirding the important role of federal benefits for this population.

SSI AND DI RECIPIENTS

According to the Office of Research, Evaluation, and Statistics of the Social Security Administration, in December 2000 almost 850,000 children were receiving SSI benefits. These children represented 12.8 percent of the over 6.5 million SSI recipients. Almost two out of three (62.5 percent) of these children had a disability based on a mental disorder, and most of them (32.8 percent of all children, or 261,200 individuals) were eligible because of mental retardation. The incidence of mental retardation as the primary diagnosis for young recipients increased with age, from 3 percent of those under age 3 to 46 percent of those ages 13-17.

Data on adult recipients of SSI as of December 2000 indicate that of the 226,792 recipients between ages 18 and 21, 50 percent were eligible because of mental retardation. The percentage of those receiving SSI due to mental retardation by age group gradually decreases with advancing age, from 45.7 percent of those ages 22-29, 33 percent of those ages 30-39, 18.6 percent of those ages 40-49, 10.7 percent of those ages 50-59, and 6.9 percent of those ages 60-64. This pattern likely reflects the increasing numbers of people with other disabilities participating in SSI, rather than a decrease in the number of beneficiaries with mental retardation.

Using estimates from December 1999 data, Parish and Braddock (2001) calculated that SSA paid out $20.64 billion in SSI benefits to 4.54 million people with disabilities in the United States in fiscal year 2000. People with mental retardation constituted 23 percent of this total and received 32 percent of the SSI payments. People with mental retardation constituted 9 percent of all people with disabilities who received DI payments, but they received only 6.5 percent of total payments. The contrast in the prevalence of people with mental retardation in the SSI program (23 percent of program recipients) and the DI program (9 percent of program recipients) underscores the fact that most people with mental retardation are not in the labor force, or are not in the labor force for the amount of time required to qualify for DI benefits. Table 2-1 shows the range and scope of benefits paid to recipients with disabilities, and specifically those with mental retardation, during fiscal year 2000.

Growth in the SSI population, particularly among children and adolescent recipients, has been a subject of contention during the past two decades. This growth may be attributed to expansion in eligibility due to legislative and regulatory changes, SSA's increased outreach efforts, limited return-to-work efforts, the increase in the number of noncitizens on the rolls, medical advances enabling people with disabilities to live longer, and state efforts to transfer people with disabilities from state-supported public assistance to SSI. In addition, the need to continue Medicaid coverage has been cited as a reason for retaining people with disabilities on SSI (Mashaw & Reno, 1996; U.S. General Accounting Office, 1995). The number of individuals with mental retardation who receive SSI grew at an average annual rate of 9 percent between 1986 and 1993. In spite of this growth, however, people with mental retardation constituted a fairly level 25 percent of total SSI recipients during the period (U.S. General Accounting Office, 1995), indicating that growth has been as fast for beneficiaries with other disabilities.

TABLE 2-1 Benefits Paid to Persons with Disabilities, Fiscal Year 2000

A. Supplemental Security Income (SSI)

	All Persons with Disabilities		Persons with Mental Retardation	
	Benefits Paid	Recipients	Benefits Paid	Recipients
SSI children	$4,439,660,478	847,063	$1,757,927,424	286,121
SSI adults	$16,203,653,578	3,690,970	$4,770,256,896	776,409
Total	$20,643,314,056	4,538,033	$6,528,184,320	1,062,530

B. Disability Insurance (DI)

	All Persons with Disabilities		Persons with Mental Retardation	
	Benefits Paid	Recipients	Benefits Paid	Recipients
DI	$44,136,908,784	4,873,560	$1,634,153,804	249,871
Adults disabled as children	$4,320,897,000	726,250	$1,515,421,383	275,975
Widows/ widowers	$1,184,645,952	197,520	$53,393,300	9,876
Total	$49,642,451,736	5,797,330	$3,202,968,487	535,722

SOURCE: Braddock, D. L., Hemp, R., Parish, S. L., & Rizzolo, M. C. (2000). *The state of the states in developmental disabilities: 2000 study summary.* Chicago, IL: University of Illinois at Chicago, Department of Disability and Human Development. Reprinted with permission.

Growth in the DI and SSI programs has been extensively analyzed, and researchers have offered various reasons for this growth. Applications for disability benefits, DI in particular, have long been directly related to the nation's overall economy. When the economy is struggling and jobs are scarce, applications for benefits tend to increase (Stapleton et al., 1998). The major reasons for growth in both applications and awards for disability benefits are related to policy changes made in the programs and to the changing business cycle, as opposed to increases in the prevalence of disability (Burkhauser, 1998). In their comprehensive analysis of the impact of an array of population-based, economic, policy, and administrative factors on the growth of disability benefits, Stapleton and his colleagues (Stapleton et al., 1998) found that the growth of applications and awards could be traced to three major causes: (1) a poor economy; (2) action by states to shift welfare spending from shared state-federal (e.g., Aid to Families with Dependent Children, AFDC) to federal (e.g., SSI and DI) programs; and (3) so-called supply changes, including SSA's outreach activities, the increasing value of benefits, work incentives, eligibility determination changes brought about by legislative and administrative activities, and court decisions on appealed cases. There is widespread agreement that the reasons for the growth in the number of beneficiaries are complex and varied (Burkhauser, 1998; Stapleton et al., 1998).

SSI AND DI IN THE CONTEXT OF FEDERAL SOCIAL SERVICE PROGRAMS

An intricate web of social services and income support programs in the United States targets diverse groups, including people who are elderly, are poor, have disabilities, or are veterans. Benefits include income transfers, food stamps, health insurance, transportation, case management, vocational training, and support for housing and education.

All of these programs are characterized by differing eligibility criteria, benefit levels, administering bodies, critical definitions (e.g., of disability), funding mechanisms, and application procedures. A person may be presumptively eligible for one program, due to his or her diagnosis, as in special education for students with mental retardation, but not eligible for other programs without participating in a lengthy and complex application procedure, as in SSI, Medicaid, and vocational rehabilitation. The complexity of this network of services and supports is heightened by the fact that some programs are funded and administered by the federal government, others by the states, others by local authorities, and others jointly among different levels of government and variously administered. This dizzying array of possibilities contributes to a service system that is extraordinarily fragmented, making it difficult for citizens to navigate.

People with disabilities, particularly mental retardation, experience great difficulty in enrolling in a program such as SSI. Considerable savvy (i.e., intellect, persistence, and ability to understand nuances) must be brought to bear to ensure that a person with disabilities gets all the services and supports to which he or she is entitled or requires to be productive. People with mental retardation face an immediate barrier in the form of demonstrating eligibility, a barrier that may be insurmountable without an advocate to pursue access to benefits for them. People with mild mental retardation whose disability is generally less visible, are less likely than people with more severe mental retardation to have an advocate to assist them in the application process.

To assess the impact that other federal programs can have on the SSI and DI programs, it is important to understand the ways programs interrelate. Researchers have recognized that some programs are complementary, while others are substitutes for one another. For example, Medicaid was considered to be complementary to the now defunct AFDC; the expansion in Medicaid benefits led to concomitant increases in AFDC participation. Researchers have also found that public health insurance is probably a substitute for private health insurance. When eligibility for Medicaid was expanded to cover children and pregnant women in the late 1980s and early 1990s, there was a corresponding reduction in private health insurance coverage (Garrett & Glied, 2000).

Given the complex interrelationships among federal programs for people with disabilities, it is useful to assess how programs interact with respect to people with mental retardation. In an evaluation of the impact of state generosity in providing AFDC benefits on child SSI participation, Garrett and Glied (2000) found that following the U.S. Supreme Court's *Zebley* decision, SSI participation by children increased more in states with lower AFDC benefits. This occurred because SSI payments are generally higher than AFDC payments, and AFDC's earned income rules were more stringent than those for SSI.

After the *Zebley* ruling, families were more likely to seek SSI and Medicaid for their children with disabilities. The welfare reform legislation of 1996 was expected to precipitate a fall in the SSI rolls due to the tightening of eligibility criteria for children with disabilities. In 1997, the eligibility of hundreds of thousands of child SSI recipients was reevaluated by SSA, and 90,000 children were removed from the program (Kubik, 2000). Similarly, cost containment processes have become a hallmark of private insurance companies, which are increasingly requiring SSI determination of disability before accepting the adult child of a policyholder as an adult dependent with a lifelong health condition (Hemingson, 1998; United Health, personal communication, 1998). For insurance companies, this practice eliminates many prospective insurees who have long-term health needs because of the long and difficult process it takes to work with two major, unconnected systems. Because of the length of the SSI determination process, some insurance companies save the costs of other prospective insurees with long-term health care needs by delaying their obligation for one, two, or even three years. SSI and DI eligibility determination can take a full year or more and extend even longer if appeals are necessary. This is an extraordinary burden for poor families and prospective beneficiaries who need health care.

Research has also assessed the extent to which Medicaid has been an incentive for families to continue to receive AFDC benefits. Yelowitz (1998) found that when Medicaid eligibility was expanded to include greater numbers of poor children, their parents were more likely to enter the workforce and discontinue participation in AFDC. This finding is important because it points to the potential for savings in welfare expenditures when Medicaid eligibility is expanded. Families with a child with a disability may also be able to eliminate their SSI use if Medicaid coverage could be continued. As observed by the National Commission on Childhood Disability in compiling its report to Congress, many families commented that the need to purchase health- and therapy-related items was a critical factor in their ongoing need for SSI (National Commission on Childhood Disability, 1995). Expansion of Medicaid eligibility is less likely to affect SSI enrollment among adults with mental retardation, however. In light of the very low competitive employment rates for adults with mental retardation, their access to private health insurance coverage is very limited. And given the reliance of state service systems on Medicaid reimbursement for community-based services for adults with mental retardation, strong incentives remain for enrollment of such adults in the SSI and, by extension, Medicaid programs.

Analyses of the relationships between SSI applications and other federal support programs reviewed for the years 1988-1992 demonstrate that reduction in welfare benefits for the poor resulted in highly significant increases in SSI applications in seven states and the District of Columbia (Stapleton et al., 1998). Although analyses of the interactions between SSI and the recently enacted Temporary Assistance to Needy Families (TANF) program have not yet been conducted, one can speculate on the potential relationship. Given that there is a five-year maximum lifetime benefit period mandated for TANF beneficiaries, it is likely that former welfare recipients will have increased incentives to apply for SSI. While certainly not all mothers who live in poverty have mental retardation, essentially all mothers with mental retardation are poor (Keltner, 1992). Therefore, the relationship between public assistance and mental retardation is a significant one. There is evidence that a substantial share of the mothers who were on AFDC have disabilities (Stapleton et al., 1998). And, because most mild mental retardation has an intergenerational nature associated with varied deprivations (diet, cognitive stimulation, substance abuse), many women and their children who are approaching the term limit for TANF may have mental retardation. The intergenerational nature of mild mental retardation may significantly complicate SSI determination, when both mother and child have mental retardation. The

SSA application process may be a formidable one for applicants with a cognitive disability.

Women with mental retardation constitute a significant proportion of individuals who have been supported by welfare programs (Stapleton et al., 1998). Because of their lifelong disability, work opportunity or vocational training for women with mental retardation may have limited success, unless it is responsive to the special training needs associated with their disability. People with mild mental retardation can certainly be employed. However, women who are also parents (usually alone) and must provide health care for dependents will struggle to fulfill their vocational potential in this context without the proper supports—supports that have not been included in the TANF program. And maximum enrollment in TANF expires in most states between August 2001 and August 2002.

The structure of the nation's economy has fundamentally changed over the last few decades, and the impact on people with disabilities has been significant. A rise in service businesses, as well as an increase in skilled work that is computer-reliant has accompanied the decline of the nation's manufacturing jobs. Although technological advances have offered new opportunities to skilled workers with physical disabilities, outcomes for people with cognitive disabilities like mental retardation have been less auspicious, given new emphasis on higher levels of education, flexibility, speed, and intellect (Krueger, 1997; Mashaw & Reno, 1996).

The economic outlook for people with disabilities, particularly cognitive ones, is bleak. That is, there is good evidence that, in spite of the Americans with Disabilities Act of 1990, they did not benefit in large scale from the economic boom of the 1990s. As noted by Burkhauser et al. (2001, p.294):

> We find that while economic expansion since 1992 has improved the economic well-being of most working age people, the gains have been much smaller for working age people with disabilities than for the rest of the working age population. Furthermore, although the gains through 1998 have returned the average person with a disability to his or her level of economic well-being in the peak year of the 1980s business cycle, the income of a large fraction of the population with disabilities in 1998

> remained well below that of their 1989 counterparts in absolute and relative terms. Finally, we find that despite a robust economic expansion, men and women with disabilities became less involved in the labor market and more dependent on public income transfers during the 1990s.

This situation has not yet been examined systematically to determine whether it is a consequence of the labor market requiring workers with better cognitive skills or whether potential workers with mental retardation did not seek employment for fear of losing their only available health insurance: Medicaid. Indeed, considering the significant policy and societal impact of the context for persons with mild mental retardation who are affected by these changes, there is a surprising lack of research and empirical resources to describe these phenomena.

IMPORTANCE OF SSI AND DI BENEFITS IN SERVICE DELIVERY SYSTEMS

Gateway to Other Services

SSI and DI benefits serve as a critical gateway to a number of other federal and state services. Recipients are able to get health care coverage and income maintenance benefits. Recipients are also able to get vocational training and other kinds of employment assistance. Here we briefly review some of these benefits that come with SSI and DI.

In the United States, eligibility for public health insurance for people with disabilities is predicated on eligibility for income transfer programs—essentially on poverty status. In 32 states, eligibility for SSI automatically qualifies an individual for Medicaid. Seven other states use SSI eligibility to establish Medicaid eligibility, but they require a separate application. In the remaining 11 states, Medicaid eligibility requirements include criteria that are more restrictive than those of the SSI program. Recipients in those states do not automatically qualify for Medicaid when they qualify for SSI. Everyone who receives DI is eligible for Medicare, although there is a two-year wait before Medicare coverage begins (Social Security Administration, 1999). And 34 states supplement federal SSI benefits with cash assistance ranging from $10 to over $250 per month for an individual living independently (State Assistance Programs for SSI Recipients January, 2000; online at http://www.ssa.gov/statistics/ssi_sap/2000/). States that supplement SSI benefits have a significantly higher rate of SSI applications, a finding that is robust over the period between 1980 and 1993 (Stapleton et al., 1998).

Current federal policy permits states to offer some home and community services as part of Medicaid state plan benefits through the Home and Community Based Services waiver program. The program does not fund costs for housing. Thus, federal SSI payments and state supplements have become increasingly critical sources of funding for community services, especially housing costs. Braddock et al. (2000) report that in 1998, federal SSI payments of $1.3 billion represented 7 percent of the nation's total publicly financed community services revenues for people with mental retardation or developmental disabilities. SSI income is considered by lenders to be a more stable source of income for purposes of purchasing a home than income from most of the jobs that people with mental retardation can hold. When accounting for the fact that these jobs also typically do not provide benefits such as health insurance, SSI becomes, again, a gateway for basic security separate from its income maintenance purpose.

Some of the services currently provided for people with mental retardation or developmental disabilities in the United States include housing, case management, employment, personal assistance, transportation, family support, and vocational training. In 1998, spending for this specialized service system totaled $25.6 billion, with Medicaid representing the largest single source of funding. Combined federal, state, and local Medicaid resources totaled $19 billion, comprising 74 percent of spending on people with mental retardation or developmental disabilities in 1998 (Braddock et al., 2000).

In 1998, the estimated population of people with mental retardation in the United States was about 2.35 million (Larson et al., 2000).

The formal service system funded out-of-home placements for nearly 417,000 individuals (almost all of whom are adults), or about 18 percent of the estimated total population with mental retardation. Of these, 238,000, or 57 percent, were living in settings of six or fewer people. The remainder were living in larger group homes or institutions. Spending exclusively for residential services in 1998 totaled $16.5 billion (Braddock et al., 2000), and represented 64 percent of total services spending for people with mental retardation or developmental disabilities in the United States (Braddock et al., 2000). Figure 2-1 shows proportionate spending for these services in the United States in 1998.

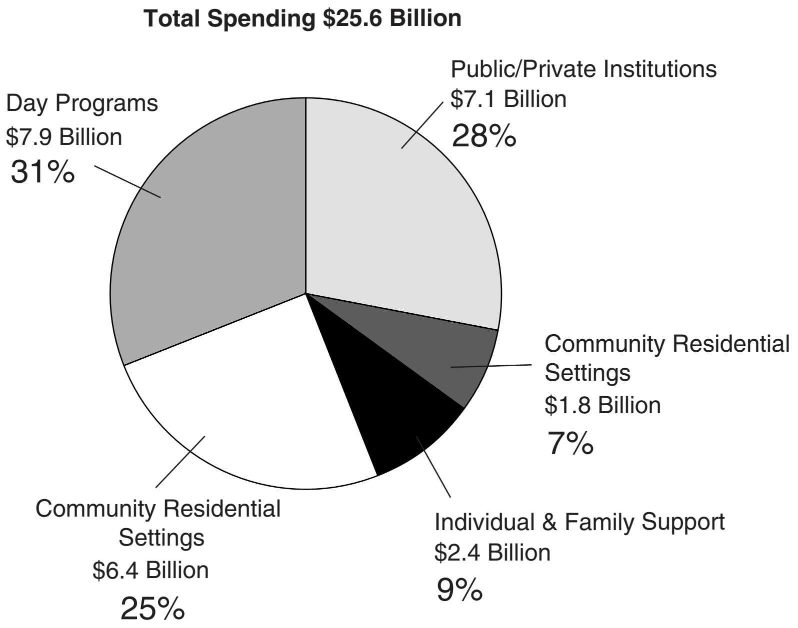

FIGURE 2-1 U.S. spending on MR/DD services by setting in 1998. Day programs include sheltered workshops, day care, transportation, case management, and other nonresidential community services. NOTE: Public/private institutions have 16 or more residents; large community residential settings between 7 and 15; small community residential settings 6 or fewer. SOURCE: Braddock et al. (2000). Reprinted with permission.

Family support services were provided to an additional 328,000 families, representing 14 percent of the estimated population of people with mental retardation. Family support services include respite care, cash assistance, and an array of in-home and out-of-home supports. Family support services spending in 1998 totaled $736 million, and represented 3 percent of total spending for this population across the states (Braddock et al., 2000). There are other public and private agencies that provide services to children and adults with mental retardation, although no national estimate on the full range of service utilization across various sectors is available.

Available estimates of publicly supported service use suggest that the vast majority of children and adults with mental retardation are not reached by the formal service delivery system. Their eligibility for and receipt of SSI (and Medicaid) thus constitute a critical component of society's safety net for this vulnerable population. Many adolescents and young adults, especially those with mild mental retardation, leave formal service systems when they complete or drop out of school. This "invisibility" of persons with mild mental retardation varies according to whether or not they were identified as needing any special education services (regardless of label) during their school years. If they were never identified, they are far less likely to know about SSI, or to be referred, screened, or identified after high school.

However, if individuals with mild mental retardation received any services from special education during middle or high school, they should have had an individualized education program (IEP). For adolescents 14 and older, the IEP must include goals pertaining to transition from high school to the worlds of work and community. Transition planning is mandated by the Individuals with Disabilities Education Act of 1990 (IDEA) and the IDEA Amendments of 1997. At least 200,000 young people with disabilities exit the school system each year and enter the adult world (Wehman, 1996). Regardless of an adolescent's specific disability, discussion of work-related issues should, theoretically, include mention of SSI. Of note is the fact that high school dropouts are unlikely to continue to pursue their IEP or transition goals, and individuals with borderline or mild mental retardation may be at especially high risk of dropping out. Thus, adolescents who drop out of high school will no longer have the support of the special education system, and they may exist outside a formal service delivery system. One challenge for SSA, then, is to develop more proactive "search and find" procedures that encourage young adults with mild mental retardation who have exited the school system to contact SSA. Such an approach should also be sensitive to the many youths who are reluctant to self-identify as mentally retarded or as in any way disabled.

Role in Reducing Poverty

The SSI and DI programs play a vital role in reducing poverty for recipients. It is important to recognize that the financial status of people with disabilities is considerably worse than that of elderly Social Security beneficiaries. A total of 32 percent of workers with disabilities are poor or near-poor, compared with 17 percent of Social Security beneficiaries (Grad, 2000). Two-thirds of SSI beneficiaries receive at least half of their total income from SSI, and Social Security is responsible for reducing the poverty gap for SSI recipients by an average of 60 percent (Grad, 2000).

For people with mild mental retardation, like all people, basic security and quality of life are context sensitive. Their abilities to secure benefits to which they are entitled, to acquire and maintain adequate health care for themselves or their families, and to participate in job training all rely on social structures that can accommodate their disabilities. The nature of mild mental retardation is such that these individuals may be more likely to be found ineligible than others during redetermination, and more likely to be confused about an appeals process. We hypothesize that eligibility redeterminations for this population are also more likely to be incorrect because of incomplete information provided by the applicant, lack of understanding of certain probes in the evaluation process, and the personal sensitivity and reluctance to present oneself as "mentally retarded." The complexity of the application, reapplication, and appeals processes are an enormous barrier to people with cognitive disabilities. The processes for application and appeal are sufficiently complex and extended that they may put people correctly diagnosed with mental retardation at a severe disadvantage in obtaining the benefits to which they are legally entitled.

People with mental retardation are often perceived to be living at the margins of society, and nowhere is this more evident than in an examination of their financial status. The connection between disability and poverty goes back several millennia in the Western world (Parish & Braddock, 2001), and people with disabilities have often faced debilitating poverty.

Recent research has shown that people with mental retardation are substantially more likely to be living below the poverty level than are other people. An analysis of the National Health Interview Survey—Disability Supplement (NHIS-D) indicated that 33 percent of adults with mental retardation are living below the poverty level, compared with 13 percent of those without mental retardation. For children ages 0-5 years, 33 percent of those with mental retardation live in poverty, compared with 22 percent of children without mental retardation (Larson et al., 2000). The complex nature of the relationship between poverty and disability is not wholly understood. Researchers have found that the risk of disability appears to be growing, and that risk is elevated in poor homes and in single-parent families (Fujiura & Yamaki, 2000). There are numerous risks associated with poverty, including poor pregnancy outcomes, child neglect and abuse, substance abuse, violence, limited access to health care, and reduced quality of life. These findings suggest that a substantial portion of the population with mental retardation is grappling with an array of issues related to poverty, in addition to the difficulties they face due to their disability. Because SSI is income based, it is intended to spare citizens from extreme poverty and constitutes the nation's basic provision of a safety net, providing a minimum level of support for people who are often marginalized in the national economy and society.

Health Care Linkages

In 1993, 10 percent of Medicare's total enrollment consisted of DI beneficiaries, with the remaining 90 percent receiving Medicare due to their status as elderly. Not surprisingly, given the generally higher costs of health care for people with disabilities, DI beneficiaries consumed slightly more than their proportionate share of Medicare program costs, at 12 percent of spending (Chirikos, 1995).

Adults with disabilities comprised 16 percent of all Medicaid recipients. Again, however, health care costs for people with disabilities were higher than costs for the rest of the Medicaid population. Spending for recipients with disabilities constituted 36 percent of total Medicaid spending in 1996 (Iglehart, 1999), in great part due to disproportionately higher expenditures for long-term residential care. Medicaid spending for children with disabilities is considerably higher than that for children without disabilities, as well. In 1998, 7 percent of child Medicaid beneficiaries had disabilities, but spending for these children consumed 27 percent of Medicaid spending for children (U.S. General Accounting Office, 2000).

The linkages between health and poverty are widely discussed in the research literature and in the ongoing public debate about viable initiatives to address child welfare issues (Fujiura & Yamaki, 2000). The importance of health insurance in the lives of poor children has been well established by the research community. Children who are eligible for Medicaid use significantly more medical care, particularly care provided in doctor's offices. In addition, Medicaid eligibility has been associated with significant reductions in child mortality (Currie & Gruber, 1996a). Rates of infant mortality in the United States are the highest in the developed world, just one of a number of indicators that poor children in this country are not receiving the quantity or quality of health care that is available to children in other developed countries (Currie & Gruber, 1996a). Increases in the Medicaid eligibility of pregnant women have also led to better prenatal care and birth outcomes for their children (Currie & Gruber, 1996b). Technology has increased survival rates among children with very low birth weight, a group at risk for mental retardation, who disproportionately come from economically disadvantaged families.

The link between SSI eligibility and eligibility for health insurance is a powerful one. Low-income families who have children with mental retardation have two "choices" to ensure adequate health insurance coverage: to become poor or to stay poor. Over 4,000 letters have been compiled from families who report not accepting raises, overtime, or promotions in order to keep the SSI benefit that ensures health care coverage for their children with mental retardation (C. Garner, personal correspondence, March 14, 2001). Many families who have middle and low-middle incomes (about $30-50,000) also have reported that some of the ways they accommodate the special needs of their children with mental retardation include second mortgages and credit card debt. Finally, some parents resort to relinquishing custody of their children in order to obtain basic services that they need (C. Garner, personal correspondence, March 14, 2001). Two other federal programs intended to provide health care for children with special needs are not available to all children with mental retardation. The Medically Needy Program, intended as a safety net, is not operational in all states. Among those states that do have the program, the numbers of awards are capped at very low levels. West Virginia, for example, caps this program at 150 children. The State Child Health Insurance Program (CHIP) is an outreach program designed to provide health care for children, but two-thirds of the states with CHIP programs have restrictions limiting access or services.

The use of preventive health care for children is influenced by financial factors, including insurance coverage and availability of free care, as well as by sociodemographic characteristics. Disadvantaged children generally receive lower levels of preventive care, including immunizations and well-child visits (Herz et al., 1998). Children who receive SSI have some of the lowest rates of utilization of preventive health care among all child Medicaid recipients (Herz et al., 1998). Although the reasons for this low rate of well-child care usage are unclear and complex, it is certain that these children, given the health care needs associated with disability, are likely to need greater levels of intervention than are other children.

The Early and Periodic Screening, Diagnostic, and Treatment (EPSDT) program for children is a mandatory Medicaid program enacted in 1967. The entitlement to these services was established by the Omnibus Budget Reconciliation Act of 1989, which expanded EPSDT requirements for states. EPSDT also provides preventive care and referrals for other health problems. Preventive services are available to children through non-EPSDT Medicaid services. The health insurance coverage offered to children under the auspices of the Medicaid program is significantly more comprehensive than the coverage typically offered by private plans. Child Medicaid beneficiaries have access to preventive care, diagnostic and evaluative treatment through the EPSDT program, and medically necessary therapies including mental health care, home health care, and office-based services (Fox et al., 1997).

There is considerable evidence that receipt of SSI, and of the Medicaid health insurance that accompanies it, is inadequate for families trying to meet disability-related costs of caring for their children. Families often report that certain of their children's needs go unmet, due to limitations on their ability to pay for services (U.S. General Accounting Office, 1999). Furthermore, eligibility rules for SSI are biased toward placing children in institutions, rather than keeping them with their families. Some families with incomes high enough to prevent them from obtaining SSI, and therefore health insurance, may seek to institutionalize their children with disabilities in order to receive Medicaid. In fiscal year 1998, there were approximately 8,200 children with disabilities living in institutions in the United States and receiving SSI (Scanlon, 2000).

Employment and People with Mental Retardation

Given that eligibility for SSI and DI involves a two-stage test, requiring (1) the documentation of a medically determinable impairment and (2) that such impairment results in an inability to engage in substantial gainful activity, we can draw a number of conclusions from available information regarding eligibility for adults with mental retardation.

First, mental retardation is highly associated with unemployment. Although many people with mental retardation work, their overall employment level is low. Approximately 85 percent are not employed in competitive jobs; studies of the employment rate of people with mental retardation consistently show very low rates of employment, high rates of layoffs, and very low wages. A recent analysis by Yamaki (1999) indicated that the employment rate for women with mental retardation was 23.5 percent, while the employment rate for men with mental retardation was 27.4 percent. The employment rate of people with disabilities overall ranges between 23 and 45 percent (Daniels & West, 1998), indicating that people with mental retardation have one of the lowest employment rates of any demographic group in society, including other people with disabilities. Even for those who do work, levels of compensation frequently are close to the level they would receive as Social Security beneficiaries. In addition to low levels of employment, people with mental retardation have very low rates of competitive employment, as opposed to employment in sheltered workshops or supported employment positions. Estimates of the employment rate for people with mental retardation suggest that only 12-15 percent of people with mental retardation are employed in jobs in the community (Mank, 2001). Therefore, it is safe to assume that the great majority of people with mental retardation do not meet the employment criteria to receive Social Security benefits.

Research on the earnings of people with disabilities has identified two sets of factors that contribute to their lower wages. One is evidence that productivity rates are affected by the type and severity of the individual's disability. The other is evidence that prejudice in the workplace, particularly by employers, contributes to lower wages (Baldwin & Johnson, 1998). People with mental retardation are subject to higher levels of prejudice than are people who acquire their disabilities as adults, or than those who have visible and nonintellectual disabilities. The severe prejudice and discrimination faced by people with mental retardation makes the impact of disability on work particularly strong (Baldwin & Johnson, 1998).

Second, people with mental retardation who are employed are more vulnerable to layoff due to changes in the work environment or the economy. Due to the nature of mental retardation, people with this disability have more difficulty adapting to changes in the workplace that necessitate acquisition of technical skills or academic instruction. Furthermore, if a person with mental retardation loses his or her job due to changes in the economy, he or she is less able to transition to another job easily. This may reflect a number of factors, including a lack of job search or interview skills, transportation problems, or the diminished availability of jobs of an entry or nontechnical nature generally.

In light of the low employment rates of people with mental retardation, it is also likely that they will be particularly hard-hit by the implementation of welfare reform initiatives, especially among those who are not currently identified as having mild mental retardation. These individuals are likely to be working in jobs at the fringes of the economy, are likely to be new workers and lack seniority, and will be among those hardest hit in an economic downturn.

Third, the employability of people with mental retardation is dependent on the existence of organized, available, and accessible vocational supports and services. There is increasing evidence that many if not most people with mental retardation could be capable of some degree of employment with the appropriate support structures (Mank, 2001). These supports include services provided by local and state educational agencies, vocational rehabilitation agencies, mental retardation agencies, community rehabilitation agencies, and public transportation. People who have access to quality education programs and later to supportive work environments can work to an extent that would be impossible otherwise.

Fourth, large numbers of people with mental retardation do not have access to the support systems they need to obtain and maintain employment. Although educational services are improving, the extent to which employment-related provisions of the IDEA have been implemented varies substantially among local educational agencies and among the states. Furthermore, failure to implement appropriate transition planning and to carry out transition strategies is one of the most frequently cited areas of noncompliance identified by federal monitors of IDEA compliance (National Council on Disability, 1996).

The availability of vocational rehabilitation services for people with mental retardation is relatively low in relationship to the overall number of people in need. Vocational rehabilitation departments exist in every state, and they spend joint federal-state funds to provide an array of employment and training services to people with disabilities. Vocational rehabilitation use and impact are not widespread for people with disabilities who receive SSI or DI benefits. In 1993, 300,000 of the 7 million SSI and DI beneficiaries (or 4.3 percent of beneficiaries) were referred for vocational rehabilitation and 6,000 (or 2.0 percent of referrals) were rehabilitated (U.S. General Accounting Office, 1995). In an assessment of the reasons underpinning the return to work of DI beneficiaries, vocational rehabilitation services and the SSA programs were not found to be significant factors (U.S. House of Representatives, Committee on Ways and Means, Subcommittee on Social Security, 1998). The U.S. General Accounting Office (1997) has found that states' disability determination service offices refer an average of only 8 percent of DI and SSI beneficiaries for vocational rehabilitation services. In 1997, a total of 58,358 people with mental retardation exited states' vocational rehabilitation systems, and 27,836 or 48 percent of them did so with employment outcomes (Teimouri, personal correspondence, 2000). Employment outcomes include enclave or mobile crew positions and various forms of supported work and do not necessarily entail competitive employment. Thus, the number of people with mental retardation who receive vocational rehabilitation services is very small in relation to the total population of such people.

In addition, state mental retardation agencies vary considerably in their ability to provide supportive employment, and the extent to which responsibility for supported work is vested in these agencies rather than the vocational rehabilitation sector differs among states. As of 1996, only 23 percent of the lead state agencies for mental retardation services funded integrated employment (Mank, 2001).

Community rehabilitation agencies, the most common providers of vocational supports for people with mental retardation, vary significantly in their ability to provide employment support. Many continue to concentrate their resources on nonintegrated and center-based programs, such as sheltered work, few of which pay their clients gainful wages (Mank, 2001).

Fifth, SSI and DI are critical income maintenance programs for people with mental retardation and often serve as an important gateway program for other services. Given their high unemployment rates and their overall marginalized status in society (Edgerton, 1993), people with mental retardation are highly dependent on SSI and DI to assist them in meeting basic needs. Furthermore, SSI beneficiary status usually confers Medicaid eligibility and in some states access to housing, subsidies, and other support programs. As such, SSI is a major component of the supports that can be utilized by people with mental retardation and that enables them to live in their home communities. In addition, for a relatively small number of people with mental retardation, it provides job supports that enable people to work by participating in work incentive programs.

Sixth, people with mental retardation want to work but are cautious about the loss of benefits that can accompany unemployment. Most available evidence clearly shows that people with mental retardation want to work and given the appropriate supports will work (Mank, 2001). At the same time, it is equally clear that supported work, which is one programmatic gateway to gainful employment, typically does not provide sufficient income to pay for minimal living expenses. This is particularly so in urban areas, where the cost of living, and housing in particular, may greatly exceed the resources provided by minimum or low-income wages. To the extent that supported work services are available and can provide a stepping-stone to employment above minimum wage, with adequate and typical fringe benefits such as health insurance, these programs can defray the need for receipt of SSI benefits. However, the impact of supported work in garnering these types of jobs for people with mental retardation has not been fully ascertained. Even when such jobs are obtained, new hires with mental retardation, lacking seniority, will remain vulnerable to layoffs associated with fluctuations in the national and local economies and changes in the labor market.

Advocates have argued that the barriers confronting people with mental retardation and other developmental disabilities in obtaining meaningful work are considerable and require particularly careful policy approaches. As individuals who often live and work at society's margins and whose earned income is frequently not substantial enough to raise them out of poverty, people with mental retardation and their allies must be cautious about the loss of benefits that can accompany employment. The Illinois Planning Council on Development Disabilities (IPCDD) stated:

> The realistic fear of loss of eligibility for Supplemental Security Income (SSI) and Social Security Disability Insurance (DI) is a major barrier which prevents many individuals with developmental disabilities from becoming employed. The policy of the IPCDD is that people with developmental disabilities must have the opportunity to be productively employed and earn real wages without jeopardizing their eligibility for programs and services. Specifically, for those wage earners with developmental disabilities who require support services,
>
> - the monetary value of such support services should not be counted as part of the individual's income when determining eligibility for services and benefits, and
> - the maximum allowable income for services and benefit eligibility should be sufficiently high so as to reflect the increased cost of living borne by people who require support services and/or adaptive equipment and supplies (Illinois Planning Council on Developmental Disabilities, 1991, p. 19).

Return to Work

The likelihood that people with mental retardation can support themselves by working has also been enhanced by the efforts of activists to secure antidiscrimination legislation in employment settings, which culminated in the 1990 passage of the Americans with Disabilities Act (Ticket to Work and Work Incentives Improvement Act of 1999). Public support for greater employment of people with disabilities has included a focus on the growth in the DI program, in which beneficiaries increased by 53% (Social Security Administration, 2001a) and the SSI program, in which beneficiaries more than doubled during the same period (Social Security Administration, 2001a). Advocates with disabilities and fiscal conservatives alike have shared an interest in increasing opportunities for people with disabilities to obtain employment. SSA (Social Security Administration, 1999) reports that less than half of one percent (0.5 percent) of DI beneficiaries return to work, and about 1 percent of SSI beneficiaries return to work, in spite of many indications that beneficiaries want to work. SSA further states that returning recipients to the workforce is one of its chief priorities, while acknowledging little success in this regard (Social Security Administration, 2000, 2002).

Employment Barriers for People with Disabilities

Barriers to work for people with disabilities include difficulty determining who can work, SSI and DI program components that impede return to work, and relatively low priority given by SSA to initiatives targeting employment acquisition (U.S. General Accounting Office, 1996). The General Accounting Office has suggested that the nature of the disability determination process itself, which necessitates that claimants prove that they cannot work, is incompatible with, or impedes development of, return-to-work goals and implementation strategies to achieve these goals (U.S. General Accounting Office, 1996, 1997). It is widely held that the current structure of both the SSI and DI programs impedes return to work, because their design and implementation actually obstruct the identification and enhancement of the productive capacities of beneficiaries who may profit from employment assistance (Petersen & Bascetta, 1999).

In addition, economists argue that the structure of the DI program, in particular, includes extensive and effective work disincentives (Hoynes & Moffitt, 1999). Basing disability determination decisions on the presence or absence of an impairment contained in the SSI *Listing of Impairments* (Social Security Administration, 2002) is problematic, in that it categorically presumes that certain disabling conditions automatically preclude the ability to work, or at least substantively impair work performance in the absence of special vocational preparation or work supports. Researchers agree that medical diagnoses are inadequate predictors of the ability to work (Fox, 1994; Mor, 1998; Petersen & Bascetta, 1999). Incentives are therefore built into the system for claimants to overstate their disabilities and understate their work abilities, because the stakes are all or nothing (Petersen & Bascetta, 1999).

In the United States, most people receive their health insurance coverage through their employment. This relationship is important to understanding a key barrier to work for people with disabilities. The National Academy of Social Insurance has identified three critical issues in the linkage between work and health care that impede employment for people with disabilities. First, employers may be reluctant to hire people with disabilities due to fears of increasing health premiums. Second, work is not economically feasible for people with disabilities who cannot obtain health insurance on the open market. Third, gaps in private health care coverage, including both managed care and fee-for-service plans, can be significant for people with disabilities (Mashaw & Reno, 1996). Between 1988 and 1992, the number of people without health insurance increased by 5 million. During this same period, employer-sponsored health insurance increasingly limited coverage for chronic conditions (Daniels & West, 1998). Not long ago, a nationally representative panel of over 1,200 disability leaders identified concern about the loss of Medicare or Medicaid as the greatest barrier to employment for people with disabilities (President's Committee on Employment of People with Disabilities, 1994). As a result of the powerful link between health care and employment, the National Academy of Social Insurance has recommended expansion of Medicare and Medicaid to allow poor working people with disabilities to buy into the programs, thus eliminating loss of health care as a disincentive to work. The National Academy of Social Insurance's assessment of the relationship between DI and SSI benefits, employment, and health care is telling:

> The [Disability] panel's basic finding is that the Social Security and SSI disability benefit programs do not pose strong incentives for Americans with disabilities to seek benefits in lieu of working. Rather, the strict and frugal design of these programs makes remaining at work preferable to benefits for those who are able to work. . . . While neither DI nor SSI cash benefits pose strong incentives for Americans to seek these benefits in lieu of working, constraints on access to health care and related services can (Mashaw & Reno, 1996, pp. 24).

Ticket to Work Legislation

The most recent legislative attempt to address employment issues for people with disabilities was passage of the Ticket to Work and Work Incentives Improvement Act (1999), which became effective in January 2001. The Ticket to Work Act contains two major sets of provisions that address barriers to employment. Under Title I of the act, people with disabilities can receive a ticket that enables them to obtain employment services such as vocational training or job placement services from vendors that may be state agencies, nonprofit organizations, or private providers. This portion of the act also contains provisions that reward providers for better employment outcomes by paying enhanced rates.

The second title of the act is intended to remove barriers that force people with disabilities to choose between work and health care. States are given the option of providing Medicaid coverage to people with disabilities ages 16-64 who are working. States can opt to allow working individuals with disabilities whose incomes are at or above 250 percent of the poverty level to buy Medicaid coverage. This title of the act also creates a new Medicaid buy-in demonstration project funded at $250 million for fiscal years 2001-2006 and offers $150 million in federal funding for improvements to the Medicaid infrastructure. Importantly, Section 112 of the law also provides for expedited reinstatement of SSI, DI, Medicaid, and Medicare benefits for many people if their benefits ended because they became employed but are again unable to work within 5 years and they still have the same condition.

The SSA supported passage of the Ticket to Work Act and estimated that 40,000 people with disabilities would leave the disability rolls as a result of the law during the first 10 years (Donkar, 2000). The act has the potential to significantly reduce SSI and DI expenditures, because it will also reduce an identified barrier to workforce participation (U.S. House of Representatives, Committee on Ways and Means, Subcommittee on Social Security, 1998). In addition, SSA increased the level at which an individual's income is considered to be "substantial gainful activity" from $500 monthly in 1998 to $740 monthly in 2001 with annual increases scheduled (Social Security Administration, 1999, 2000).

Advocates with disabilities have also hailed the new legislation as the first major attempt by Congress to address some of the serious barriers to employment that confront people with disabilities. Advocates have also expressed concerns that since many of the law's new provisions are optional, states will need incentives or have to be pressured to adopt the new programs. Provisions in the Ticket to Work Act that expand eligibility for Medicaid to workers with disabilities are likely to discourage growth in DI and SSI because they address the health care-employment linkage that encourages people with disabilities to maintain their SSI or DI status in order to retain health insurance benefits (Mashaw & Reno, 1996).

Given the high costs of supporting people with disabilities through the DI and SSI programs, there has been increasing interest in strategies and programs that return people with disabilities to work, and the economic trends of the 1990s seemed to support these initiatives. A strong economy and a tight labor market create incentives for employers to work to retain their employees (Travarozzi, 2000) and to hire from traditionally underrepresented groups, including people with disabilities. As mentioned earlier, in spite of the economic boom of the 1990s, however, people with disabilities have not realized major advances in their employment rates (Burkhauser et al., 2001); the reasons for this are not well understood.

Research has discounted the notion of returning to work as a static event ending a period of work-related disability (Baldwin & Johnson, 1998). However, return-to-work initiatives in the SSI, DI, and vocational rehabilitation programs are not equipped to address the changing nature and needs of people with disabilities. For many people with mental retardation, however, the Ticket to Work Act may offer opportunities to maintain their health care coverage and enter the workforce for the first time.

Although the Americans with Disabilities Act prohibits discrimination in the workplace, the burden of proving discrimination is on the wronged employee. Seeking remedies requires substantial emotional, cognitive, and often financial resources, which are likely to be particularly difficult for people with mental retardation to muster. These considerations may make the promise of the Ticket to Work Act less fruitful for people with mental retardation than for those with other disabilities.

CONCLUSIONS AND RECOMMENDATIONS

The committee's review of the scientific and policy literature leads us to make the following conclusions:

First, gainful employability of people with mental retardation is heavily dependent on variables that are external to the individual. Furthermore, the conditions that allow people with mental retardation to be employed often go unmet. Unlike certain other disabilities, for which the provision of short-term interventions such as vocational rehabilitation can result in long-term gains in employment skills, from an employment perspective mental retardation is a lifelong disability. Although employment is often possible and desirable, the supports, services, and networks needed by this population in order to be continuously employed are significantly underdeveloped. The lack of SSI benefits is likely to leave vulnerable individuals without critical supports.

Second, given their dependence on external supports and their vulnerability to changes in the work environment, it should be made easier for people with mental retardation to quickly resume receipt of benefits throughout their lives when they become unemployed, even beyond the provisions of the Ticket to Work Act. The knowledge that eligibility can be reestablished easily may encourage those already receiving benefits to seek employment. There is evidence that some, who might otherwise work, continue to receive benefits instead, because they fear for their future security should they lose their jobs.

Third, the earned income of gainfully employed persons with mental retardation is extremely low, and this population is highly vulnerable to unemployment in economically depressed periods. SSA has developed a variety of work incentive programs to aid recipients in their return to work. Although these services hold great promise and provide supports that are consistent with the vocational assistance needs and aspirations of people with mental retardation, they are not frequently utilized. Low utilization rates may reflect fears about a loss of critical income support and health care coverage. Counterbalanced against this realistic fear is the desire of many people with mild mental retardation to be gainfully employed, tax-paying members of society. A critical step in resolving this tension is the provision of effective support services that assist individuals to maintain employment and the provision of flexibly administered benefits that provide ongoing and predictable income and health care coverage. Given individual needs for supports and services, work incentive programs pose great potential to improve the gainful employability of people with mental retardation. However, these programs should not be time-limited because of the enduring nature of mental retardation and its impact on the individual's ability to work.

Fourth, benefits provided by Medicaid programs are more generous than those typically available in privately purchased health care plans. And people with mental retardation who are gainfully employed may work in settings that do not provide employer-based health insurance or they may require employee contributions that are prohibitively costly for low-wage workers. Fear of losing health care coverage dampens the willingness of many people with mental retardation to seek gainful employment. Extending Medicaid coverage to them regardless of their eligibility for SSI benefits would provide greater incentive for these individuals to seek employment.

Fifth, given the importance of SSI and DI benefits in the lives of adults with mental retardation, efforts should be made by school personnel to refer individuals to secure eligibility determination prior to adulthood. Many school systems do not use the label "mental retardation" because of its social stigma. Thus, reliance on diagnostic labels associated with the provision of special education will be insufficient to identify potentially eligible children. Although many with mental retardation qualify for benefits as children, others whose families do not meet income guidelines do not. Eligibility data show that many people with mental retardation do not establish SSI eligibility until well into their twenties or even later. This kind of planning should be part of transition planning for all youth with mental retardation. The Social Security Administration and the Department of Education could sponsor joint initiatives at the federal and state levels to promote timely determinations and information dissemination to teachers and school administrators.

Based on our review of the literature and the conclusions listed above, the committee makes the following recommendations concerning the context of the SSA disability benefit programs.

Recommendation: The Social Security Administration should remove disincentives for people with mental retardation to seek employment by:

- **Considering individuals with mental retardation to be presumptively reeligible for benefits throughout their lives, if they have previously received benefits, subsequently secured gainful employment, and then lost that employment.**
- **Encouraging the use of work incentive programs for people with mental retardation, with appropriate and necessary protections of each program's role as a safety net for income support.**
- **Permitting individuals with mental retardation to retain eligibility for Medicaid independent of their employment status.**

Chapter 3

The Role of Intellectual Assessment

For many years, only scores from intelligence tests (IQs) were used in the diagnosis of mental retardation. As professionals and the public came to understand better the limitations of intelligence theory and IQ tests, finding other useful measures for assessing mental retardation became more urgent, especially because of allegations of racial, cultural, and gender bias in standard IQ assessment instruments. Yet constructs like adaptive behavior have proven at least as difficult to assess as intelligence, and IQ still looms large in determining eligibility for a diagnosis of mental retardation. To address the many misunderstandings about intelligence and its assessment, this chapter covers the following topics: (1) intelligence theory and test use from a historical perspective; (2) intelligence tests used commonly in the diagnosis of mental retardation; (3) assessment conditions that affect examinees' assessed cognitive performance; (4) the use of total test scores, like full-scale IQs, and subscores (part or scale scores) in the diagnosis of mental retardation; (5) the use of comprehensive as opposed to restricted measures of intelligence; and (6) psychometric considerations in the selection and application of intelligence tests for diagnosing mental retardation, including test fairness.

HISTORICAL PERSPECTIVE ON THEORY AND PRACTICE

History of Development of Tests of Intelligence

The use of intelligence tests in the process of diagnosing mental retardation dates back to the turn of the 20th century, when Alfred Binet and Theodore Simon developed an intelligence test for that purpose. In the course of the implementation of universal education laws in France at that time, debates arose over the relative benefits and methods of educating schoolchildren with subnormal intelligence. As a result of this educational movement, Binet and Simon developed and in 1905 published what has come to be known as the first "practical" intelligence test (Sattler, 1988).

Binet Scale

Three years after its initial publication in 1905, the Binet-Simon Scale was revised by Binet and Simon (Binet & Simon, 1916) and then again by Binet in 1911. The instrument was noticed by researchers in the United States and was brought to this country by Goddard (1908). Three independent researchers, Huey, Kuhlmann, and Wallin, translated the Binet Scale into English in 1911 (Thorndike & Lohman, 1990), and use of the instrument and the general practice of assessing intelligence for many purposes spread quickly.

Lewis M. Terman was responsible for making the Binet Scale a recognized and accepted professional tool. Terman adopted, then revised, and renormed the instrument several times at Stanford University (Terman, 1916; Terman & Merrill, 1937, 1960, 1973), and from the early 1900s it became the principal tool for assessing the intelligence of children, adolescents, and young adults. As a result of Terman's research and development efforts, the original Binet-Simon scale eventually become known throughout the United States as the Stanford-Binet Intelligence Scale, which is currently in its fourth edition (Thorndike et al., 1986).

Pioneer Nonverbal Assessments

In tandem with Binet and Simon's work, European clinicians also attempted to develop methods for assessing the cognitive functioning of children who could not or would not speak. This effort, designed to determine latent cognitive functioning in the absence of manifest language abilities, initiated the field of nonverbal intellectual assessment. In the widely celebrated case of Victor, the Wild Boy of Aveyron, Jean Itard sought to determine the cognitive abilities of a feral youth and help the boy acquire functional language skills (Carrey, 1995; Itard, 1932).

In addition to Itard's pioneering work, even earlier historical figures pursued the problem of assessing the intellectual abilities of children who could not or would not speak more directly. Seguin (1856) is possibly best known for his development of unique instrumentation to aid in the assessment of children's abilities through nonverbal means. Seguin's performance-based nonverbal measure of cognition required the puzzle-like placement of common geometric shapes into openings of the same shape. The instrument and its many derivatives have become widely used and are known universally as the Seguin Form Board (DuBois, 1970). The current edition of the Stanford-Binet Intelligence Scale includes a Seguin-like form board, which has resulted in the merger of efforts by Binet, Simon, and Seguin, the three pioneer European test developers, in a contemporary American instrument.

Nonverbal intelligence testing has a history paralleling that of traditional language-loaded intelligence tests with the publication of many nonverbal scales during the early 1900s. In the lineage of nonverbal intelligence tests, the Leiter International Performance Scale (Leiter,

1948) is one of the only surviving historical instruments, although the number of new nonverbal tests available has grown since 1990.

Group Language and Nonverbal Assessments

The parallel development of verbal and nonverbal intellectual assessment continued during the group mental testing movement that stemmed from the country's need to assess military recruits during the First World War. According to the Examiner's Guide for the Army Psychological Examination (U.S. Government Printing Office, 1918), military testing was deemed necessary to classify soldiers according to mental ability, create organizational units of equal strength, identify potential problem soldiers (e.g., recruits with cognitive disability), assist in training and assignments, identify potential officers, and discover soldiers with special talents or skills. The Army Mental Tests resulted in Group Examination Alpha and Beta forms. The Group Examination Alpha (Army Alpha) was administered to recruits who could read and respond to the written English version of the scale. Because the Army Alpha was not useful as a measure of ability when recruits had limited English proficiency or were insufficiently literate to read and respond reliably to verbal items, the Group Examination Beta portion of the Mental Tests (Army Beta) was developed as a nonverbal supplement to the Army Alpha.

Wechsler Scales

Since the onset of mental testing with the Stanford-Binet and the application of group intelligence testing procedures, a plethora of individual and group tests have been developed in the United States for assessing overall intelligence and diagnosing subnormal intellectual functioning in infants, children, adolescents, and adults. Most prominent among the post-Binet instruments was a series of intelligence tests developed by David Wechsler (1939, 1949, 1955, 1967, 1974, 1981, 1989, 1991). Although the Binet scale was preeminent during the early to mid-1900s, the Wechsler scales quickly replaced the Binet as the test of choice among psychological examiners. The Wechsler and Binet scales remain the two dominant, language-loaded, individually administered intelligence tests used for the diagnosis of mental retardation in the United States.

The Wechsler Scales of Intelligence employed the Army Alpha and Beta approach to assessment by creating a collection of language-oriented subtests (verbal scale) and a collection of language-reduced subtests (performance scale), which combine to create a full scale IQ (FSIQ). Historically, the performance scale has been used as a nonverbal test because of its reduced language demands, but it is not truly nonverbal in that it requires the examinee to comprehend lengthy and complex verbal instructions.

During the past 20 years, a number of intelligence tests have been published as alternatives to the Stanford-Binet and the Wechsler scales. Currently psychologists have an impressive array of instruments differing in their features, theoretical orientations, length and complexity, and technical quality from which to select.

Correlates of Assessed Intelligence

The widespread use of intelligence tests in the diagnosis of mental retardation is a consequence of research outcomes that have definitively demonstrated that, of all the social science variables that have been studied, intelligence tests remain the single best predictors of most important life events and outcomes (Jensen, 1981; Neisser et al., 1996; Sattler, 1990; Wilson, 1978). Intelligence tests predict such diverse outcomes as academic achievement, attainment, and deportment (Beck et al., 1988; Martel et al., 1987; Paal et al., 1988; Poteat et al., 1988; Roberts & Baird, 1972; Venter et al., 1992); language development, comprehension, and communication (Ackerman-Ross & Khanna, 1989; Bolla et al., 1990; Bracken, Howell, & Crain, 1993; Bracken, Prasse, & McCallum, 1984; Caplan et al., 1992; Lindsay et al., 1988; Morton & Green, 1991; Mitchell & Lambourne, 1979); psychosocial adjustment (Cunningham et al., 1991; Denno, 1986; Drotar & Sturm, 1988; Greenwald et al., 1989; Kohlberg, 1969; O'Toole &

Stankov, 1992; Poon et al., 1992; Siebert, 1962; Stogdill, 1948; Windle & Blane, 1989); family and home environment (Bracken et al., 1993; Luster & Dubow, 1992); short-term memory (Miller & Vernon, 1992); and employment success (Arvey, 1986; Burke et al., 1989; Faas & D'Alonzo, 1990; Hunter, 1986; Hunter & Hunter, 1984; Schmidt & Hunter, 1981; Thorndike, 1986).

With such a diverse and wide array of correlates, intelligence tests have become highly instrumental in the identification of levels of cognitive functioning, including differentiating levels of mental retardation and the prediction of concomitant behavioral, social, and economic consequences.

Theories of the Structure of Intellectual Abilities

Research and theorizing on the structure of intellectual abilities has a history that is virtually as long as the history of work on the measurement of intelligence. As Binet and Simon were hard at work developing their seminal scale for intelligence, Charles Spearman (1904a, 1904b) published two groundbreaking statistical papers, one on basic methods of correlational analysis and the other that laid the foundation for factor analysis. The factor analytic techniques that Spearman (1904a) proposed were specially geared for testing his theoretical notions regarding ability structure, but the value of the generalized factor analysis model was recognized almost immediately by other researchers. Factor analysis has become the standard way to investigate the structure of the ability and other domains for over half a century.

An interesting conundrum in ability research is the continuing disconnection between techniques for assessing intelligence, or general intellectual ability for practical decision making, and research on the structure of intellectual abilities. When assessing intelligence to make decisions about individuals, attention has been paid almost exclusively to general intelligence, as reflected in a composite intelligence quotient, or IQ. That is, a single number, embodied in the IQ, is used to portray an individual's mental ability. This focus on a single dimension of general intelligence is consistent with the theory outlined by Spearman (1904a). The need to consider more than a single factor to represent correlations among ability tests was recognized only a few years after Spearman first described his theory. Furthermore, the need for more than a single factor has been widely acknowledged for over 60 years; the key disagreements in the field relate to how the structure of multiple abilities is portrayed, understood, and used in decision making. This continuing disconnection between theory and practice in the structure and measurement of human cognitive or intellective abilities is a central issue in psychometrics today (McArdle & Woodcock, 1998).

Signs of a closer connection between theory and practice in the measurement of abilities are apparent, and the next decade is likely to show even greater influence of ability theory on the range of mental abilities for which IQs can be obtained. We now review the major theories of the structure of intellectual abilities, which point toward an emerging consensus on the major ability dimensions that constitute the ability domain. This information undergirds the committee's recommendations regarding the intelligence test scores that can best be used for eligibility decisions.

Spearman's Two-Factor Theory

Spearman (1904a) developed factor analytic techniques to test his hypothesis that a single dimension accounted for correlations among all tests of mental ability. Spearman called this dimension "general intelligence." To avoid contaminating the scientific construct of general intelligence with any ideas associated with the notion of intelligence in common parlance, Spearman signified the scientific construct derived from correlations among ability tests with the letter *g*, which stood for general intelligence. Spearman argued that *g* represented a new scientific construct, the meaning of which would be established only with substantial empirical research.

Spearman was perhaps the first to notice what is called the *positive manifold*, which refers to the finding of uniformly positive correlations among tests of ability. This positive manifold is a hallmark of the ability domain and is a distinctive attribute of the domain in comparison with others. Spearman reasoned that, if all tests of ability are positively intercorrelated, a single entity might influence all tests and thus be in common among the tests. Tests that correlate highly with other tests would be more heavily saturated with this common entity, whereas tests that tended to correlate at lower levels with other tests would be less saturated with the common entity. Spearman (1904a) presented techniques for estimating the saturation of each test, based on its correlations with other tests, and he continued to refine and extend these techniques for the remainder of his career.

Spearman's theory is frequently called the two-factor theory, reflecting the hypothesis that two factors account mathematically for the variance of each measured variable. One of these factors is g, the factor of general intelligence; and the second factor is s_j, a factor that is specific to manifest variable j. Thus, the two-factor theory postulates two classes of factors. One class has a single member, g, the factor of general intelligence, which is the single influence that is common to all tests of ability. The second class of factors has as many members as there are tests of ability, one specific factor for each different test of ability.

As a theoretical metaphor, Spearman (1927) borrowed from the Industrial Revolution. Arguing that g, or general intelligence, could be likened to or identified with mental energy, Spearman also hypothesized that individual differences in mental energy were largely genetic in origin. This mental energy can be directed toward any kind of intellectual task or problem, and the greater the amount of mental energy devoted to a task, the better the performance on the task. Individuals with a high level of g have a high level of mental energy to devote to intellectual pursuits, whereas persons with low levels of g have much lower levels of mental energy at their disposal when confronting intellectual problems or puzzles. Consequently, individual differences in g

reflect individual differences in mental energy, and individual differences in mental energy lead to individual differences in performance on all ability tests and therefore account for the correlations among all tests of mental ability.

The specific factor for variable j, s_j, is composed theoretically of two components—a reliable component that is specific to variable j, and a stochastic or random component that represents random error of measurement. (This specific factor is sometimes referred to as residual variance.) In most research situations, these two components cannot be separated, so emphasis is laid on the combined specific factors. Spearman equated the specific factor s_j for a given test j with an engine. General intelligence, or g, provides the mental energy to power the engine that is used to solve a particular type of problem. Thus, one engine would be used to solve the problems on a verbal comprehension test, another engine would be used for numerical problems, and so forth. For certain types of problems, the general factor g is of primary importance, leading to a high g-loading for such a test and a relatively low contribution to explained variance by the engine, or specific factor, for the test. But, for other tests, g is of less importance, and the engine for the test accounts for the majority of the variance. The specific factor s_j for a test is an opportunity for the environment or experience to play a part in performance on mental ability tests.

When conducting research to test his hypothesized ability structure, Spearman often conducted analyses so that the results would conform to his theory. For example, Spearman (1914) dropped a test from his analyses because its inclusion resulted in a failure to satisfy his statistical criterion for adequacy of a single factor. Once the test was dropped from the analysis, the remaining tests satisfied the mathematical criterion, supporting the adequacy of a single factor for the set of tests. This approach—discarding tests that led to failure to confirm his theory—was a common one for Spearman, who discarded recalcitrant tests in several analyses reported in his major empirical work on mental abilities (Spearman, 1927). As a result, Spearman's two-factor theory has equivocal support, because any indication of lack of fit was effectively swept under the rug. But the two-factor theory is important for several reasons, including its status as the first theory of the structure of mental abilities, the clarity with which the theory and its predictions were stated, and the close interplay between psychological theory and the mathematical and statistical tools developed to test it.

Thurstone's Primary Mental Abilities

During the 1930s, L.L. Thurstone and his colleagues pursued a program of research designed to identify the basic set of dimensions that span the ability, or intelligence, domain. Rather than beginning with a strong a priori theory about the structure of mental abilities as Spearman had done, Thurstone and his collaborators took a very different approach. Specifically, they collected a large battery of tests comprising all conceivable types of intellectual tasks, administered the battery to a large sample of subjects, and then analyzed the correlations among the tests in this battery to determine the number and nature of the dimensions required to account for the correlations. If the same dimensions continued to emerge from their analyses across several samples of subjects and different but largely overlapping batteries of tests, then the dimensions would serve as a framework for representing the ability domain.

In several early studies, Thurstone and his colleagues (1938a, 1938b) found seven interpretable factors that were replicated across several analyses; these seven factors were termed *primary mental abilities*. The seven primary mental abilities that consistently appeared across samples were identified as: (1) verbal comprehension (V), or the ability to extract meaning from text; (2) word fluency (W), subsuming the ability to access elements of the lexicon based on structural characteristics (e.g., first letters, last letters), rather than meaning; (3) spatial ability (S), or the ability to rotate figural stimuli in a two-dimensional space; (4) memory (M), involving the short-term retention of material typically presented in paired-associate format; (5) numerical facility (F), reflecting the fast and accurate response to problems involving simple arithmetic; (6) perceptual speed (P), or the speedy identification of stimuli based on their stimulus features; and (7) reasoning (R), which represented inductive reasoning in some studies, deductive reasoning in other studies, and general reasoning in still others.

As for an interpretation of the nature of mental abilities, Thurstone (1938a, 1938b) was not specific. He repeatedly referred to ability dimensions as representing "functional unities," by which he meant that the tests loading on a given factor had some functional similarity that was hypothesized to be the same across tests. Thurstone did believe that the future would bring a mapping of mental abilities onto brain areas, such that each ability factor would be tied to particular brain areas that supported its functioning. But brain mapping was in its initial stages and Thurstone could only voice this as a hope for the future. He did think that cognitive psychology held hope for understanding the underpinnings of mental abilities, stating that psychologists should move into the laboratory to devise studies that would illuminate why a given set of tests loaded on a given factor (Thurstone, 1947). Once again, the field of psychology was not ready for this recommendation, and cognitive investigations into the processes underlying mental test performance began in earnest about 30 years after Thurstone's encouragement to pursue this avenue of research.

In the initial studies by Thurstone and his collaborators (e.g., 1938a, 1938b), the primary mental ability factors were rotated orthogonally, so they were statistically uncorrelated with one another. But after the development of the mathematical theory for oblique rotations (Tucker, 1940), Thurstone and Thurstone (1941) quickly applied oblique rotations to the primary mental abilities and found substantial correlations among the seven dimensions. The correlations among the primary mental abilities were well described by a single second-order factor, which Thurstone and Thurstone argued provided a way to reconcile Spearman's theory with their own. That is, at the level of the primary mental abilities, seven dimensions were required to represent the relations among a large set of tests. But correlations among the primary mental abilities could be explained by a single second-order factor. Thus, one could argue that Spearman pursued work on the ability domain at the second-order level, whereas Thurstone and his colleagues worked to specify well the dimensions that constituted the first-order level of factoring. Although this would provide a way of integrating the Spearman and Thurstone models, not all researchers agreed with this position. Indeed, Spearman (1939) argued that the primary mental abilities were rather trivial and narrow, and that the second-order general factor, or *g*, should be considered the principal or primary factor, rather than being relegated to second-order importance.

British Hierarchical Theorists

As early as 1909, Burt performed analyses that demonstrated the need to consider more than a single factor for explaining the correlations among a set of manifest indicators of ability. In this early publication, Burt (1909) provided little indication of a meaningful multiple factor structure, but 40 years later, he presented a theoretical summary of research that provided a three-level structure of mental abilities (Burt, 1949). At the first level, Burt postulated the presence of basic sensory and perceptual dimensions, including dimensions such as sound discrimination thresholds. The second level contained dimensions that were more cognitive and intellective in nature; here, typical ability dimensions such as verbal comprehension and spatial ability resided. The third level has a single dimension, the general factor of Spearman.

Vernon (1950, 1961) provided the most comprehensive and integrative review of the hierarchical theory; Vernon's focus was at its highest levels. The topmost level had a single dimension, the general intelligence factor, *g*, of Spearman. Below *g* were two subgeneral abilities: *v:ed* (or verbal:educational), and *k:m* (or spatial:mechanical). Below the *v:ed* subgeneral dimension fall factors such as verbal comprehension, verbal fluency, numerical facility, and reasoning, whereas under the *k:m* subgeneral dimension are factors such as spatial rotation, mechanical and technical information, and various psychomotor abilities. Vernon presented the hierarchical structure of abilities as a way of summarizing the previous three decades of research and considered the several versions of the hierarchy to be tentative and subject to revision in the future. However, both Vernon and Burt believed strongly in the nature of the general factor *g* as representing a single entity that was common to all tests of ability.

The third member of the British hierarchical group was Godfrey Thomson (1951), who supported the general hierarchy of abilities even as he espoused a rather different theoretical basis for it. Thomson's hierarchical factor pattern was similar to Vernon's, with a general factor aligned with Spearman's *g* at the apex of the hierarchy, followed by rather broad subgeneral factors, and finally a series of much more narrow factors at the bottom of the hierarchy.

However, Thomson believed that the ability hierarchy was based on a radically different set of processes. Indeed, he repudiated the notion of a single entity common to all tests of ability. Instead, the human mind may be composed of a virtually infinite set of bonds or potential bonds that are independent of one another. When working on a particular type of test, a given set of bonds is required to arrive at a correct answer. When a different type of test was administered, a different but overlapping set of bonds was activated. The more highly overlapping the sets of bonds required by two tests, the higher the correlation between the tests. Conversely, if the sets of bonds sampled by two tests showed little overlap, then the tests would correlate positively but at a low level. The upshot of Thomson's sampling theory was this: no single entity (i.e., bond) may be found that is common to all tests of mental ability, so the hierarchical structure of human mental abilities simply indicates the degree of overlap among the bonds sampled by tests of mental ability.

The Thomson explanation for the hierarchy of mental abilities may lead to a number of reactions. One may become highly suspect of factor analytic approaches, as one set of empirical results, with a dominant general factor, is consistent with diametrically opposed generating mechanisms—a single entity common to all tests (e.g., Spearman) versus no single entity common to all tests of ability (e.g., Thomson).

But another reaction to these findings is to become attuned to the need to marshal evidence beyond the pattern of tests loading on factors. The loadings of tests on factors may suggest the presence of an entity common to all tests of ability. But additional evidence of different types may be relevant to the choice between a single common entity and the absence of a single common entity. This additional evidence may then tip the balance in favor of one or the other of the two competing positions.

Guilford's Structure of Intellect

Based on considerable research during World War II on army recruits and a thorough review of cognitive psychological research, J.P. Guilford (1967) developed a model he termed the structure of intellect, or SOI. He and his colleagues spent over two decades attempting to confirm the basic hypotheses of SOI theory, work summarized by Guilford (1967) and Guilford and Hoepfner (1971). The Guilford theory was well recognized as a competing model of ability structure until Horn and Knapp (1973) published a reanalysis of many of the data sets used by Guilford and his colleagues to corroborate SOI theory. They found that Guilford's own data gave much stronger, in fact almost perfect, support for Thurstone's hypotheses than for hypotheses generated by SOI theory. An interesting pair of commentaries on the Horn and Knapp (1973) study by Guilford (1974) and Horn and Knapp (1974) left the main findings by Horn and Knapp (1973) unchallenged. SOI theory is no longer recognized as a useful conceptualization of the structure of human abilities.

Cattell–Horn Theory of Fluid and Crystallized Intelligence

Capitalizing on a much earlier observation (Cattell, 1941), Raymond B. Cattell (1963) proposed a new theory of ability structure, subsequently referred to as the theory of fluid and crystallized intelligence. According to the initial theory sketched by Cattell (1963), two very broad and important dimensions of intelligence—fluid intelligence, or *Gf*, and crystallized intelligence, or *Gc*—could be distinguished, rather than the single dimension of *g* hypothesized by Spearman. Cattell conceived of fluid intelligence in ways that were reminiscent of Spearman's theorizing about *g*. In particular, *Gf* was thought to be a reservoir of reasoning ability that could be directed toward many different kinds of content, hence its identification as a fluid form of intelligence. Furthermore, *Gf* was thought to be largely genetically determined.

As fluid intelligence was expended on a given kind of content or intellectual problem, the individual would develop knowledge stores related to the particular content or type of problem as well as mental algorithms for solving such problems. The knowledge and mental algorithms developed through the application of *Gf* on given tasks are therefore crystallizations of the influence of *Gf*. Thus, verbal comprehension, or the ability to extract meaning from text, is a crystallized ability assessed using tests of vocabulary, paragraph comprehension, and the understanding of proverbs, among others. All of these tests require one to extract the meaning from text using stored meanings of words in the lexicon. Numerical facility is a crystallized ability that subsumes knowledge of simple numerical facts (e.g., addition facts, subtraction facts) as well as algorithms for solving numerical problems that cannot be solved easily mentally (e.g., long division, multiple-place multiplication). The higher a person's level of *Gf*, the greater the amount of fluid intelligence invested on particular tasks, and therefore the higher that person's general levels of performance crystallized ability on tasks. Because *Gf* influences performance on all crystallized ability tasks, these tasks should correlate with one another and therefore define a general crystallized intelligence factor, or *Gc*.

Because *Gf* was a fluid ability to reason with new material, Cattell (1963, 1971) argued that *Gf* was best measured using either novel stimuli or problems or with highly overlearned stimuli with which a person is instructed to perform some novel operation, like doing simple math with letters of the alphabet. Theoretically, *Gf* was largely genetic in origin, and any learning that affected tests for *Gf* would be haphazard learning that occurred in the context of daily life. In contrast, *Gc* was best measured using tests of standard cultural knowledge (vocabulary, information, similarities) or tests of material like numerical facility that was highly practiced in standardized cultural settings such as school. One hypothesis regarding the pattern of tests loading on factors that distinguishes *Gf-Gc* hypotheses from those of the British hierarchical theorists has to do with tests of mechanical knowledge. Cattell argued that tests of mechanical knowledge should load on the *Gc* factor, which is closest to Vernon's *v:ed* factor, because mechanical knowledge is systematically taught in schools, rather than on the *k:m* factor, as Vernon had hypothesized. In addition, boys should have an advantage on mechanical knowledge over girls relative to other indicators of *Gc*, due to the more consistent teaching of mechanical knowledge to boys than to girls. These hypotheses were confirmed, lending support to structural hypotheses of *Gf-Gc* theory over those associated with the hierarchical model of Vernon.

Cattell (1971) made a further contribution to the understanding of mental abilities by distinguishing between the order and stratum of a factor. The order of a factor is a superficial, methodological aspect of the analysis in which a factor is identified, whereas the stratum a factor occupies is a deeper, theoretical concern regarding the nature and breadth of the factor. Factors that are obtained from analyzing the correlations among observed variables are termed first-order factors. If the first-order factors are rotated obliquely, factoring the matrix of correlations among first-order factors leads to the identification of second-order factors. Multiple orders of factoring may be continued as long as at least three oblique factors are identified at a given level. In contrast, the stratum a factor occupies depends on its breadth and the generality of its influence.

To make the distinction between order and stratum clearer, consider the following two research scenarios. In the first scenario, suppose that a researcher included in a battery of tests three tests of word fluency, three tests of associational fluency, and three tests of ideational fluency. Factoring these nine tests would lead to the identification of three first-order factors, one each for word fluency, associational fluency, and ideational fluency. If the correlations among these three factors were factored, a single general fluency factor (or *Glr*, for general long-term retrieval from memory) could be derived as a second-order factor, and the three first-order fluency factors would load on this second-order factor. In this research scenario, the first-order factors are also first-stratum factors, representing the narrowest dimensions that would be fruitful to research. In addition, the second-order general fluency factor is a second-stratum factor, with broader influence on each of several types of more narrow fluency.

In the second research scenario, given constraints in testing time, the second researcher could administer only a single test for word fluency, a single test for associational fluency, and a single test of ideational fluency. In this scenario, the researcher could not identify first-stratum factors for word fluency, associational fluency, and ideational fluency, because only a single manifest variable for each dimension was available, and one must have at least two, and preferably three, tests for a given factor to identify it as a factor. Factor analyzing the three fluency tests would lead to a first-order factor on which the word fluency, associational fluency, and ideational fluency tests loaded. Now this factor is a first-order factor, because it was derived from the correlations among measured variables. But, because the variables loading on it represented different types of fluency, the first-order factor reflects general fluency, or *Glr*, a second-stratum dimension.

The distinction between the order and the stratum of factors enables one to place results in a hierarchical structure based on the stratum of the factors found in different studies. The current version of *Gf-Gc* theory has been outlined in several papers by John L. Horn (1985, 1988, 1998). The ability structure for *Gf-Gc* theory posits at least 55 primary or first-stratum factors. When correlations among the first-stratum factors are analyzed, nine second-stratum factors are found. These nine second-stratum factors are: (1) *Gc* (crystallized intelligence), which has verbal comprehension, semantic relations, numerical facility, mechanical knowledge, syllogistic reasoning, verbal closure, and general information factors as indicators; (2) *Gf* (fluid intelligence), which subsumes first-order factors such as induction, general reasoning, figural relations, concept formation, and symbolic classification; (3) *Gv* (general visualization), with loadings from first-stratum factors for visualization, speed of closure, flexibility of closure, spatial orientation, figural fluency, and figural adaptive flexibility; (4) *Ga* (general auditory processing), with loadings from first-stratum factors, such as listening, verbal comprehension, temporal tracking, sound pattern discrimination, and auditory memory span; (5) *Gsm* (general short-term memory, also identified at times as SAR, for short-term acquisition and retrieval), which subsumes first-stratum factors of associative memory, span memory, meaningful memory, and memory for order; (6) *Glr* (general long-term memory, also sometimes identified as TSR, for tertiary storage and retrieval), which represents a variety of fluency dimensions, such as delayed retrieval, associational fluency, expressional fluency, ideational fluency, word fluency, and originality; (7) *Gs* (general speediness or processing speed), covering first-stratum dimensions of perceptual speed, numerical facility, and writing and printing speed; (8) *Gt* (decision speed, also identified at times as CDS, for correct decision speed), reflecting choice reaction time, decision speed, and simple reaction time; and (9) *Gq* (general quantitative knowledge), representing dimensions such as applied problems, quantitative concepts, numerical facility, and general reasoning.

The preceding results related to the loading of first-stratum factors on the nine second-stratum factors may be termed structural results. But in the continued development of *Gf-Gc* theory, Horn (1998) has always monitored several additional kinds of information. One of these additional types of information is derived from developmental studies and consists both of kinematic trends (developmental growth and decline of abilities over the life span) and of the dynamic effects of ability dimensions on one another. The differential kinematic, life-span trends for the various second-stratum abilities have been replicated many times.

These trends show that both *Gf* and *Gs* begin to decline early in adulthood, around the age of 30, whereas *Gc* continues to increase in mean level until perhaps age 70 before declines begin. This is perhaps the strongest evidence against attempting to define a higher-stratum general factor analogous to Spearman's *g*, an argument that Horn has made repeatedly. The moderate correlations among the nine second-stratum ability dimensions have been an impetus to many researchers to factor analyze these correlations and obtain a higher-order general factor. But Horn argued that, with very different life-span trends for the second-stratum dimensions, any general factor would be constructed out of the mixing of cognitive apples and oranges. This would lead to a hopelessly confounded and uninterpretable general factor showing essentially no change in level during the adult years, a pattern that none of the second-stratum factors actually displays. The dynamic effects mentioned above involve the hypothesized lead-lag relations among abilities. The most often cited of these is the hypothesis that *Gf* leads to later increases in *Gc* due to the investment of *Gf* on intellectual problems. Studies of these dynamic hypotheses have not been strongly supportive of hypothesized relations, but the current development of better models to test these hypotheses may lead to more definitive results.

Horn (1985, 1998) evaluated still other kinds of research evidence, which are discussed here only briefly. Although still somewhat premature for drawing final conclusions, neurocognitive studies appear to support the hypothesis that different ability factors are subserved by different brain areas. As these findings become more firmly established, they will provide additional support for the hypotheses of *Gf-Gc* theory. Another type of evidence is derived from studies of heritability. *Gf-Gc* theory makes certain predictions regarding heritability, or the degree of genetic variance in ability factors. One such prediction is that *Gf* should have higher heritability than *Gc*. Here, the evidence is not obviously supportive of *Gf-Gc* theory, as most estimates of heritability show about equal heritabilities for *Gf* and *Gc*. A final kind of evidence comes from studies of achievement, in which achievement in particular curricular areas is related to second-stratum dimensions of ability. Horn (1998) noted the difficulties with such studies but concluded that achievement studies tend to support differential relations between achievements in distinct curricular areas and associated second-stratum factors of ability.

In summary, *Gf-Gc* theory is a complex and far-reaching enterprise. The theory makes predictions in the structural domain concerning the loading of first-stratum abilities on the broad second-stratum factors, but also makes clear predictions in several other domains. Although empirical results to date are not fully supportive of all predictions of the theory, a sufficient number of predictions have been confirmed that *Gf-Gc* theory is the most comprehensive and widely supported ability theory currently available. The frequent replication of the differential life-span trends for different abilities has resulted in *Gf-Gc* theory being the primary theoretical framework now used in studies of adulthood and aging. Moreover, the well-replicated structural results are leading the developers of intelligence tests to incorporate measures of *Gf* and *Gc*, in addition to an overall IQ in the scoring of their instruments.

Carroll's Three–Stratum Theory

In 1993, John B. Carroll published a monumental tome that reported the reanalyses of over 450 sets of data. The aim of this project was to reanalyze all previous ability studies using a constant and well-justified set of factor analytic techniques, trusting that this would lead to a more consistent set of results across studies. The factor analytic results reported by Carroll are similar to the Horn-Cattell structural results in most respects, so little detailed description is needed here. We merely recount the broad strokes of the Carroll approach.

The upshot of the reanalysis of 477 studies was the identification of approximately 65 narrow, first-stratum factors. When correlations of these first-stratum factors were analyzed, eight second-stratum factors were located. When the correlations among second-stratum ability factors were analyzed, Carroll identified a single third-stratum factor, which he interpreted as corresponding to Spearman's *g*. One interesting advance by Carroll was to identify both level and speed components of abilities, where appropriate. The level component involves power tests in which time limits have little effect on individual differences on the tests. In contrast, the speed (or rate) component contains tests on which time limits or the rate of presentation of information, and therefore the speediness of performance, is important to measuring individual differences on the tests. One second-stratum ability had only level indicators, and two second-stratum dimensions had only speed indicators. The remaining five second-stratum factors had both level and speed (or rate) indicators.

The eight second-stratum factors identified by Carroll (1993) are: (1) *Gf* (fluid intelligence), with level first-stratum factors of general reasoning, induction, and quantitative reasoning and a speed factor of speed of reasoning; (2) *Gc* (crystallized intelligence), with level indicators of language development, verbal comprehension, spelling, and communication and speed indicators of oral fluency and writing ability; (3) *Y* (general memory and learning), with a level first-stratum factor of memory span and rate (related to speed) indicators of associative memory, free recall memory, meaningful memory, and visual memory; (4) *V* (broad visualization), with a level factor of visualization and speed indicators of spatial relations/orientation, speed of closure, flexibility of closure, and perceptual speed; (5) *U* (broad auditory perception), with level indicators of hearing and speed thresholds, speech sound discrimination, and musical discrimination and no clear speed or rate indicators; (6) *R* (broad retrieval), with level indicators of originality and creativity and speed indicators of ideational fluency, associational fluency, expressional fluency, word fluency, and figural fluency; (7) *S* (broad cognitive speediness), with no level indicators but speed indicators of rate of test taking and numerical facility; and (8) *T* (processing speed and/or decision speed), once again with no level indicators but speed indicators such as simple reaction time, choice reaction time, semantic processing speed, and mental comparison speed.

Despite the clear similarities between the Horn-Cattell and Carroll models, some important differences are apparent. The key difference concerns the presence and nature of a general intelligence factor.

Carroll (1997) argued that his work provided perhaps the strongest and most comprehensive support yet for the general intelligence factor, a position Cattell would probably have seconded. Carroll also identified the general factor as corresponding to Spearman's *g*, the mental ability common to all tests of ability, also a position that Cattell might have favored. However, for more than 25 years, Horn has been responsible for the current synthesis of the Horn-Cattell model. He has long disclaimed the utility of a general factor, despite the positive correlations among the second-stratum abilities. Based on other information, such as the trends of growth and decline over the life span for second-stratum abilities, any overall score approximating general intelligence would represent a changing mixture of abilities, a general level of a person's profile of second-stratum abilities, or "intelligence in general" or "on average," rather than a single element common to all tests that retains its unitary nature across development. This striking difference of scientific opinion is reminiscent of the conflicting views on the nature of the general factor held by Vernon and Thomson, discussed above. The monumental work by Carroll (1993) was concerned almost exclusively with structural information about how variables load on factors. Carroll dismissed other forms of data, particularly differential life-span aging trends, by claiming that the data and their implications for theory were not yet sufficiently well established. In contrast, Horn has always studied structural information, but he has also monitored and integrated information from numerous other sources, such as kinematic or life-span trends, dynamic relations between abilities over time, and neurocognitive studies. Taking all of these kinds of information into consideration, Horn has argued that the existence of a single, unchanging entity common to all tests of ability cannot be supported.

The Horn-Cattell and Carroll models exhibit additional, but less important, differences. One of these is the absence of *Gq*, or general quantitative ability, as a second-stratum dimension in the Carroll model. Carroll considered the *Gq* dimension of the Horn-Cattell theory to be too narrow and lacking a sufficient research base to be accorded a position as a second-stratum factor. Also, some differences in the first-stratum factors subsumed by second-stratum dimensions can be found.

Aside from the preceding differences, eight of the second-stratum ability dimensions from the Horn-Cattell and Carroll models fall in a rather clear one-to-one relation with one another. Some second-stratum dimensions have differing names and identifying symbols across the two systems. Still, the eight second-stratum dimensions represent the current state of the science with regard to the broad abilities that span the intelligence domain. An overall score, whether corresponding to Spearman's *g* or to a changing composite reflecting "intelligence in general," may be a useful summary index of a person's general level of functioning, regardless of whether one believes the score corresponds to a particular identifiable entity.

Other Theories

The preceding theories were developed in connection with the use of factor analysis, which was used to derive the dimensions underlying batteries of tests and thereby confirm or disconfirm the hypotheses put forward by the groups of researchers. In addition to these theories based on factor analysis, several additional theories of the structure of mental abilities have been developed. Most of these other theories have been based on a priori theory or summaries of previous research, but have relied much less or not at all on sophisticated measurement techniques such as factor analysis. As a result, the utility of these theories for applied work on the assessment of intelligence is much more limited, although the future may see greater application of the ideas.

The first of these other theories is embodied in the PASS model of Das, Naglieri, and Kirby (1994). PASS stands for planning, attention, simultaneous processing, and successive processing, which are processes or mental functions associated with particular brain areas by Luria (1966a, 1966b). Planning refers to processes governing cognitive control and self-regulation, enabling a person to develop or plan courses of intelligent action to be followed. Attention subsumes the processes by which a continual focus on cognitive problems is maintained. Simultaneous processing involves processing of stimuli in which the stimulus as a whole must be comprehended or in which elements must be integrated into a meaningful whole. Successive processing concerns processes in which the sequence of the processing of elements is crucial, such as language. Factor analytic studies of the PASS model have been less than fully successful, failing to establish planning and attention as empirically distinct entities. Despite this, the Cognitive Assessment System (Naglieri & Das, 1997) provides a standardized battery to assess the components of the PASS model.

A second theoretical approach encompasses information processing theories derived from cognitive psychology. For example, Campione and Brown (1978) offered an initial model that was further developed by Borkowski (1985). Information processing models of cognitive ability often distinguish the architectural and executive systems, roughly equivalent to the hardware and software components, respectively, of a computer. The architectural system is assumed to be genetically, or at least biologically, based and consists of basic operating parameters of cognitive processes, encompassing individual differences in (1) amount of information that can be processed, which is assessed using memory span, (2) durability of information storage, or the retention of memory traces, and (3) efficiency of processing, or the speed of encoding and decoding information. The executive system encompasses components that are environmentally based and guide processes comprising problem solving. The executive system subsumes components such as (1) one's knowledge base, or declarative knowledge of facts; (2) control processes, which include strategies or heuristics to aid memory or problem solution; and (3) metacognition, which involves, among other things, knowing how problems should be solved and then monitoring progress toward problem solution and evaluating outcomes to ensure successful solution of the problem. Researchers using the information processing approach have paid little attention to converting theoretical insights into usable measures of intelligence.

Sternberg (1985, 1986, 1996) has offered several theories of human intelligence, theories that have been radically reformulated over time. The most recent incarnation is Sternberg's notion of successful intelligence. The three components of successful intelligence are (1) analytic abilities, which aid in defining problems, setting up solution strategies, and monitoring solutions and presumably include many of the dimensions outlined in the Horn-Cattell and Carroll models; (2) creative abilities, which involve generating new problem solving options and attempting to convince others of their worth; and (3) practical abilities, which subsume skills in ensuring that one can implement solutions and see that they are carried out. As with information processing approaches, at present no standardized batteries are available to assess constructs within Sternberg's triarchic theories.

The final theory discussed in this section is the theory of multiple intelligences, described by Gardner (1983). According to this theory, at least eight different types of intelligence can be identified: (1) linguistic intelligence, subsuming language and communication skills; (2) musical intelligence, involving individual differences in rhythm and pitch and skills in composing music; (3) logical-mathematical intelligence, including logical reasoning and number abilities; (4) spatial intelligence, or the ability to understand spatial relations; (5) bodily-kinesthetic intelligence, assessed by skills in dancing, acting, and athletics; (6) intrapersonal intelligence, or knowledge of one's self, feelings, and motives; (7) interpersonal intelligence, or skills in discerning the feelings, beliefs, and intentions of others; and (8) naturalist intelligence, involving seeing and understanding patterns in nature. Gardner has done little research to validate his theory on the types of intelligence. To the extent that evidence supports the notion of different intelligences, the evidence is consistent with the Horn-Cattell and Carroll theories. For example, Gardner's linguistic intelligence is most similar to *Gc* in the Horn-Cattell model. As a result, little empirical evidence is available that uniquely supports Gardner's theory. Moreover, no standardized measures of the constructs in this theory are available.

Summary

During the 20th century, theories of the structure of mental abilities have evolved from the two-factor theory of Spearman, which hypothesized only a single factor common to all tests of ability, to the more differentiated structure of the Horn-Cattell and Carroll models. In these models, the two most widely studied of the second-stratum factors are *Gc* and *Gf*. *Gc*, or crystallized intelligence, reflects stored cultural knowledge and corresponds closely with the verbal factor often reported in factor analyses of the Wechsler batteries. *Gf*, or fluid intelligence, is a dimension representing reasoning or thinking skills; the performance factor identified in factor analyses of the Wechsler batteries appears to be an amalgamation of *Gf* and *Gv* (or visualization skills).

Some movement has already taken place in structuring intelligence tests to acknowledge the utility of the Horn-Cattell and Carroll models. For example, the Stanford-Binet IV yields a composite IQ, but it was based on a theoretical model that included subareas for crystallized abilities (verbal reasoning and quantitative reasoning), fluid-analytic abilities (abstract/visual reasoning), and short-term memory. Furthermore, one battery—the Woodcock-Johnson—was explicitly designed to assess all second-stratum dimensions in the Horn-Cattell model. During the next decade, even greater alignment of intelligence tests and the IQ scores derived from them and the Horn-Cattell and Carroll models is likely. As a result, the future will almost certainly see greater reliance on part scores, such as IQ scores for *Gc* and *Gf*, in addition to the traditional composite IQ. That is, the traditional composite IQ may not be dropped, but greater emphasis will be placed on part scores than has been the case in the past. As this movement to part scores develops, it will most likely occur first for *Gc* and *Gf*, the most central of the second-stratum factors, and then extend to other second-stratum dimensions as they are determined to be useful for differential prediction.

INTELLIGENCE TESTS COMMONLY USED IN THE DIAGNOSIS OF MENTAL RETARDATION

Given the widespread development of intelligence tests during the past 100 years, but especially during the past 20 years, many instruments with different theoretical orientations and quality can be employed to diagnose mental retardation. Table 3-1 identifies 13 instruments commonly used in the assessment of intelligence and the diagnosis of mental retardation among infants, children, adolescents, and adults. In addition to these 13 instruments, additional comprehensive intelligence tests are available to psychologists (e.g., McCarthy Scales of Children's Abilities—McCarthy, 1972; Cattell Infant Intelligence Scale—Cattell, 1940). However, these additional tests are not included in the table because they lack norms or because their norms and stimulus materials are too outdated to recommend their use.

Also, several brief or unidimensional intelligence tests are currently available for the screening of intellectual functioning (e.g., Kaufman Brief Intelligence Test, KBIT—Kaufman & Kaufman, 1990; Test of Nonverbal Intelligence: Third Edition, TONI-III—Brown et al., 1997). Although these brief tests may have merit for use as cognitive screeners, they are best suited for low-stakes decision making because of their brevity and limited sampling of important theoretical facets of intelligence. Consequently, intellectual screening instruments are not included in the table.

Finally, a considerable number of group-administered intelligence tests (e.g., Otis-Lennon School Ability Test, OLSAT—Otis & Lennon, 1979) are also available. But such group-administered instruments, while suitable for group screening and decision making, are not designed or appropriate for high-stakes individual disability diagnosis and decision making. Therefore, of all the intelligence tests published and available in their many forms, the instruments cited in Table 3-1 include current, individually administered, comprehensive tests of intelligence suitable for disability diagnosis and eligibility determination. It should be stated, however, that at some point in the future each of the instruments listed may also become outdated, unless they

TABLE 3-1 Comprehensive Tests of Intelligence

Intelligence Test	Age Range[a]	Publication Date	Publisher Level[b]
Bayley Scales of Infant Development-II	Birth to 42 months	1993	C
Cognitive Assessment System	5-0 to 17-11	1997	C
Differential Ability Scale[c]	6-2 to 17-11	1990	C
Kaufman Assessment Battery for Children[d]	6-2 to 12-6	1983	C
Kaufman Adolescent and Adult Intelligence Test	11-0 to 85+	1993	C
Leiter International Performance Scale-Revised[e]	2-0 to 20-0	1997	C
Mullen Scales of Early Learning[c]	Birth to 68 months	1995	C
Stanford-Binet Intelligence Scale: Fourth Edition[b]	2-0 to 24	1986	C
Universal Nonverbal Intelligence Test[e]	5-0 to 17-11	1998	C
Wechsler Adult Intelligence Scale-III	16 to 89	1997	C
Wechsler Intelligence Scale for Children-III	6-0 to 16-11	1991	C

blication te	Publisher Level[b]	Appropriate for MR	Appropriate Scores
93	C	Conditional[c]	Mental development index
97	C	Yes	Full-scale standard score
90	C	Yes	Verbal ability Nonverbal ability General conceptual ability
83	C	Yes	Mental processing composite
93	C	Yes	Fluid scale Crystallized scale Composite intelligence scale
97	C	Yes	Full-scale IQ
95	C	Conditional[c]	Early learning composite
86	C	Yes	Abstract/visual reasoning Verbal reasoning SAS composite
98	C	Yes	Reasoning Memory Full-scale IQ
97	C	Yes	Verbal scale Performance scale Full-scale IQ
91	C	Yes	Verbal scale Verbal comprehension index Performance scale Perceptual organization index Full-scale IQ

Continued on next page

TABLE 3-1 Continued

Intelligence Test	Age Range[a]	Publication Date	Publishe Level[b]
Wechsler Preschool and Primary Scale of Intelligence	11-2 to 7-3	1989	C
Woodcock-Johnson Psycho-Educational Battery-III	2-0 to 90+	2001	C

NOTE: Comprehensive intelligence tests are those that assess intelligence or early cognitive development through multiple subtests and factors, and assess a variety of cognitive processes.

[a] Ages are specified in years-months: 5-0 is 5 years, 0 months of age.

[b] Test publishers use criteria for purchasing tests, with different levels of tests requiring different levels of training and/or credentials. Most comprehensive intelligence tests are known as Class C tests, which require the highest level of training and credential to purchase. Qualification guidelines used by The Psychological Corporation, which is similar to other publishers, to purchase a Class C test requires: "Verification of a PhD-level degree in psychology or education or the equivalent in a related field with relevant training in assessment OR Verification of licensure or certification by an agency recognized by The Psychological Corporation to require training and experience in a relevant area of assessment consistent with the expectations outlined in the 1985 *Standards for Educational and Psychological Testing*."

are revised and renormed. In addition, new instruments may be developed and considered appropriate for inclusion on the list of appropriate instruments. Thus, the list presented in Table 3-1 should be viewed as being valid today, but the equivalent list of appropriate tests is likely to change over time as old tests become outdated and new tests are developed.

The instruments listed in Table 3-1 can be thought of in a variety of ways. For example, some instruments, like the Cognitive Assessment System (CAS) and the Kaufman Assessment Battery for Children (K-ABC) are designed as "process" oriented tests that are intended to be sensitive to the processing aspects of intelligence and are based on neuropsychological theories, such as Luria's conceptualization of brain function and activity. Other instruments, like the Stanford-Binet

blication te	Publisher Level[b]	Appropriate for MR	Appropriate Scores
89	C	Yes	Verbal scale Performance scale Full-scale IQ
01	C	Yes	General intellectual ability

[c] Infant scales may be used for identifying developmental delay that is in the mentally retarded range of functioning, but many psychologists and professional groups defer diagnosis of mental retardation based on developmental scales during the infant/toddler years.

[d] The K-ABC is currently undergoing revision and will be available in two or three years.

[e] The Leiter-R and UNIT are explicitly designed to assess intelligence in a nonverbal administration format. Such tests are employed when language-loaded intelligence tests may provide distorted portrayal of the client's current level of intellectual functioning due to limited English proficiency, language-related disabilities (e.g., verbal learning disability, speech disorders), certain psychiatric conditions (e.g., autism, selective mutism), or some neurological disorders.

Fourth Edition and the Wechsler scales, are product-oriented measures that tend to assess the outcome of a lifetime of knowledge acquisition. Two instruments, the Leiter International Performance Scale-Revised (Leiter-R) and the Universal Nonverbal Intelligence Test (UNIT), were designed specifically for use when an examinee's limited language facility makes it difficult to assess his or her overall cognitive functioning. This could occur with ethnic minorities, individuals who speak English as a second language, individuals who are deaf or hard of hearing or autistic or selective/elective mutes, and others. In such instances, language-based intellectual assessments may produce "construct irrelevant variance." That is, test scores may be contaminated by variance related to a confounding influence like poor English facility, resulting in an IQ that is not a good indicator of the person's true ability.

Another distinction among the instruments listed in Table 3-1 is the number and types of intellective factors or abilities assessed. The Wechsler scales, for example, are composed of two major subscales (verbal and performance), with three or four cognitive factors that better explain the tests' true theoretical underpinnings. At the other extreme of sheer numbers of abilities assessed by a test, the cognitive battery of the Woodcock-Johnson Psycho-Educational Battery-III (Woodcock et al., 2001) purports to assess seven distinct cognitive factors. Most of the instruments cited in the table assess between three and five cognitive factors, with support for their theoretical underpinnings adequate to warrant their use in the diagnosis of mental retardation.

ASSESSMENT CONDITIONS THAT AFFECT INTELLIGENCE TEST SCORES

Intelligence test scores have considerable weight in diagnostic determination of mental retardation. Because of the importance placed on IQs, it is essential that examiners ensure that these scores are obtained in the most objective, clinically appropriate, and standardized fashion.

Test scores too frequently are assumed to be precise estimates of an individual's intellectual functioning, without thoroughly considering the conditions under which the scores were obtained. Four major influences on an individual's performance on an intelligence test should be considered in making diagnostic decisions or recommendations for intervention (Bracken, 2000). Each of the four poses threats to the validity of assessment results and all of them can be controlled to some considerable degree: (1) examinee characteristics, (2) examiner characteristics, (3) environmental influences, and (4) psychometric characteristics of tests. Table 3-2 summarizes these characteristics and gives examples of each. Test results should not be used for making diagnoses or program eligibility decisions without a statement from the examiner attesting to the validity of the evaluation results with regard to these four threats to validity.

Examinee Characteristics

Examinees approach intellectual assessments from differing sociocultural backgrounds and experiences. Some examinees may mistrust the “system” that is mandating the assessment, whereas other examinees may be challenged and highly motivated to participate. In a program like SSI that uses a form of intellectual means testing to identify participants, examiners must be aware of the risk of intentional faking or malingering by or on behalf of examinees. That is, examinees may be motivated to perform poorly intentionally in an effort to receive desired benefits or preferential treatment. Similarly, when parents or other parties who may benefit from assessment outcomes are involved in the diagnostic process by answering background information questions or responding to adaptive behavior measures, the veracity of the participants’ responses also needs to be considered and evaluated.

Because the results of intellectual assessments typically are associated with important decisions and outcomes, examinees should not be assessed unless they appear suitably healthy and well rested. If they exhibit symptoms of poor health, like cold or influenza symptoms, or symptoms of psychological disorders or distress, like depression or acute anxiety, that could adversely affect the assessment of the examinee’s cognitive functioning, the evaluation should be rescheduled for a later date after these conditions have abated or after they have been addressed adequately. In instances in which examinees have had an ongoing history of physical illness or mental health problems, the effects of these conditions on the examinee’s cognitive functioning must also be considered.

Examiners must also ensure that examinees have the requisite skills to perform all intelligence test tasks and activities when selecting instruments or assessment procedures. An examinee with impaired vi-

TABLE 3-2 Potential Threats to the Validity of Intelligence Test Results

Sources of Threat	Examples
Examinee Characteristics	
Transient health conditions	Influenza, colds, fever, minor injuries
Chronic health conditions	Otitis media, speech impairment, diabetes
Transient mental conditions	Recent trauma, acute situational anxiety
Ongoing mental conditions	Psychiatric disorders
Attitudinal conditions	Malingering, oppositional /defiant, uncooperative
Physical conditions	Hearing, vision, motor, neurological limitations
Social/cultural conditions	Linguistic/cultural effects, mistrust of examiner
Examiner Characteristics	
Nonstandardized administration	Failure to administer test in standardized manner
Communication	Failure to establish and/or maintain rapport
Attitude/approachability	Personal bias, prejudice, inability to fairly work with some clients (e.g., certain racial groups, sexual orientations)
Competence/clinical skill	Lack of experience working with some clients (e.g., preschoolers, elderly)
Behavior management	Inability to manage examinee behaviors (e.g., disruptions, poor motivation)
Environmental Characteristics	
Furniture	Inappropriately sized, textured furniture
Examining room conditions	Too cluttered, too cold or hot, poor lighting conditions
Distractions	Excessively noisy, too visually distracting, phones ringing, extraneous noises

sion that is not suitably improved with corrective devices should not be examined using materials that require visual acuity and discrimination. Similarly, individuals with impaired motor skills should not be examined using materials that require fine motor dexterity or processing speed. Examinees with vision, motor, or visual-motor handicapping conditions might better be assessed on verbally loaded measures of intelligence to remove the construct-irrelevant influences of these noncognitive handicapping conditions on their cognitive assessment results. Similarly, examinees who are hard of hearing or who have known speech or language disabilities or who speak and understand English with limited proficiency should not be assessed with language-

TABLE 3-2 Continued

Sources of Threat	Examples
Psychometric Characteristics	
Norms	Old norms, nonrepresentative norm samples, insensitive norm tables, small normative samples
Reliability	Excessive measurement error in scale, internal consistency, stability, interrater judgment
Validity	Threats to internal and external sources of test validity—nonsupportive factor structure, poor criterion-related validity, poor convergent/ discriminative validity
Item gradients	Too few items to allow for fine levels of ability discrimination
Ceilings/floors	Ceilings and floors that artificially limit an examinee's level of performance
Skill demands	Inappropriate skill demands for certain clients (e.g., language demands for examinees who speak English as a second language; performance tasks for motorically disabled clients)
Spoiled subtests/scales	Subtests that are spoiled for any reason (e.g., examiner, examinee, environmental)

loaded intelligence tests, but rather should be assessed with nonverbal tests of intelligence. In all such instances, the examiner must use sound professional judgment when selecting appropriate instrumentation to render a valid assessment of the examinee's true cognitive functioning. Social Security Administration (SSA) evaluators should ensure that they do not apply the same psychological battery to every client, without regard for its appropriateness. Examiners should carefully craft assessment batteries to fit the unique needs of each client.

Psychological examiners are responsible for ensuring that examinees are sufficiently healthy, motivated, and cooperative and that they have the requisite skills and abilities to participate in the assessment before attesting to the validity of test results. When examinees' mental or physical health or their effort or requisite skill levels are such that the validity of the test results are threatened, examiners have an obligation to select more appropriate assessment procedures or make known their reservations about the validity of the test results. Diagnoses should be deferred whenever test results are considered insufficiently valid to contribute meaningfully to such important decisions.

Examiner Influences

Examiners have the potential to significantly affect examinees' test performance, and therefore they should have had proper training, supervision, and experiences to conduct individual intellectual assessments for all the clients with whom they work. In addition, examiners must present an overall demeanor that creates an optimal assessment environment. Examiners must hold the required credentials to perform intellectual assessments in their respective locales, and they should ensure that they administer tests only in the manner in which the instruments were standardized and intended to be used. For example, they should use full-scale, normed versions of the instrument rather than employing abbreviated versions and should not modify test directions.

The ethical standards of American professional and scientific psychological associations, like the American Psychological Association and the National Association of School Psychologists, require that psychologists not engage in services for which they lack competence. Whether the examiner is a psychologist or holds other acceptable credentials to provide psychological assessment services, it is essential that examiners provide only those services they are competent to perform. Because Supplemental Security Income (SSI) benefits are distributed to people across the entire life span from infant to adult and are allocated to members of all racial, ethnic, and linguistic groups, examiners must be properly trained and experienced to work with such a diverse clientele.

Examiners should refrain from providing assessment services to any demographic group member whom they feel inadequately prepared to test. For example, examiners who are comfortable working with school-age children and adolescents may not have experience assessing infants and preschool children. Similarly, examiners may not have the linguistic or cultural competencies to fairly assess examinees whose nations of origin are other than the United States and whose primary languages are other than English. Examiners lacking the prerequisite skills and experience should acquire them through postgraduate or in-service training with supervision prior to attempting assessments with such a diverse clientele. Good professional practice requires that examiners who do not possess the required skills and experience refer clients to other examiners who do.

Examiners should also ensure that rapport is well established and that a businesslike atmosphere conducive to intellectual assessment is created and maintained during testing. Examinees should be comfortable and optimally engaged throughout the assessment process, and the pace of the assessment should be established to minimize examinee fatigue, boredom, distractibility, or other detrimental conditions associated with either a too slow or too rapid assessment pace. The examiner should describe the extent to which these potential threats to assessment validity adversely affected the examinee's performance.

Environmental Conditions

Intellectual assessments should be conducted in settings that are optimal for eliciting the examinee's best performance. Office furniture should be appropriately sized and safe for clients of all ages; for example, preschool children should be seated in small chairs for safety and comfort. Office decor should not unduly distract examinees or interfere with the examinee's ability to focus on stimulus materials and tasks. Examining rooms should be properly ventilated, and the physical climate should be comfortable, with possible sources of distraction such as telephones and beepers eliminated during the assessment. Examiners should ensure that the environment allows examinees to demonstrate optimally their full range of intellectual talents and abilities.

Psychometric Considerations

Bracken (1988) identified 10 psychometric reasons why similar tests sometimes produce dissimilar results. When two tests intended to assess the same construct produce dissimilar results, one or the other or both tests may possess some unique psychometric characteristics that diminish the accuracy of its results for certain populations. Some of these psychometric characteristics, such as limited floors or steep item gradients, are often not as readily noted as other, more obvious psychometric characteristics, like low reliability, yet they must be identified through careful analysis before tests are employed.

Too frequently, examiners assume that test publishers have ensured that tests are equally useful for examinees of all ages and ability levels, but such assumptions are not always warranted. As an example, in the cognitive domain of the Battelle Developmental Battery (Newborg et al., 1984), the item gradient of the memory scale at the 24-35 month age level is too steep for reasonable discrimination of examinees' abilities. On this scale, a raw score of 9 produces a percentile rank of 18, while a raw score of 14 has a percentile rank of 95. Thus, only five items must discriminate across a range of nearly 3 standard deviations, from nearly –1 to +2. Therefore, examiners must acquaint themselves with the examiner's manuals for the instruments they use to determine which instruments may be inappropriate for certain demographic groups. The section on test standards in this chapter addresses these relevant psychometric considerations in more detail.

USE OF TOTAL TEST SCORES AND PART SCORES

Whenever the validity of one or more part scores (subtests, scales) is questioned, examiners must also question whether the test's total score is appropriate for guiding diagnostic decision making. The total test score is usually considered the best estimate of a client's overall intellectual functioning. However, there are instances in which and individuals for whom the total test score may not be the best representation of overall cognitive functioning.

In a compelling article, Jensen (1984) presented a sound argument for generally using total test scores in decision making. His recommendation was to use instruments' total scale composite scores (for example, a composite IQ) rather than the instruments' respective part scores when making diagnostic decisions. Jensen argued that total test scores are more "*g*-saturated"; that is, they are better representations of general intelligence than part scores because they combine and reflect the contributions of all the respective individual part scores from the instrument. Whereas part scores tend to reflect specific abilities like verbal skills or performance abilities, total test scores combine examinees' various skills and abilities to better reflect the client's overall cognitive functioning.

In a similar vein, Spearman (1927) argued that because psychometric *g*, an instrument's average loading on the general ability factor in exploratory factor analyses, permeates all cognitive tasks to some considerable degree, the test content or the specific abilities like memory, spatial ability, and reasoning assessed by tests may be less important considerations when selecting instruments than their *g*-loadings. Two tests of differing theoretical orientations or content may be equally strong measures of psychometric *g* and, as such, may represent equally good measures of general intelligence. Spearman (1927) also coined the term "indifference of the indicator" to describe the phenomenon that test content, process, or theoretical orientation is secondary to how well the test measures *g*, and that tests with comparable *g*-loadings can be used interchangeably as overall measures of intelligence regardless of how they go about assessing *g*.

From a practical standpoint, the total test score of an intelligence test best approximates an instrument's overall *g*-loading. Total test scores produce the highest percentage of a test's explained, reliable variance, and they are reliably better predictors of more external criteria than are part scores (see Jensen, 1981, 1998). The intelligence test total score is also the single overall fairest predictor for individuals of differing ages, genders, races, and ethnic backgrounds (Jensen, 1980; Reschly, 1981b; Reynolds & Kaiser, 1990).

Total Test Scores

All of the instruments listed in Table 3-1 are first and foremost measures of general intelligence, which is best represented in each instrument's total test score. Although the names applied to the instruments' total test scores vary across instruments (e.g., full scale IQ, mental processing composite, general cognitive index), these global scores tend to be highly correlated and share a common source of variance: general intelligence or psychometric *g*. In this sense, there is little practical difference between what the total test scores are called; they are all representations of overall intelligence and historically have been referred to as IQ or full scale IQ. Given the high correlations among the comparable mean total test scores provided by the instruments cited in the table, these instruments can be thought of as collectively measuring the same construct, general intelligence, although their respective subparts may measure a diverse collection of additional specific cognitive abilities.

It is important to note that these tests can also be thought of as largely interchangeable, except in specific situations related to the unique characteristics of individual examinees, for example, those with limited English proficiency, or because of the psychometric foibles associated with specific tests. Some tests, for example, may have inadequate or barely adequate floors for the diagnosis of mild mental retardation at specific ages. With a minimal raw score of a single item answered correctly on each of the appropriate subtests of the Wechsler Preschool and Primary Scale of Intelligence-R for a child who is two and one-half years old, the test FSIQ barely meets the negative two standard deviation criterion commonly used for the diagnosis of mild mental retardation, which is an FSIQ of 68. This instrument would not be capable of differentiating among mild, moderate, and profound levels of retardation at this age level. In most instances and for most examinees, the instruments cited in Table 3-1 produce comparable total test scores across much of the ability continuum, and in most cases the total test score is the preferred score to use in the diagnosis of mental retardation.

Part Scores

There are occasions when a total test score may not be the best indicator of an individual's overall intellectual functioning, and the examiner must resort to interpreting one of the instrument's part scores as the best indicator of overall intellectual functioning. In such cases, the instrument's total test score may offer little more than an awkward and artifactual "average" of a number of relatively disparate subtests or subscales (i.e., part scores). Whenever an examinee's test performance is highly variable across subtests or subscales of an instrument, the validity and meaningfulness of the total test score must be questioned as a reflection of overall intellectual ability. Before an examiner chooses to employ a part score in place of a total test score for a diagnosis of mental retardation, however, four issues must be considered: the statistical significance of scale differences, the meaningfulness of scale differences, which abilities are appropriate for FSIQ replacement, and the actual magnitude of the composite IQ.

Statistical Significance

The first issue to be addressed when considering replacing a total test score with a part score in the diagnosis of mental retardation is whether a statistically significant difference exists between the subscales that contribute to the total test score. When differences between part scores do not differ significantly from each other, the total test score is unequivocally the best indicator of overall cognitive functioning and should be used for decision making.

Most of the instruments cited in Table 3-1 provide interpretative information that identifies when one or more of the instrument's subtests or subscales differs significantly from the mean of the remaining subtests or subscales or when subtests differ significantly from each other. In such comparisons the examiner frequently has a choice between using 85, 90, 95, or 99 percent confidence levels. Determining that two scales differ significantly in magnitude depends in part on the alpha level used for the basis of significance and the level of confidence desired. Statistically significant differences between scales or subtests are necessary but not sufficient criteria for judging that the total test score is not an optimal representation of the examinee's overall intellectual functioning.

Meaningful Differences

The second issue is the meaningfulness of the difference between two or more statistically disparate part scores. It is common to find that two intelligence test subscales differ significantly (e.g., $p < .05$) from each other for individuals, meaning that the differences in the client's respective intellectual abilities are not likely to have occurred by chance alone. However, differences of such magnitude and larger are quite common in the general population. For example, a difference of one standard deviation (15 IQ points) between the simultaneous and successive subscales of the Cognitive Assessment System (Naglieri & Das, 1997) is statistically significant for the individual, but it occurs among 31 percent of the general population. A similarly significant difference of one standard deviation between the verbal and performance scales of the Wechsler Intelligence Scale for Children-III (WISC-III) occurs among 24 percent of the general population; and the same one standard deviation difference between UNIT primary scales (reasoning and memory) occurs among 28 percent of the general population. Differences of this magnitude, although statistically significant, are not unusual or rare occurrences in the general population.

Before determining that a total test score is not an optimal representation of the examinee's overall intellectual functioning, the examiner must consider both the statistical significance of the difference *and* the relative rarity of those significant differences. Scale differences that occur with less frequency than approximately 25 percent in the general population could be considered unusually rare events and therefore may be considered significant threats to the utility of the total test score as a representation of the individual's overall intelligence. When subscale difference scores are statistically significant *and* are relatively rare occurrences in the general population, then examiners should consider whether the total test score is the best indication of overall functioning or whether one or another of the appropriate subscale scores might be a better representation of the client's overall level of functioning. For example, examinees with limited English proficiency who are tested on the Wechsler scales frequently produce score differences that are both statistically significant and relatively rare between the instruments' verbal and performance subscales. In such cases, the total test score would generally be considered invalid as a measure of the examinee's "true" overall intellectual functioning because limited English facility, and not limited overall intelligence, is likely to have adversely affected and rendered invalid the examinee's assessed verbal IQ and, consequently, the composite IQ.

The client's language difficulty consequently would have had the adverse effect of reducing the composite IQ in direct proportion to its influence on the person's performance on the verbal scale. In contrast to the verbal scale, the examinee's performance on the language-reduced performance scale would have probably resulted in a significantly higher performance IQ. Hence, when there is a significant and relatively rare verbal and performance scale difference for individuals who speak English as a second language, the conclusion to be reached would be that the performance IQ is likely to be the best estimate of the client's overall intellectual functioning.

It should be noted, however, that even the performance scales of the Wechsler series or comparable subscales on other instruments, like the Stanford-Binet IV's abstract visual reasoning subscale, pose considerable language demands on examinees who are not proficient in English. As such, performance scales should be viewed only as a better measure of ability, but not necessarily the best measure of ability.

Appropriate Cognitive Abilities

The third issue addresses which of the various instruments' subscales or the intellectual subskills they measure are of sufficient importance, contributing significantly to the understanding of intelligence, to warrant their individual consideration in the diagnosis of mental retardation. Although there are different theoretical approaches to the construct of intelligence, the Cattell-Horn-Carroll (CHC) theory in particular appears to be more developed than others. Because all facets of even this model are not considered to be equally important and the facets vary in predictive value, an essential question arises: which factors can be used individually in the determination of mental retardation? That is, which factors are sufficiently credible measures of general intelligence to contribute to such important decisions? Some are more obvious than others, most prominently *Gc* and *Gf*.

Historically, both crystallized (*Gc*) and fluid (*Gf*) abilities have been considered substantive facets of intelligence. From a Thurstonian perspective, *Gc* maps closely onto the construct of verbal comprehension and *Gf* maps closely onto Thurstone's concept of reasoning. In a multiple-instrument factor analysis of the Woodcock-Johnson Psycho-Educational Battery and the Cognitive Assessment System, subtests from both of these broad ability factors load at moderate to high levels (.60s - .70s) on the *g*-factor (Timothy Z. Keith, personal communication, June, 2001), whereas subtests from some other areas, such as long-term retrieval, *Glr*, short-term memory, *Gsm,* and auditory processing, *Ga*, load at much lower levels (.30s - 50s). Visual (*Gv*) and spatial (*Gs*) subtests tend to be moderate *g*-loaders, ranging in the .50s and .60s. Kaufman (1975, 1979, 1994) suggested a convention for rating the value of subtests *g*-loadings: .70 and above are considered "good" *g*-loaders, .50 to .69 are "fair" *g*-loaders, and *g*-loadings below .50 are considered "poor." Given this convention, part scores derived from crystallized, fluid, and visual/spatial measures appear to be acceptable measures of general ability, in addition to the specific abilities they assess.

The traditional and practical dichotomy of verbal comprehension (*Gc*) and spatial reasoning (*Gf*) appears to represent a reasonable collection of subscales that could be used differentially in the diagnosis of mental retardation because subtests in these domains typically are considered "good" *g*-loading tasks. Such a dichotomy would allow the differential use of the Wechsler verbal and performance scales (or preferably, the factorially purer verbal comprehension and perceptual organization indices) and the verbal reasoning and abstract visual reasoning scales of the Stanford-Binet Intelligence Scale, Fourth Edition, for the purpose of diagnosing mental retardation.

Similar divisions within other instruments might also be considered appropriate, such as the simultaneous and successive scales of the CAS and K-ABC, but such decisions should be based on the instruments' respective subtest *g*-loadings, with only those scales that have subtests that are predominately moderate to high *g*-loaders being used. For example, the two primary scales (memory and reasoning) of the UNIT (Bracken & McCallum, 1998) comprise six subtests that measure either complex memory or reasoning. Of the three memory and three reasoning subtests, all have *g*-loadings above .70, except one, the reasoning subtest (mazes < .50). Using the criteria of substantive contribution to *g*, either of these two primary scales may be considered appropriate for use in the diagnosis of mental retardation because of their significant *g*-loadings.

Magnitude of the Total Test Score

The last issue when considering whether a part score should be used in place of a total test score is the magnitude of the existing total test score. That is, when scale score discrepancies meet the previously mentioned criteria of significance and meaningfulness, the total test score may be simply too high to support a diagnosis of mental retardation. For example, one scale score might barely qualify for a diagnosis of mental retardation (e.g., verbal IQ near 70), while the second scale score may be considerably higher (e.g., performance IQ in the average range). In such cases, which are usually rare once significance and meaningfulness have been assessed, the resulting composite IQ would be in the low average range, and, in the committee's opinion, the individual would not be likely to truly have mental retardation, despite one scale score in the retarded range. Accuracy of diagnosis is vitally important to the individual client and to SSA, because the stakes are so high. It is as important to include appropriate individuals as it is to exclude inappropriate ones from the SSI and Disability Insurance benefits programs. In the committee's view, comprehensive intelligence tests provide the greatest technical adequacy and construct sampling and result in the best assessment of intelligence. Therefore, the final criterion for deciding whether or not to use part scores in place of the total test score in the diagnosis of mental retardation is that, no matter how great the discrepancy between relevant subscales, individuals with total test scores greater than 75 should not be diagnosed as having mental retardation.[1]

Composite scores from intelligence tests should be used routinely in mental retardation diagnosis, except when the composite IQ validity is in doubt, in which case an appropriate part score may be used in its place. Significant and meaningful variation among an instrument's respective part scores may indicate evidence of compromised validity for one or more of them (for example, a low verbal scale score for an individual with a suspected speech disorder), which in turn would threaten the validity of the composite IQ. In such situations, appropriate part scores may better represent the individual's true overall level of cognitive functioning.

---

[1] Committee member Keith Widaman disagrees with this statement. Dr. Widaman believes that IQ part scores representing crystallized intelligence (*Gc,* similar to verbal IQ) and fluid intelligence (*Gf,* related to performance IQ) have clear discriminant validity and represent broad, general domains of intellectual functioning. Therefore, a score of 70 or below on either of these part scores from any standardized, individually administered intelligence test that reports such scores should be deemed sufficient to meet the listings for low general intellectual functioning regardless of the level of the composite score, providing that the part scores have adequate psychometric properties (e.g., high reliability, low standard error of measurement).

However, only part scores derived from scales that demonstrate high *g*-loadings (e.g., crystallized, fluid, visual/spatial measures of intelligence) should be used in place of the composite IQ score when its validity is in doubt. Many intelligence tests assess several facets of intelligence, but not all facets are equally important or predict life events equally well. Those intellectual facets that are heavily *g*-saturated provide the best sources for replacing the composite IQ score when its validity is questionable.

The characteristics of comprehensive IQ tests are such that, even when part scores are used in making disability determinations for mental retardation, the composite IQ score from an instrument should never be higher than 75. Furthermore, if a part score is used in place of the composite IQ score in SSA decision making, the part score should not exceed 70.

The committee considered a number of alternatives before recommending, under certain circumstances, the use of part scores in disability determination for mental retardation. Alternatives included: (1) recommending that SSA continue with its current practice of allowing the use of part scores in diagnosing mental retardation; (2) recommending against any use of part scores, with eligibility determinations made solely on the basis of composite IQs; and (3) recommending that the composite IQ be used, but also allowing for the use of part scores from various instruments, in certain circumstances.

The committee first considered endorsing current SSA practice, which allows the use of a valid verbal performance or full-scale (composite) IQ from an individually administered intelligence measure. In common clinical practice, this usually results in the use of a Wechsler VIQ, PIQ, or FSIQ, a situation that unfairly privileges one set of intelligence tests and has the effect of discouraging innovation on the part of other test developers. Furthermore, the Wechsler part scores VIQ and PIQ have poor theoretical and weak or mixed empirical support for their distinctive status. The current science of the structure of intelligence suggests that the Wechsler Verbal Comprehension and Perceptual Organization Indexes are better measures of *Gc* and *Gf* than the VIQ and PIQ.

The committee also considered recommending that SSA allow the use of only composite IQs for disability determinations. This recommendation would have made SSA's definition of mental retardation consistent with that used by other professional associations and health-related organizations, all of which identify significantly subaverage general intellectual functioning as characteristic of mental retardation. In this situation, significantly subaverage *general* intellectual functioning would suggest that the deficits must be evident on the overall index of functioning, or the composite IQ. The committee decided against such a recommendation for two reasons. First, the practical consequences of declaring that a practice long used by SSA was invalid would have caused significant disruption for the agency and for disability benefit recipients. Second, and also important, was the recognition that there are circumstances, described earlier in this chapter, in which the composite IQ does not represent the person's true intellectual functioning and is instead a meaningless artifact.

The recommendation eventually adopted by the committee advises that part scores not be used routinely in mental retardation determinations, except in those cases in which the composite IQ is thought to be invalid. Only then can an appropriate part score be used as the measure of the person's intellectual functioning. The committee opted to bring SSA's definition of the intellectual functioning dimension of mental retardation more in line with that of the other professional associations and health-related organizations, which focus on the summary measure of intelligence. Since there are some situations in which the composite IQ is invalid, part scores may more accurately reflect a person's intellectual functioning. The committee's examination of the structure of intelligence suggests that part scores that measure crystallized and fluid intelligence are the most appropriate part scores to use in these situations. Also, the committee recognizes that many of these abilities are measured by a wide number of intelligence tests, not just Wechsler measures, and therefore recommends that SSA expand in its listings the use of examples of other apropriate tests that yield *g*-loaded part scores. The text of this recommendation appears at the end of this chapter.

In rare instances it may be impossible to develop reliable and valid assessments of intellectual functioning even with the use of specially designed instruments that attempt to limit the effects of language differences, sensory or neuromotor impairments, and severe emotional disturbance. In such cases all of the summary scores, both the composite and part scores, may be suspected of being invalidly low. Invalid intelligence test results in the range of mental retardation, whether too low or too high, should always be ignored and other methods used to confirm or disconfirm a diagnosis of mental retardation, such as case history information, educational performance, social functioning across a variety of settings, adaptive behavior, and interviews with the individual and significant others. The principle of convergent validity should be applied to the interpretation of this information (see Chapter 5) and diagnostic decisions made based on the preponderance of evidence.

MULTIDIMENSIONAL VERSUS UNIDIMENSIONAL MEASURES OF COGNITIVE FUNCTIONING

The *Standards for Educational and Psychological Testing* (Joint Committee on Standards for Educational and Psychological Testing of the American Educational Research Association, the American Psychological Association, and the National Council on Measurement in Education, 1999) discuss the importance of using tests of differing length and psychometric quality for high-stakes versus low-stakes decisions. Intelligence testing intended for high-stakes decision making, such as disability diagnosis or eligibility determination, should include multidimensional measures of important intellectual factors like high *g*-loading tasks rather than unidimensional measures. The unidimensional Peabody Picture Vocabulary Test (PPVT—Dunn, 1959) originally reported an IQ as its total test score—a score that was once used for high-stakes placement and eligibility testing. During the 1960s and 1970s, the field came to the realization that, although the PPVT assessed a singularly important aspect of intelligence, verbal comprehension (crystallized abilities), it correlated to a relatively modest degree with comprehensive, multidimensional tests of intelligence. As a result, this unidimensional test was deemed insufficiently comprehensive to warrant using the term IQ. The revised PPVT (PPVT-R—Dunn & Dunn, 1981) did not continue the practice of using the term IQ for the total test score. Bracken et al. (1984) recommended further that the instrument not be considered or used as a general measure of intelligence.

In addition to tests designed as unidimensional ability measures, like the PPVT-R and the Raven Progressive Matrices (Raven et al., 1986), abbreviated versions of comprehensive tests (WISC-III short forms, the Wechsler Abbreviated Scale of Intelligence—Wechsler, 1999) and screening tests (KBIT—Kaufman & Kaufman, 1990) have been developed. These shortened tests typically have limited construct sampling and consequently have reduced levels of reliability and validity compared with comprehensive measures of intelligence, and consequently they should be reserved for low-stakes decision making.

When intelligence testing is conducted for high-stakes purposes, multidimensional, full-scale instruments should be used in the diagnostic, decision-making process because these instruments provide the most convincing evidence of technical adequacy, construct sampling, and validity. Comprehensive intelligence tests assess multiple facets of the construct, and they more thoroughly sample the domain of intelligence. The instruments listed in Table 3-1 represent a current compendium of comprehensive measures of intelligence for infants, children, adolescents, and adults.

PSYCHOMETRIC STANDARDS

The psychometric quality of tests should guide examiners' selection of tests used to contribute to the diagnosis of mental retardation. Due to the nature of this disability and the unique characteristics of individual intelligence tests, all comprehensive tests may not be appropriate for this application—at least, all tests may not be appropriate for all examinees. Both pragmatic and empirical aspects of test quality should guide test selection and inform decision making. Bracken (1987, 1988, 1998; Wasseman & Bracken, 2002) have proposed criteria to guide the selection of cognitive tests. In some instances these guidelines include more breadth and specificity than the *Standards for Educational and Psychological Testing* (Joint Committee on Standards for Educational and Psychological Testing of the American Educational Research Association, the American Psychological Association, and the National Council on Measurement in Education, 1999), yet in general they are consistent with recommendations provided by such psychometricians as Anastasi and Urbina (1997), Cattell (1986), Cicchetti (1994), Nunnally and Bernstein (1994), and Salvia and Ysseldyke (2001).

Intelligence Test Norms

The adequacy of a test's norms is of paramount importance when selecting a norm-referenced intelligence test. The quality of test norms is dependent on several factors, including sample size and population representation. Cronbach (1949) posed the following questions, which remain pertinent today, when assessing the quality of test norms: "(1) Are the norms based on a sufficiently large group? (2) Is the standard group representative? (3) Does the standard group resemble the persons with whom we wish to compare our subject?" (pp. 75-76).

The primary goal in normative sampling is to accurately reflect population parameters, which allows inferences based on obtained scores to be generalized to the population. The goal of intelligence test norms is to accurately represent the U.S. population because the goal of assessment is to identify the degree to which an individual deviates from normative expectations. In test norm development, sampling plans should sample representatively from among all potential examinees to reflect the entire distribution of ability, including individuals who have mental retardation.

Some test developers employ truncated selection procedures that do not sample the entire population and systematically exclude individuals with impairments or other special needs (McFadden, 1996). Such norming practices should be avoided and tests used for the identification of mentally retarded individuals should include a representative proportion of such individuals in the test normative sample as are found in the general population. Hollon and Flick (1988) recommended that when tests are developed for use with special populations, norms still should be based on fully representative samples.

Sampling plans should be thoroughly described. Two of the principal assumptions of random sampling are that every individual in the target population has an equal chance of being selected and that every sample selection is made independently. However, true random sampling is an ideal that is rarely if ever achieved in test norming. Given the geographic expanse of the United States and its population of approximately 280 million citizens, random sampling from the entire U.S. population is typically not economically feasible or practical. As a reasonable compromise in test norming, intelligence test norms should be gathered in a stratified sampling manner that results in a sample that is demographically representative of the population, including all of its relevant characteristics.

Normative samples should be sufficiently large. Intelligence test normative samples should be sufficiently large to provide stable estimates of population parameters, thereby reducing sampling error to acceptable levels and meeting assumptions for requisite statistical analyses. Although large-scale group tests may involve 10,000 to 20,000 students per grade or age level, samples for individually administered intelligence tests generally are considerably smaller. Carefully drawn samples of 150 to 200 participants per grade or age level are typically considered appropriate and are frequently employed with individually administered tests. The smaller the sample size, the less likely the sample is to be normally distributed or to accurately reflect population parameters. Therefore, tests with norms based on samples smaller than the minimal level noted above should be avoided, unless additional evidence that supports their use is available.

Normative samples should reflect appropriate demographic pa-

rameters. Research has shown that certain demographic variables are more related to levels of cognitive functioning than are other variables. It is important that norm samples are stratified and selected on the basis of these identified variables. Variables used for selection and stratification when gathering samples for intelligence test norms generally include age, grade level, sex, ethnic origin, race, geographic region, urban or rural residence, and socioeconomic status. Intelligence tests used for the diagnosis of mental retardation should include carefully selected samples that fully represent these important demographic characteristics to the degree that they are found in the general population.

Many intelligence tests also appropriately include individuals with handicapping conditions and educational exceptionalities in their normative samples. The inclusion of exceptional individuals in norming samples is based on the logic that the intended function of the normative sample is to represent accurately the population, and the intended function of the test is to serve a comprehensive group of individuals rather than only people without known deficits or gifts (Elliott, 1990). If an intelligence test is intended to diagnose and serve individuals with mental retardation, then the test should include proportionate representation of this population in the normative sample.

Sampling should be representative and precise. The accuracy and precision of a stratified sample is most readily determined by the degree to which the sample matches the sampling plan. The degree to which the composition of an acquired sample reflects census proportions should be assessed through examination of not only single demographic characteristics like gender for the entire sample but also by examining combined demographic sampling cells (e.g., gender by race within individual age levels). It is in these smaller cells that sampling plans typically fail most often. Examiners should carefully examine sampling outcomes to ensure that selection variables are accurately represented not only across the entire norm sample, but also within each level of the norm sample (say, for 5-year-olds or 20-year-olds) and for each group sampled, such as blacks or females.

All statistical transformations used to develop interpretive scores should be reported in the examiner's manual. Because raw scores have limited norm-referenced interpretative value, they must be transformed into more meaningful metrics, like standard scores. The statistical procedures by which raw scores are transformed into standard scores should be clearly documented in the test manual, including procedures used to smooth, normalize, or stretch distributions during the transformation process.

One consideration in transformation from raw score to standard score is whether scores were manipulated through sample weighting. Weighting is not necessary with most carefully normed intelligence tests; however, sometimes weighting is done to "correct" errant samples when the stated goals of the sampling plan have not been adequately met. When specific demographic strata have been undersampled, score weighting is sometimes used to statistically correct this methodological slight. It should be recognized, however, that weighting scores often increases sampling error because the "corrected" scores are based on smaller and probably less representative samples than appropriate. Weighted scores in general should be viewed as an undesirable characteristic of test norms and should be carefully considered when selecting tests.

Similar to weighting is the issue of extrapolated score development. When normative samples are not sufficiently diverse in their range of talent, it sometimes happens that there are too few low- or high-functioning individuals to properly generate norms for individuals who are functioning in the mentally retarded or gifted range. In situations in which exceptional individuals are excluded from the norming process, there may be too few people with mental retardation assessed to establish norms at this level. Consequently, test publishers often "stretch" norms beyond their actual range through linear extrapolations. Extrapolation provides the benefit of extending norms farther than would otherwise be the case; however, extrapolated norms provide no assurance of accuracy because they are not based on obtained data. It is not known for certain whether cognitive functioning progresses through the population in a linear fashion, and the application of linear extrapolation is merely a best guess at what the norms would have been if sufficient numbers of exceptional individuals had been included in the sampling plan.

Standardization examiners and procedures should be clearly described. Procedures used to recruit, qualify, and train standardization examiners should be carefully described in intelligence test manuals. Quality assurance procedures intended to correct invalid administrations and to identify invalid test protocols also should be detailed. Ideally, standardization examiners should have the same credentials and experience as the professionals who will be administering the test.

Test manuals should carefully describe the standardized test conditions under which the test norms were established. These conditions should be the same when the test is employed in clinical practice. Any changes in artwork or stimulus materials, instructions, and test or item sequence after standardization should be described in the test manual.

The standardization sample should be current. Research suggests that intelligence in the entire population increases at a rate of approximately 3 IQ points per decade, which approximates the standard error of measurement for most comprehensive intelligence tests. Thus, tests with norms older than 10 to 12 years will tend to produce inflated scores and could result in the denial of benefits to significant numbers of individuals who would be eligible for them if more recent norms had been used. Disability examiners who use tests with outdated norms may be systematically if unintentionally denying benefits to those who are legally entitled to them. The examiners also risk losing their licenses for ethical violations of their professional codes. Proper test usage is essential for accurate testing and diagnosis and ultimately for equitable disability determination.

In several meta-analyses, James Flynn (1984, 1987, 1994, 1999) has demonstrated that the age of intelligence test norms may be one of the most important considerations when selecting tests for use. Internationally, Flynn has demonstrated that intelligence test norms "soften" at a rate of about 3 IQ points per decade. That is, a test with 20-year-old norms will tend to produce IQs that are approximately 6 points (> one-third of a standard deviation) higher than a recently normed instrument. Thus, administering outdated forms of intelligence tests like WISC-R instead of the WISC-III may have the unintentional and undesired result of failing to qualify individuals for services or benefits that they would otherwise qualify for. The Flynn effect is noticeable in samples as young as infants (Bayley, 1993; Campbell et al., 1986) and appears to continue throughout childhood and adolescence. Chan et al. (1999) demonstrated that a variety of cognitive abilities, especially those involving more semantically laden content and procedures that measure crystallized abilities, tend to be most susceptible to population changes over time.

This issue is particularly salient for psychologists who habitually use older tests such as the McCarthy Scales of Children's Abilities (McCarthy, 1972) or previous editions of revised tests like the WISC or WISC-R, rather than the newer WISC-III. Tests of this vintage may have norms that are as old as three decades or in some instances even older. Norms of this age would predictably and reliably fail to identify large numbers of individuals who would otherwise qualify for services or benefits. For example, given the 9-point IQ inflation that would be associated with using either the McCarthy or WISC-R rather than a current generation test, many and possibly most individuals who are functioning in the 60-70 IQ range would fail to be properly identified as having mental retardation on either instrument. This issue is also important because ethical codes admonish psychologists from using outdated tests and norms. This view is also supported by many state psychological associations.

Verification of this norm-softening can be seen throughout the literature wherein early researchers discovered that the most recent edition of various intelligence tests produced scores that were significantly lower than the previous edition of the same instrument (e.g., Kaufman, 1979). Similarly, new instruments just entering the field typically produce total test scores that are significantly lower than the scores obtained on the traditional "old standards" used in convergent validity studies—leading to criterion contamination as a major threat to the validation of the newer instrument. For these reasons, professional organizations like the American Psychological Association and the National Association of School Psychologists and the joint AERA, APA, and NCME *Standards for Psychological and Educational Testing* (Joint Committee on Standards for Educational and Psychological Testing of the American Educational Research Association, the American Psychological Association, and the National Council on Measurement in Education, 1999) admonish psychologists to not use outdated instruments.

Given these generational changes in intelligence test norms, tests should undergo normative update, restandardization, or revision at intervals corresponding to the time expected to produce one *SEM* of change. For example, the commonly used WISC-III has a composite IQ *SEM* of 3.20. Given an *SEM* of this magnitude, the WISC-III norms would be expected to soften a significant degree, (3 to 4 points), in 10 to 11 years (Wasseman & Bracken, 2002). Therefore, the WISC-III and most other intelligence tests might be considered inappropriate for the diagnosis of mental retardation when their norms are more than 10 to 11 years old.

A related issue is the length of time that an obtained IQ (or IQ equivalent) can be considered valid. Because intelligence is a quite stable construct, especially among older children, adolescents, and adults, IQs of record may be useful for a number of years beyond the date they were obtained, with the exception of the occurrence of any known condition that might threaten the validity of the obtained score, such as physical or emotional trauma. Despite its general stability, cognitive development proceeds most rapidly during the infant and toddler years and slows thereafter through childhood and adolescence (Bloom, 1964). For adults, formal learning-dependent knowledge (crystallized abilities) and long-term memory continue to improve into advanced years, but fluid abilities like novel problem solving and clerical speed generally decline fairly rapidly after peaking in adolescence (Horn, 1985).

Therefore, during the infant and toddler years, when cognitive growth and development are most rapid and consequently least stable, total test scores should be obtained at the time they are to be used in diagnosis or disability determination. For children between the ages of 3 and 6, total test scores might reasonably be considered valid for one year. Among children and adolescents between the ages of 6 and 16 years, total test scores should be considered valid for as long as three years. For adults ages 18 to 50 living in stable conditions and with stable health, total test scores should be considered valid for as long as five years. After age 50, total test scores might be considered reasonably valid for three years, but separate intellectual abilities, like *Gf-Gc*, might become important considerations. This lack of stability in elderly individuals' specific cognitive abilities is typically due to debilitating factors associated with aging, and, although their IQs may change over the years, their diagnostic status is unlikely to change. That is, adults with mental retardation are likely to become more retarded in their functioning as they age.

Norms should reflect adequate item difficulty gradients. *Item gradients* reflect the degree to which standard scores change as a function of success or failure on a single item (Bracken, 1987). The larger the resulting standard score difference in relation to a change in a single raw score, the less sensitive and discriminating the test is. For a test to have adequate sensitivity at all levels of cognitive functioning, it must have adequate item density across the ability range. Bracken (1987, 1998) has suggested that item gradients should not be so steep that a single item passed or failed would result in a standard score change of more than one-third of a standard deviation.

Similarly, norm table gradients should be sufficiently sensitive that when the same raw score is entered into two adjacent age tables, that score should not produce standard score changes of more than one-third of a standard deviation (Wasseman & Bracken, 2002). For example, the norm tables on the McCarthy Scales of Children's Abilities are insufficiently sensitive at the younger age levels. A child who is 2 years, 7 months, and 16 days old could earn McCarthy total test general cognitive index scores that are more than two-thirds of a standard deviation (11 points) apart when the same raw score is entered into adjacent norm tables (Bracken, 1988). That is, a single day's difference in the child's chronological age would result in the child's graduating from one norm table to the next, and with that movement into another norm table the child would appear to be as much as two-thirds of one standard deviation less intelligent. Such insensitivity in item or norm table gradients could easily lead to misidentification or misdiagnosis, especially among low-functioning individuals.

Test norms should have adequate floors and ceilings. When tests are used to identify individuals who may have mental retardation or giftedness, it is important that the tests have sufficient discriminating power in the extreme ends of the distributions for accurate differentiation of ability and diagnosis. At a minimum, intelligence tests should have floors sufficiently strong to differentiate the extreme lowest 3 percent of the population from the top 97 percent (Bracken, 1984, 1987, 1998; Bracken & McCallum, 1998). Preferably, intelligence tests should be able to discern more severe levels of retardation from mild mental retardation. Although not pertinent to the diagnosis of mental retardation, intelligence tests should also have ceilings that are sufficiently high to differentiate the extreme upper 3 percent from the lower 97 percent.

Evidence of Test Score Validity

The validity of a test is characterized by the extent to which it exclusively measures its targeted constructs (construct validity) and its scores meaningfully guide decision making. Increasing emphasis is being placed on the extent to which test scores serve their intended purposes and proposed applications (Messick, 1995). Construct validity can be supported with two broad classes of evidence, internal and external, which parallel the threats to validity typically considered in research designs (Campbell et al., 1963; Cook & Campbell, 1979).

Internal Evidence of Test Validity

Internal sources of validity include procedures to systematically examine the characteristics of a test, especially its content, assessment methods, structure, and theoretical underpinnings. Internal evidence of test validity can be found in investigations of face validity, content validity, theory-based validity, and structural validity.

Face validity refers to the degree to which a test appears to measure what it purports to measure. During casual examination, test items may be judged for face validity by the extent to which they appear to appropriately measure the targeted construct and objectives. Although not considered a source of validity in a technical sense, face validity has been shown to be related to examinee motivation and effort, as well as social desirability biases, labeling, and fairness (Bornstein, 1996). Most tests selected for the diagnosis of mental retardation include activities and tasks of sufficient difficulty that they readily appear to measure the construct of intelligence.

Content validity can be described as the degree to which a test adequately samples the domains of interest. Content validity varies with the purpose of the test and the nature of the inferences that may be drawn from test scores (Messick, 1993). Inferences made from tests with inadequate content validity may be suspect, even when other indices of validity are satisfactory (Haynes et al., 1995). Ideally, content should remain consistent throughout the age range of a test to ensure that the same construct is being measured (Bracken, 1988). The Stanford-Binet Intelligence Scale, Fourth Edition, includes subtests in which the content assessed is not consistent across the age range. For example, the vocabulary subtest begins with a picture vocabulary format and then graduates to an oral vocabulary format. When test content and item formats change in this manner, it is difficult to interpret an examinee's test performance, because it is no longer clear which construct is being interpreted, receptive or expressive vocabulary.

The formulation of test items and procedures based on and consistent with a theory has been termed *substantive validity* (Loevinger, 1957) and is closely related to content validity. Psychology has produced rich and cohesive theories of behavior and cognition—theories that have led to the development of new tests and assessment practices (e.g., the K-ABC—Kaufman & Kaufman, 1983; CAS—Naglieri & Das, 1997). As Crocker and Algina (1986) suggest, "psychological measurement, even though it is based on observable responses, would have little meaning or usefulness unless it could be interpreted in light of the underlying theoretical construct" (p. 7). Tests used for the diagnosis of mental retardation should be based on reasonable and supportable theories, and these theoretical orientations should be presented in the test manual for consideration.

Composite scores should be supported through factor analyses. Exploratory factor analyses allow for examination of the natural structure of an instrument and the psychological meaningfulness of the dimensions or factors that emerge (Gorsuch, 1983). This criterion refers to the degree to which factor analytic results match the composite scales or subscales of the test. The mismatch between factor structure and composite indices has been shown to render test interpretation more difficult (Chattin & Bracken, 1989) on such tests as the Stanford-Binet Intelligence Scale, Fourth Edition (Thorndike, 1986), and the McCarthy Scales of Children's Abilities (McCarthy, 1972).

Exploratory factor analyses provide a methodology by which the underlying dimensions assessed by a test may be separated or summarized. Floyd and Widaman (1995) suggest that exploratory factor analyses for clinical assessment instruments should routinely report principal component analysis or common factor analysis, initial communality estimates (or squared correlations of observed variables with the factors), the method of factor extraction, the criteria for retaining factors, the eigenvalues and the percentage of variance accounted for by the unrotated factors, the rotation method and rationale, all rotated factor loadings, factor intercorrelations, and the variance explained by the factors after rotation.

Competing models or theories should be tested with confirmatory factor analyses. Confirmatory factor analyses are conducted to evaluate the congruence of the test data with an a priori theoretical model, as well as to measure the relative fit of competing models. Floyd and Widaman (1995) recommend that confirmatory factor analyses should report proposed model(s), number and composition of factors, orthogonal versus correlated factors, secondary loadings, correlated error terms, other model constraints (fixed and free parameters), method of estimation, goodness of fit, overall fit, relative fit, parsimony, any model modification to improve model fit to data, factor loadings and standard errors, communality, and factor correlations and standard errors with statistical significance. Comprehensive treatment and inclusion of such information allows test users to better understand the extent to which the test fits its proposed model compared with competing models and provides support for the interpretation of the instrument's respective subscales and composite scores.

External Evidence of Validity

External evidence of test validity considers the extent to which a test relates to or predicts other variables or outcomes in differing populations. Tests should be validated with regard to the purposes for which they are employed and the consequences of their use. In this section, we describe external classes of evidence for test construct validity, including criterion-related validity, consequential validity, and generalizability.

Criterion-related validity. Campbell and Fiske (1959) originally proposed that test scores should be related to external measures of the same psychological construct (*convergent* evidence of validity), and they should be comparatively unrelated to measures of different psychological constructs (*discriminant* evidence of validity). In criterion-related validity, criterion measures can be obtained concurrently (*concurrent* validity) or at some future date (*predictive* validity). An intelligence test that is proposed for use in the process of diagnosing mental retardation should demonstrate convergent validity with other extant intelligence tests before the instrument is accepted for this purpose. Similarly, as a class of instruments, intelligence tests should demonstrate higher correlations among themselves than with measures of other psychoeducational constructs (e.g., academic achievement, adaptive behavior).

Tests should meaningfully guide decision making. Contrasted groups methodology is commonly used for validating psychological tests. In this approach to validation, the test performance of two samples that are known to be different on the criterion of interest is compared. For example, a sample of people who are known to have mental retardation should perform on an intelligence test at a level significantly below the performance of a second group that is known to not have mental retardation. Decision-making classification accuracy should be determined by examining sensitivity, specificity, positive predictive power, and negative predictive power.

Tests should provide evidence of consequential validity. A form of validity that emphasizes the societal impact of test results on individuals and groups is known as *consequential validity*. Consequential validity evaluates the utility of score interpretation as a basis for action, as well as the actual and potential consequences of test use (Messick, 1989). Messick (1995) argued that examination of the consequences of test use as a trigger to social and educational actions, such as equitable application of SSI benefits, is a necessary element of validating tests. Consequential validity is especially relevant to issues of bias, fairness, and distributive justice.

Generalizability of validity. External evidence of test validity is especially important when test results are to be generalized across contexts, situations, and populations, and when the consequences of testing reach beyond the test's original intent. Intelligence test manuals should demonstrate the extent to which the test validity generalizes across subpopulations, such as racial or ethnic minority groups, gender, or age levels. Examiners who wish to use tests for purposes not stated or supported in the examiner's manual, such as using a language instrument for discerning levels of cognitive functioning, must demonstrate the validity of the new application prior to its application.

Test Score Reliability

The reliability of test scores refers to the reproducibility (precision, consistency, and repeatability) of test results, or the degree to which test scores are free from measurement error. Measurement precision can be assessed by examining the instrument's internal consistency, temporal stability, and interrater agreement. Reliability can only be evaluated in the context of test use (Nunnally & Bernstein, 1994).

Internal consistency. The internal consistency of a test is a reflection of the uniformity and coherence of test items and content. All variance generated by a test can be classified as either reliable variance or error variance. In classical test theory, reliability is based on the assumption that measurement error is distributed normally and equally for all score levels. By contrast, item response theory posits that reliability differs between individuals with different response patterns and levels of ability but generalizes across populations (Embretson & Hershberger, 1999).

Internal consistency is usually coefficient alpha or split-half reliability. Several psychometricians (Bracken, 1987; Clark & Watson, 1995; Nunnally & Bernstein, 1994) have recommended that minimal levels of internal consistency should average across age levels at or above .80 or .90, depending on the nature and applications of the test scale to low-stakes or high-stakes applications, respectively.

Consistent with Nunnally's (1978) original standards, Bracken (1987, 1998; Wasseman & Bracken, 2002) recommended that total test or total scale internal consistency of high-stakes test applications, such as for clinical diagnosis or eligibility decision making, should equal or exceed .90 when averaged across the age levels. Instruments used for the high-stakes purposes of diagnosing mental retardation for SSI should approximate this minimal level of reliability, recognizing that the inverse of reliability is measurement error and that error only confounds correct decision making.

Local reliability. Local reliability refers to measurement precision at specified levels or ranges of scores that are at or near the decision-making point for mental retardation. For example, a test with high local reliability at low ability levels would be more appropriate for use with low-functioning individuals than one with less local reliability. Local reliability can be measured by approaching it from classical test theory orientation or by using item response theory. Whichever approach is used, local reliability should be measured and the data made available for disability determination examiners so they can use the most appropriate tests for their clients.

Total test short-term stability. Test scores must be reasonably stable to have practical utility when diagnosing known stable conditions such as mental retardation and to be predictive of future performance. Stability is typically estimated through use of test-retest stability (correlation) coefficients across two points in time. Bracken (1987) suggested that for short-term test intervals of two to six weeks the total test stability coefficient should be greater than or equal to .90 for high-stakes test applications. Test-retest reliability is in part a measure of construct stability, but its interpretation in clinical contexts can be influenced by several factors like the deleterious effects of degenerative disorders or the positive effects of successful therapeutic interventions, which should be remembered in individual studies of test stability.

Generalizability of test score reliability. As an extension of validity generalization (Hunter & Schmidt, 1990; Schmidt & Hunter, 1977), reliability generalization investigates the stability of reliability coefficients across varying samples. In order to demonstrate measurement precision for the populations for which a test is intended, the test should show comparable levels of reliability across various demographic subsets of the population, as well as salient clinical and exceptional populations like individuals with mental retardation.

Fairness in Testing

Fairness has not been considered historically as a leading criterion by which test selection decisions are made, but increased social sensitivity and recent court decisions have elevated its importance. Tiedeman (1978) has noted, “Test equity seems to be emerging as a criterion for test use on a par with the concepts of reliability and validity” (p. xxviii). As such, tests intended for use with all subsets of the U.S. population, as in SSA evaluations, should provide ample evidence of psychometric fairness and equitable treatment of examinees.

Wasseman and Bracken (2002) consider fairness to be the extent to which test scores are (a) statistically shown to be free from evidence of psychometric bias, (b) comparably reliable and valid across demographic groups, and (c) equitably applied and equally predictive in real-life consequences and pragmatic impact. Fairness transcends psychometrics and includes philosophic, legal, and practical considerations.

Test bias refers to elements of a test and its usage that are construct irrelevant *and* that yield systematic errors that in turn lead to erroneous decisions related to specific demographic group membership. Bias results in differential outcomes for individuals of the same ability levels but from different ethnic, sex, cultural, or religious groups (Hambleton & Rodgers, 1995). Test bias has also been described as "a kind of invalidity that harms one group more than another" (Shepard et al., 1981, p. 318)

Internal Evidence of Fairness

As with internal evidence of validity, test fairness rests in part on the structural features of the instrument, including theoretical underpinnings, item content, assessment procedures, differential item functioning, and an invariant factor structure.

Theoretical underpinnings. The theory on which a test is built may have an inherent sensitivity to issues of fairness and should be fully discussed in the test manual. Several illustrations of these implications may be presented with regard to measures of cognitive and intellectual ability. For example, tests that emphasize speed may be less fair for Hispanics, because time is considered a less salient concept in many Hispanic cultures (Scheuneman & Oakland, 1998). Individuals who speak English as a second language also may be disadvantaged by traditional language-loaded intelligence tests, even on performance-based measures like the Wechsler Performance Scale that include lengthy and conceptually laden test directions (Bracken & McCallum, 1998; Duran, 1989; Geisinger, 1992; Oakland & Parmelee, 1985). In addition, measures of crystallized ability and knowledge are inextricably linked to culture (Carroll, 1997) and accordingly may show differential performance across culturally different groups, whereas fluid abilities tend to show less differential performance across groups.

Multicultural bias and sensitivity reviews. The use of multicultural reviewers to examine the type, content, and format of test items for potential bias is a common practice among test publishers. Usually the goal of bias review panels is to identify offensive or controversial material and unfair material, remaining sensitive to population diversity. Among the considerations of such reviewers are language usage, ethnocentric item content, minority group representation in the norms, and minority group portrayals in test stimulus materials (Sireci & Geisinger, 1998).

All tests should present items in a sensitive manner for all gender, culture, age, and racial groups. Stimulus artwork should depict people performing similar or equivalent roles and activities, regardless of gender, age, race, and cultural backgrounds. Stimulus artwork that portrays facial expressions, such as happiness, anger, or fear, or indicators of physical limitations like eyeglasses, hearing aids, or wheelchairs, should be evenly distributed across representations of differing demographic groups. Stereotyping of any sort in test artwork and stimulus materials should be avoided.

Differential item function (DIF). Differential item function (DIF) refers to a family of statistical procedures used to identify whether test items display different statistical properties in different group settings after controlling for differences in the abilities of the comparison groups (Angoff, 1993). The concept of DIF has been extended by Shealy and Stout (1993) to include a test level of analysis known as *differential test function* (DTF). DTF is important because tests may produce a small number of offsetting items that are identified as biased against both comparison groups, such as males and females, using DIF procedures. Because the number of biased items are offsetting, the overall effect (DTF) of these few items on the fairness of the test can be minimal (Waller et al., 2000).

Invariant factor structure and scale reliabilities. The examination of comparable reliability and validity across separate demographic groups should be conducted to investigate test fairness. Jensen (1980) noted that if test reliability and validity coefficients differ significantly for designated subgroups of interest, then "it is clear that the test scores are not equally [reliable or valid] measures for both groups" (p. 430). With respect to validity, Meredith (1993) asserted that strict *factorial invariance* is required for test fairness and equity to exist.

Geisinger (1998) noted the importance of comparable reliabilities across subsamples, stating that "subgroup-specific reliability analysis may be especially appropriate when the reliability of a test has been justified on the basis of internal consistency reliability procedures (e.g., coefficient alpha)" (p. 25). The demonstration of comparable reliabilities across samples that differ on the basis of gender, race, or ethnicity has been studied in some current-generation intelligence tests with positive outcomes (Bracken & McCallum, 1998; Matazow et al., 1991; Vance & Gaynor, 1976; Zhu et al., 1999).

External Evidence of Test Fairness

The external features of test fairness are evident in the relationship between test scores and various external criteria, including equality of prediction and consequential impact. It is important to examine external evidence of validity in addition to internal sources of evidence like DIF when investigating test fairness. Focusing solely on internal evidence of fairness may fail to capture subtle yet important sources of test bias (Shepard et al., 1981).

Comparable prediction. The demonstration of equivalent predictive validity across demographic groups constitutes an important source of fairness that is related to validity generalization. Intelligence tests used for the diagnosis of mental retardation should predict future external outcomes, such as employability or independent functioning, in a comparable manner across differing demographic groups.

Minimize adverse impact and selection bias outcomes. A second form of external bias includes the differential incidence of adverse outcomes or differential selection rates across groups. Mean score differences between groups on tests are not inherently an indication of bias and may yield comparable prediction rates. Still, disparate group mean scores can have the undesirable effect of producing disproportionate negative impact for one group as opposed to another (Thorndike, 1971). Such consequential aspects of test bias are commonly referred to as selection bias (Jencks, 1998). When test scores produce adverse, disparate, or disproportionate impact for one group over another, even when that impact is construct relevant, test users should consider the societal and legal implications of such selection bias.

CONCLUSIONS AND RECOMMENDATIONS

Review of the extensive literature on the assessment of intellectual functioning reveals that because of differential rates of development across the life span, the most accurate estimates of intellectual functioning can be made only from recently administered, comprehensive IQ tests. This means that intelligence testing for infants (birth through age 2) is best done at the time of the eligibility determination, within the last year for children between the ages of 3 and 6, and within three years between the ages of 6 and 16. For adults between the ages of 18 and 50 who are living in stable conditions and are in stable health, composite IQ scores are valid for as long as five years; and, after age 50, composite IQs could reasonably be considered valid for three years.

Research also suggests that intelligence in the entire population increases at a rate of approximately 3 IQ points per decade, which approximates the standard error of measurement for most comprehensive intelligence tests. Thus, tests with norms older than 10 to 12 years will tend to produce inflated scores and could result in the denial of services to significant numbers of individuals who would have been eligible for them, if more recent norms had been used.

Because intelligence is a complex and multidimensional construct, it is imperative that intelligence tests used for diagnosis be comprehensive (multifactored) and assess more than a single cognitive attribute. Also, because test length and comprehensiveness are directly related to an instrument's technical adequacy and construct sampling, brief or abbreviated tests compromise test quality or comprehensiveness for brevity.

Language-loaded intelligence tests are not appropriate for people who would be disadvantaged due to language limitations (e.g., deafness, limited English proficiency, elective/selective mute, autism). Whenever language facility constitutes a source of construct-irrelevant variance for examinees, language-loaded instruments (both verbal and performance scales) create an unfair additional challenge. In such cases, examinees should be assessed in their native language or with intelligence tests that do not require receptive or expressive language.

Since the skills and training of the examiner can affect the accuracy of an IQ test, examiners should meet publishers' requirements for the use of Class C tests. Class C instruments are those that require the highest level of training, professional credentials, and supervision. Examiners (not their supervisors) should meet this minimal professional standard. Furthermore, examiners who administer and interpret intelligence tests should possess the skills and competencies to assess clients with uncommon characteristics, such as deafness, extreme youth or age, or a nonmajority cultural or linguistic background. Not only should examiners be competent to administer and interpret intelligence tests, but they should also have the knowledge and experience to work effectively with clients of all ages, exceptionalities, and cultural/linguistic backgrounds to ensure valid assessment results.

Almost a century of intelligence test development has shown that the most valid and accurate results are obtained when tests meet minimal psychometric standards, as outlined in this chapter, for use in high-stakes decision making like SSA disability determination. The tests should demonstrate adequate floors, item gradients, reliability, validity, norm table sensitivity, population representation, as well as sufficient convincing evidence of fairness and lack of bias.

Composite scores from intelligence tests should be used routinely in mental retardation diagnosis, except when the validity of a composite IQ above 70 is in doubt, in which case an appropriate part score may be used in its place. Significant and meaningful variation among an instrument's part scores may indicate evidence of compromised validity for one or more of them (for example, a low verbal scale score for an individual with a suspected speech disorder), which in turn would threaten the validity of the composite IQ. In such situations, appropriate part scores may better represent the individual's true overall level of cognitive functioning or it may be necessary to use other methods to support a diagnosis of mental retardation (see Chapter 5).

However, only part scores derived from scales that demonstrate high *g*-loadings (e.g., crystallized, fluid, visual/spatial measures of intelligence) should be used in place of the composite IQ score when its validity is in doubt. Many intelligence tests assess several facets of intelligence, but not all facets are equally important or predict life events equally well. Those intellectual facets that are heavily *g*-saturated provide the best sources for replacing the composite IQ score when its validity is questionable.

The characteristics of comprehensive IQ tests are such that, even when part scores are used in making disability determinations for mental retardation, the composite IQ score from an instrument should never be higher than 75. Furthermore, if a part score is used in place of the composite IQ score in SSA decision making, the part score should not exceed 70. Therefore:

> **Recommendation: A client must have an intelligence test score that is two or more standard deviations (SD) below the mean (e.g., a score of 70 or below, if the mean = 100 and the standard deviation = 15).**
>
> - *Composite score is 70 or below:* **If the composite or total test score meets this criterion, then the individual has met the intellectual eligibility component.**
> - *Composite score is between 71 and 75:* **If the composite score is suspected to be an invalid indicator of the person's intellectual disability and falls in the range of 71-75, a part score of 70 or**

below can be used to satisfy the intellectual eligibility component.

- *Composite score is 76 or above:* **No individual can be eligible on the intellectual criterion if the composite score is 76 or above, regardless of part scores.**[2]

The committee recommends continuation of the criterion of presumptive eligibility for persons with IQs below 60.

---

[2] Committee member Keith Widaman dissents from this part of the recommendation. Dr. Widaman believes that IQ part scores representing crystallized intelligence (*Gc*, similar to verbal IQ) and fluid intelligence (*Gf*, related to performance IQ) have clear discriminant validity and represent broad, general domains of intellectual functioning. Therefore, a score of 70 or below on either of these part scores from any standardized, individually administered intelligence test that reports such scores should be deemed sufficient to meet the listings for low general intellectual functioning regardless of the level of the composite score, providing that the part scores have adequate psychometric properties (e.g., high reliability, low standard error of measurement). Dr. Widaman notes that, without any clear justification, SSA currently accepts either a composite IQ score from any standardized, individually administered intelligence test or a verbal or performance IQ score, any one of which can be 70 or below. SSA does not stipulate that the composite IQ must be below a certain score for a part score to be used. Dr. Widaman's position provides a rationale for current SSA use of part scores, but it (a) aligns the acceptable part scores with the constructs of *Gc* and *Gf* used in contemporary theories of mental abilities and (b) argues that usable part scores for *Gc* and *Gf* should not be limited to those derived from any particular test instrument.

Chapter 4

The Role of Adaptive Behavior Assessment

NATURE AND DEFINITION OF ADAPTIVE BEHAVIOR

Adaptive behavior has been an integral, although sometimes unstated, part of the long history of mental retardation and its definition. In the 19th century, mental retardation was recognized principally in terms of a number of factors that included awareness and understanding of surroundings, ability to engage in regular economic and social life, dependence on others, the ability to maintain one's basic health and safety, and individual responsibility (Brockley, 1999). Today, fulfillment of these personal and social responsibilities, as well as the per-

---

This chapter contains material drawn from an unpublished paper commissioned by the committee from Sharon Borthwick-Duffy, Ph.D., University of California, Riverside.

formance of many other culturally typical behaviors and roles, constitutes adaptive behavior.

By the close of the 19th century, medical practitioners diagnosing mental retardation relied on subjective or unsystematic summaries of such factors as age, general coordination, number of years behind in school, and physiognomy (Scheerenberger, 1983). These practices persisted over that century because of the absence of standardized assessment procedures. And many individuals who would currently be considered to have mild mental retardation were not included in these early definitions.

Professionals voiced early caution about diagnosing mental retardation solely through the use of intelligence testing, especially in the absence of fuller information about the adaptation of the individual. In addition, mitigating current circumstances (not speaking English) or past history (absence of schooling) were often ignored in the beginning years of intelligence testing (Kerlin, 1887; Wilbur, 1882). At the turn of the century, intelligence assessment placed primary emphasis on moral behavior (which largely comports with the current construct of social competence) and on the pragmatics of basic academics. (Chapter 3 provides details on the development of intelligence assessment.)

Alternative measures to complement intelligence measures began to appear as early as 1916. Edger Doll produced form board speeded performance tests, which were analogues to everyday vocational tasks. During the 1920s, Doll, Kuhlmann, and Porteus sought to develop assessment practices consistent with a definition of mental retardation that emphasized adaptive behavior and social competence. Their work in this area sparked broadened interest in measurement of adaptive behavior among practitioners serving people with mental retardation (Doll, 1927; Kuhlman, 1920; Porteus, 1921; Scheerenberger, 1983).

Doll emerged as a leader in the development of a psychometric measure of adaptive behavior, called social maturity at that time. His work emphasized social inadequacy due to low intelligence that was developmentally arrested as a cardinal indication of mental retardation

(Doll, 1936a, p. 35). Doll objected to the definition of mental retardation in terms of mental age, which had proven problematic in IQ testing (because it resulted in classification of a significant proportion of the population). In 1936, he introduced the Vineland Social Maturity Scale (VSMS—Doll, 1936b), a 117-item instrument. The VSMS, which measured performance of everyday activities, was the primary measure used to assess adaptive behavior, social competence, or social maturity for several decades. One concern that emerged over time was that it was developed and normed for use with children and youth. It did not cover adults and had a limited range of items tapping community living skills (Scheerenberger, 1983).

The assessment of adaptive behavior became a formal part of the diagnostic nomenclature for mental retardation with the publication of the 1959 manual of the American Association of Mental Deficiency (Heber, 1959, distributed in 1961). The 1961 manual (Heber, 1961) discussed adaptive behavior with respect to maturation, learning, and social adjustment. This framework, reiterated in 1983, described adaptive behavior limitations consisting of "significant limitations in an individual's effectiveness in meeting the standards of maturation, learning, personal independence, or social maturity that are expected for his or her age level and cultural group, as determined by clinical assessment and, usually, standardized scales" (Grossman, 1983, p. 11).

The 1983 manual characterized the tasks or activities encompassed by adaptive behavior (and, plausibly social competence) as:

- In infancy and early childhood: sensorimotor development, communication skills, self-help skills, socialization, and interaction with others;
- In childhood and early adolescence: application of basic academic skills in daily life activities, application of appropriate reasoning and judgment in mastery of the environment, and social skills—participation in group activities and interpersonal relations; and
- In adolescence and adult life: vocational and social responsibilities.

During the 1960s, a wider variety of adaptive behavior measures was developed and disseminated (e.g., Allen et al., 1970; Balthazar & English, 1969; Leland et al., 1967). Indeed, by the late 1970s, the number of available adaptive behavior measures, largely interview or observational in format, had burgeoned, including checklists pertaining to vocational behaviors (Walls & Werner, 1977). Measures developed in the 1960s have typically been updated in subsequent editions with enhanced psychometric characteristics and scoring (e.g., Sparrow & Cicchetti, 1985).

Over the past 25 years there has also been further refinement of the parameters and structure of tests of adaptive behavior and social competence. This refinement was based on large samples of research participants and data from service registries (McGrew & Bruininks, 1990; Siperstein & Leffert, 1997; Widaman et al., 1987, 1993). Novel frameworks for conceptualization of adaptive behavior have been proposed (American Association on Mental Retardation, 1992), and conventional frameworks have been endorsed for application in differential diagnosis and classification practices (Jacobson & Mulick, 1996). Finally, the difficulties and complexities of differentiating mild mental retardation from its absence or from other disabling conditions (e.g., Gresham et al., 1995; MacMillan, Gresham, et al., 1996; MacMillan, Siperstein, & Gresham, 1996) have remained an enduring concern in both professional practice and policy formulation.

Differing Conceptualizations

In Chapter 1 we summarized the history of definitions of mental retardation and discussed their relevance to the Social Security Administration's definition. At first glance, current definitions seem to be quite similar; however, there are subtle differences in the conceptualization of adaptive behavior that may affect the outcomes of diagnostic decisions for individuals with mental retardation, particularly those in the mild range.

In the recent *Manual of Diagnosis and Professional Practice in Mental Retardation* (Jacobson & Mulick, 1996), Division 33 of the American Psychological Association put forth a definition of mental retardation that emphasizes significant limitations in intellectual functioning and adaptive behavior. The definition also views adaptive behavior as a multidimensional construct, in that the definition is expanded to include "two or more" factor scores below two or more standard deviations. In describing mild mental retardation, there is minimal reference to adaptive behavior problems, except for the inclusion of "low academic skill attainment."

It is important to note that the Division 33 definition places equal importance on the constructs intellectual functioning and adaptive behavior. The definition speaks to the presence of significant limitations in intellectual functioning and significant limitations in adaptive behavior, which exist concurrently. The term "concurrently" suggests an interdependent relationship in which both constructs are equally important. In this definition, the order of the constructs can be switched without affecting the validity of the definition.

The *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition* (DSM-IV), published by the American Psychiatric Association (1994), definition of mental retardation also has a cutoff of two standard deviations below the mean for intelligence, making an IQ cutoff of 70 to 75 acceptable for a diagnosis of mental retardation. In contrast, there is no mention of a standardized score or cutoff point for operationalizing any "significant limitations in adaptive behavior," even though it is suggested that one or more instruments be used to assess different domains from "one or more reliable independent sources" (p. 40). The implicit rationale for not providing any statistical criteria for adaptive behavior testing is based on the existing limitations in instruments that measure adaptive behavior, specifically in terms of the comprehensiveness of measuring all domains and the reliability of measuring individual domains. Furthermore, issues are raised about the degree to which existing instruments are able to take into account the cultural context in assessing an individual's adaptive behavior. One of the key themes throughout the DSM-IV definition is the cultural aspect of adaptive behavior. For example, adaptive behavior is defined in terms of effectively coping with common life demands and the ability to meet the standards of personal independence for a particular age group with a specific sociocultural background.

The DSM-IV definition identifies four levels of mental retardation based on IQ: mild, moderate, severe, and profound. No mention is made of the degree of severity of adaptive deficits for each of these levels, nor of the number or types of impaired adaptive behavior domains at each level. The DSM-IV definition places a greater emphasis than the Division 33 one on intelligence than on adaptive behavior, defining mental retardation as "significantly sub-average general intellectual functioning *accompanied* by significant limitations in adaptive functioning" (p. 39). In using the term "accompanied," the definition suggests that adaptive behavior is a supplementary variable to intelligence, although both criteria must be present.

The World Health Organization (1996) also includes a definition of mental retardation in its *International Classification of Diseases and Related Health Problems, Tenth Revision* (ICD-10). ICD-10 views the relationship between intellectual functioning and adaptive behavior as causal, with deficits in adaptive behavior resulting from deficits in intellectual functioning.

In describing the different severity levels of mental retardation, the ICD-10 guide presents IQ levels not as strict cutoffs but as "guides" to categorizing individuals with mental retardation. There is no mention of any standardized cutoffs for adaptive ability, except for mention of the use of "scales of social maturity and adaptation" in the measurement of adaptive behavior.

In the characterization of mild mental retardation, the ICD-10 guide points out that, "some degree of mild mental retardation may not represent a problem." It goes on to state that the consequences will only be apparent "if there is also a noticeable emotional and social immaturity." This statement implies that for individuals with mild mental retardation, intellectual deficits are apparent only when represented by problems in adaptive behavior (emotional and social immaturity). Furthermore, "behavioral, emotional, and social difficulties of the mildly mentally retarded . . . are most closely akin to those found in people of normal [range of] intelligence." It is important to note that the terminology used in the ICD-10 is international English rather than North American English, and that, as a result, word usage in ICD-10 is not entirely consistent with contemporary North American terminology with respect to functional limitations or depiction of social performance.

The most cited definition in the field is that of the American Association on Mental Retardation (AAMR). In their most recent classification system (American Association on Mental Retardation, 1992), AAMR defines mental retardation as subaverage intellectual functioning existing concurrently with limitations in adaptive skills. These limitations in adaptive skills are operationally defined as limitations in two or more of ten applicable adaptive skill areas (e.g. self-care, home living, social skills, self-direction, health and safety, etc.). The definition also includes the notion that adaptive skills are affected by the presence of "appropriate supports" and with "appropriate supports over a sustained period, the life functioning of the person with mental retardation will generally improve."

AAMR departs significantly from other organizations by eliminating the grouping of individuals with mental retardation into levels of severity. AAMR no longer differentiates, either qualitatively or quantitatively, differences in intellectual or adaptive functioning of individuals with mild, moderate, severe, and profound mental retardation. Instead, they differentiate individuals with mental retardation based on the supports they need. The result is that the unique aspects and characterization of individuals with mild mental retardation are no longer the basis for differentiating them from more moderately and severely involved individuals. In so doing, AAMR ignores the substantial theoretical and empirical foundation that validates the difference between individuals with mild mental retardation and other individuals with mental retardation (MacMillan et al., in press).

Among these four definitions, there is little variation in the intelligence construct for individuals with mental retardation. The differences occur rather in their consideration of the contributing role of adaptive behavior. In some definitions (Division 33 and AAMR), adaptive behavior is construed as distinct from intellectual functioning and of equal importance, while in other definitions it is considered a result of deficits in intellectual functioning. The definitions also vary as to whether they consider adaptive behavior to be made up of a single factor or to have multiple factors or domains. In the definitions that imply a multifactor construct, deficits in adaptive behavior must be specified in a certain number of areas/domains. With regard to identifying decision-making criteria, Division 33 presents the only definition that employs a statistical cutoff based on standard norms. In contrast, the other definitions employ more qualitative terms, which are open to interpretation in describing deficits and limitations in adaptive behavior.

Dimensions of Adaptive Behavior

Structure

Multidimensional or Unidimensional? Answers to this question have been mixed. Meyers et al. (1979) concluded from their review of factor analytic studies that adaptive behavior was definitely multidimensional and that the use of a total score would be inappropriate to indicate a general level of adaptation. Their view has been both supported and disputed in the past two decades, and there are currently firm adherents on each side of this issue. McGrew and Bruininks (1989) and Thompson et al. (1999) have concluded, for example, that the number of factors emerging from factor analyses depends on whether data were analyzed at the item, parcel, or subscale level, with fewer factors found for subscale-level data than item- or parcel-level data.

They also found that it was not the selection of the instrument that determined the number of factors**.** This important finding has direct implications for definitions that require limitations to be observed in a specific number of areas. If there is actually one underlying domain that "causes" behaviors in all different conceptual domains, and there is relatively little unique variance found in each domain, then a total score with a single cutoff point could reliably distinguish those with and without significant limitations. If not, diagnosticians would have to consider a profile of adaptive behavior deficits that takes all domain scores into account. Widaman et al. (1991) and Widaman and

McGrew (1996) concluded that evidence supported a hierarchical model with four distinct domains: (1) motor or physical competence; (2) independent living skills, daily living skills, or practical intelligence; (3) cognitive competence, communication, or conceptual intelligence; and (4) social competence or social intelligence. Widaman and McGrew (1996) further argued that agreement on a common set of terms for domains of adaptive behavior (in contrast to the use of "or" as above) would contribute to a better consensus on the structure of adaptive behavior.

The review by Thompson et al. (1999) is the most recent summary of studies using factor analysis; it concludes that adaptive behavior is a multidimensional construct. The three most common dimensions found were in these broad categories: (1) personal independence, (2) responsibility, i.e., meeting expectations of others or getting along with others in social contexts, and (3) cognitive/academic. Physical/developmental and vocational/community dimensions were found less often. Thompson et al. concluded: "No single adaptive-maladaptive behavior assessment instrument completely measures the entire range of adaptive and maladaptive behavior dimensions. . . It is clear that different scales place different levels of emphasis on different adaptive behavior domains. No one instrument produced a factor structure that included all of the domains" that were identified by the American Association on Mental Retardation (1992).

Breadth of Domains. The domains assessed by adaptive behavior scales, and thus the individual items included on them, depend in part on the context, target age group, and purpose of the measure. Thus, considerable variation has been found in the content covered by different scales (Holman & Bruininks, 1985; Thompson et al., 1999). Measures used in schools may not need a work domain, for example, if students are too young for employment or the school does not have a work experience program. Conversely, adult scales would not need items on school-related behaviors (Kamphaus, 1987a). In their review, Thompson et al. (1999) suggest that this incongruity reflects the problem noted by Clausen (1972) and Zigler et al. (1984), that adaptive behavior lacks a unifying theoretical foundation. A consequence of this, according to Thompson et al., is the inability to develop precise measures of adaptive behavior that would objectively differentiate individuals by disability. An alternative explanation is that adaptive behavior must be understood in the context of the individual's relevant daily and social life, which is determined by age, culture, and context (Thompson et al., 1999).

Independence of Domains. The 1992 AAMR definition requires that an individual show significant limitations in at least 2 of the 10 adaptive skill areas. A danger of accepting "erroneous domains that are not truly distinct from one another" (Thompson et al., 1999, p. 17) is that it can lead to the inconsistent application of eligibility criteria and unequal treatment across groups of people. Thus, characteristics of the factor structure of a measure of adaptive behavior have important implications for diagnosis.

Thompson et al. (1999) reviewed studies that reported factor analyses of adaptive behavior measures. They made two important points before summarizing their findings: (1) highly correlated factors may indicate that they do not represent independent dimensions and (2) different methods of factor analysis can support different factor structures.

Domains Missing from Adaptive Behavior Scales

Greenspan (1999) noted that a drawback to the factor analytic approach to determining the dimensional structure of adaptive behavior is that this statistical method cannot determine whether some domains do not make conceptual sense (i.e., items should not have been included on tests in the first place) or whether missing content domains should have been included.

Social Skills Dimension of Social Competence. Most adaptive behavior scales contain factors addressing interpersonal relationships or social skills, but they do not address overall social competence. For individuals whose diagnosis is most in question because their measured IQs are near the cutoff, this vital area may determine the presence or absence of mental retardation. Gresham and Elliott (1987) and Greenspan (1999) have argued that social competence has received too little attention in the conceptualization and measurement of adaptive behavior (Figure 4-1). Their model divides social competence into two overall dimensions: (1) adaptive behavior, which includes the factors contained on most adaptive behavior scales (independent functioning, self-direction, personal responsibility, vocational activity, functional academic skills, physical development) and (2) social skills, including domains that are likely to be most key to identifying mental retardation at the borderline levels (interpersonal behaviors, self-related behaviors, academic-related skills, assertion, peer acceptance, communication skills). The dimensions of adaptive behavior and social skills in the Gresham and Elliott model are surprisingly similar to the 10 adaptive skill areas in the 1992 AAMR definition of mental retardation.

Gullibility/Credulity Component of Social Competence. Greenspan and colleagues (Greenspan, 1999; Greenspan & Driscoll, 1997; Greenspan & Granfield, 1992) have argued that social intelligence, some aspects of which are not contained on any current scales of adaptive behavior or social skills (e.g., credulity, gullibility), should be a key determinant of a diagnosis of mental retardation for adults (Figure 4-2). Greenspan and Driscoll (1997) proposed a "dual nature of competence." They suggest that intelligence, as measured by IQ, is typically viewed as an independent variable that predicts outcomes, whereas personal competence is the combination of what individuals "bring to various goals and challenges as well as their relative degree of success in meeting those goals and challenges" (p. 130).

Greenspan (1999) argues that the victimization of people with mental retardation, observed in social and economic exploitation, is "a more central (and generally more subtle) problem that goes to the heart of why people with mental retardation are considered to need the protections (ranging from in-home services to conservators) associated

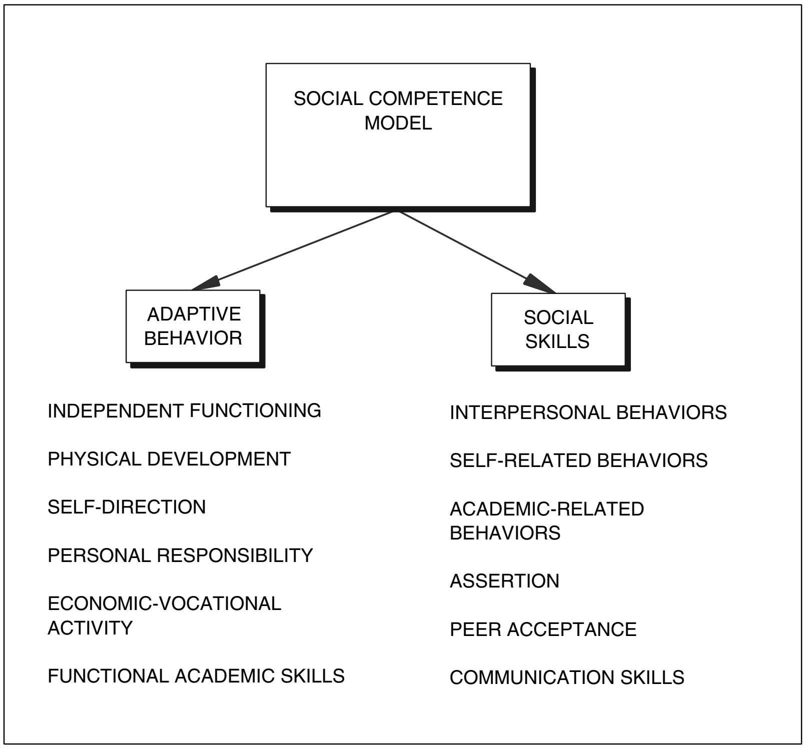

FIGURE 4-1 Social competence model. SOURCE: Gresham & Elliott (1987). Copyright 1987 by PRO-ED, Inc. Reprinted with permission.

with the label" (p. 69). Very recently Greenspan (1999) proposed ideas for assessing vulnerability in a comprehensive assessment of adaptive behavior or social competence. As there is no research yet on credulity in people with mental retardation, these proposals for assessment are unlikely to be found in practice in the next several years. Nevertheless, there is merit to the idea of considering these subtle indicators of social competence, i.e., vulnerability, gullibility, and credulity, as important indicators of adaptive behavior in people with mild cognitive impairments.

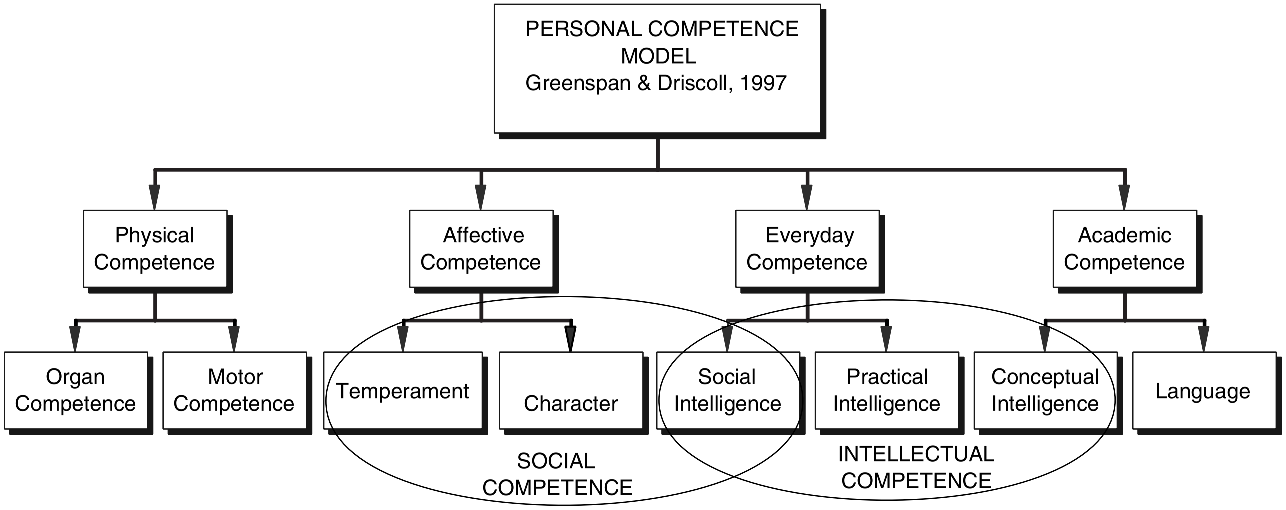

FIGURE 4-2 Personal competence model. SOURCE: Greenspan and Driscoll (1997). Reprinted with permission.

Maladaptive Behavior

Many adaptive behavior scales contain assessments of problem or maladaptive behavior, but relationships between domains of adaptive and maladaptive behavior are generally low, with correlations tending to be below .25 (and a tendency to be higher in samples of persons with severe or profound retardation—Harrison, 1987). Division 33 makes it clear that the presence of clinically significant maladaptive behavior *does not* meet the criterion of significant limitations in adaptive functioning (Jacobson & Mulick, 1996). Hill (1999) also emphasized that behaviors that interfere with a person's daily activities, or with the activities of those around him or her, should be considered maladaptive behavior, not the lack of adaptive behavior. Refusal to perform a task that a person is capable of doing is also a reflection of problem behavior and should not be considered in relation to adaptive behavior. The classroom form of the Vineland Adaptive Behavior Scales (Sparrow & Cicchetti, 1985) does not include a section on maladaptive behavior, which also suggests that these authors viewed measures of problem behavior as irrelevant to diagnosis or eligibility. Greenspan (1999) also has argued for many years that the presence of maladaptive behavior, or mental illness, is irrelevant for the purpose of diagnosing of mental retardation.

If it is assumed that maladaptive behavior ratings should not contribute to diagnostic decisions about adaptive functioning, then problems in their measurement need not affect this process. However, because several adaptive behavior scales contain maladaptive components, it is worth noting important challenges to reliable measurement. Specifically, several roadblocks to meaningful ratings of maladaptive behavior were noted after publication of the original AAMD Adaptive Behavior Scales (ABS). Scales developed subsequently improved on the simple rating format found in the ABS, which contained a finite list of problem behaviors rated according to the frequency of occurrence. These improvements notwithstanding, the complexity of balancing frequency and severity of problem behavior occurrence will continue to pose problems of score interpretation.

ASSESSMENT OF ADAPTIVE BEHAVIOR

Assessment Dimensions

The assessment of adaptive behavior is complex. One must consider not only general competencies across relevant domains but also the level, quality, and fluency of those behaviors. In addition, there is the issue of the ability to perform behaviors (i.e., can do) versus the actual performance of those skills (i.e., does do). In order for the assessment to be clinically and scientifically meaningful, it is important that the assessor be sufficiently trained in using and interpreting appropriate instruments. A high level of training is necessary in order to capture and distinguish the level, quality, and pattern of adaptive behavior displayed by a given subject, as viewed by the eyes of the respondent (parent, teacher, or caregiver).

The frequency of performance can be classified along a dimension from "never" to "usually or always." The number of choice points varies by specific instrument or by the variation in the clinical interpretation of the assessor when a formal assessment instrument is not used. The quality of performance may be somewhat more subjective, but a key feature is the appropriateness of a given level of adaptive behavior performance. For example, one needs to distinguish between an individual's deficit in a specific adaptive behavior skill, as opposed to a deficit in a larger domain.

Assessment Methods

There are a number of ways to assess the level, quality, and pattern of adaptive functioning, each with its own strengths and weaknesses. These include clinical assessment by interview methods (unstructured, structured, semistructured, direct observation), usually with the aid of clinical instruments that are completed by the evaluator during the interview, and the use of checklists that are completed either by an observer or by the individual being assessed.

In an unstructured interview, the clinician applies personal, experience-based clinical norms to the adaptive behavior assessment. The advantage of the method is that it frees the clinician from using a set of criteria that may be perceived as restrictive. The disadvantage is that each clinician imposes his or her own subjective criteria, a process that threatens both the reliability and the validity of the assessment.

Both structured and semistructured interviews, when performed by well-trained and experienced clinicians, appear to be the best available safeguard against threats to the reliability and the validity of adaptive behavior assessment. These procedures, however, need to be employed using an instrument that is reliable, has valid criteria for evaluating adaptive behavior, and uses empirically based norms. In fact, semistructured interviews require the highest level of professional expertise, as the questioning and interpretation of answers requires a high level of training.

Since the adaptive behaviors that need to be assessed are those found in the context of a broad range of everyday living situations displayed across a wide variety of settings, an assessment of adaptive functioning by direct observation is usually not practical. It would be difficult to set up situations in which individuals can demonstrate their ability to perform a wide variety of social, communicative, and daily living behaviors.

Checklists completed by teachers, parents, or other caregivers are often used to rate individuals' behavior for a broad variety of suspected conditions (e.g., mental retardation, autism, other pervasive developmental disorders, attention deficit disorder). However, the simplicity and lack of reliability or validity of many such procedures render them less useful than more complex measures administered professionally. Checklists may add valuable information and insights, but they are seldom solely sufficient for diagnostic purposes. In order to make reliable and valid judgments about the presence or absence of many behaviors, the items may need such extensive clarification as to obscure the meaning of such behaviors for many respondents.

The issues of cross-cultural, racial, ethnic, and subcultural biases are of concern to some who view many aspects of adaptive functioning as culturally determined (Boyle et al., 1996; Valdivia, 1999—for a general discussion see the section "Sociocultural Biases"). The issue of sociocultural bias also arises in the context of the adaptive behavior interview. Administration of adaptive behavior scales generally follows one of two possible formats. One is an interview with a professionally trained interviewer and a respondent who knows the individual being assessed well. The other consists of a person who also knows the individual being assessed well but who independently completes a checklist of specific items without assistance. Other scales permit someone to help the person answer questions that cannot be answered without assistance. Some scales can be administered either way. When trained professionals use an interview format, the phrasing of items contained in the record booklet is not used. In this format, the professional has the opportunity to ask questions that are at the appropriate level of sophistication and also appropriate to the cultural group of the respondent.

Adaptive behavior is generally not a mental health issue, since the focus is on developing positive behaviors, rather than deficits. Thus, some of the concerns about cultures that are less accepting of mental illness labels than the majority culture are much less relevant to adaptive behavior assessment.

There seems to be little evidence that adaptive behavior assessment is as prone to cultural, racial, and ethnic bias as other areas of psychological testing. For example, adaptive behavior tests are not as culturally or ethnically bound as tests of intelligence (Hart, 2000; Hart & Risley, 1992; Sparrow et al., 1984a; Walker et al., 1994). However, a recent surgeon general's report (U.S. Department of Health and Human Services, 2001) focuses on the miscommunication that may exist when the interviewer and the respondent speak different languages. The report notes that "several studies have found that bilingual patients are evaluated differently when interviewed in English as opposed to Spanish." It is also possible that different subcultural expectations about independence or religious or medical causes for certain behaviors may affect the validity of reports. In such instances, if a same-language or same-culture interviewer is not available, the clinician needs to be very aware of such possible miscommunications in order to obtain a valid interview. The surgeon general's report emphasizes that more research is needed to better understand how, when, and if culture affects interview-based assessments.

Psychometric Concerns in Using Adaptive Behavior Scales

The primary use of adaptive behavior scales in the classification of mental retardation has frequently been confirmatory (i.e., to confirm that a low IQ is associated with delayed acquisition or manifestation of everyday personal and social competencies). This use may result from concerns among clinicians about the robustness of adaptive behavior measures. For the most part, such concerns result from considerations of the structure of measures (e.g., as related to items and other factors mentioned in this section), procedures for obtaining information used to complete the protocols, and issues surrounding informant bias.

Such concerns arise in part because intellectual performance, the other criterion associated with mental retardation, is measured by comprehensive intelligence tests that are the most thoroughly researched forms of psychological assessment (Neisser et al., 1996). Research studies in the past decade that employ adaptive behavior measures have used them as outcome measures or to study the structure or dimensions of adaptive behavior, rather than behavioral development. Clinicians may consequently believe adaptive behavior to be less well understood than intelligence. Nonetheless, there is a rich literature documenting differential outcomes for quality of life, autonomy, and clinical decision making for adaptive behavioral development as measured by existing assessment instruments (Jacobson & Mulick, 1996). Newer adaptive behavior scales evidence more robust psychometric properties than older scales. In this section, we discuss a variety of psychometric features of adaptive behavior scales that have implications for decision making about mental retardation.

Floor and Ceiling Effects

The initial, and probably primary, application of adaptive behavior scales in clinical practice has been to assess the behavioral development of children thought to have mental retardation. Thus, most norming samples, item development, and scale selection have been targeted at groups ages 3 to 18 or 21. This facilitates the early identification of preschool children at risk of mental retardation and permits confirmation of persisting developmental delays. Adult norming samples are often included as well, but they tend to consist of people with already identified disabilities. Thus, adaptive behavior scales have particular relevance in application with preschoolers and with teens, who are often participants in Supplemental Security Income (SSI) determinations or redeterminations. However, depending on the age range of adult participants without disabilities sampled during norming studies, the ceiling (i.e., the highest level of behavioral performance assessed) may differ across scales and may affect the characterization of the degree of delay manifested. Measures of behavioral functioning or responsiveness of children younger than 36 months have not been strengths of many adaptive behavior measures. Infants and toddlers may more appropriately be assessed with more specialized measures in most cases.

Developmental Range Effects

Floor and ceiling effects are also evident as developmental range effects. Scales typically include items that permit behavioral assessments for young children and adolescents without disabilities (i.e., superior behavioral development or skill). For older adolescents, ages 18 to 21, the difficulty level of items often permits identification of either delayed or typical skills. Thus, to the extent that a young adult with mild mental retardation has selected skills that are well developed relative to others, it may not be accurate to describe those skills in developmental terms. Instead, it may be possible to establish only that their skills are superior to those achieved by other young adults with mild mental retardation, and they may sometimes fall in the normal range of performance of similar age peers. Some data suggest that ceiling and developmental range effects hinder the full description of skill assets for some individuals with mild mental retardation. In unpublished data on some 27,000 people with mild mental retardation, between 75 and 100 percent of participants obtained perfect scores (100 percent) on three of five indices of one scale (J.W. Jacobson & C.S. Brown, personal correspondence, June 17, 2001).

Item Sampling in Relation to Age–Typical Behavior and Settings

Because adaptive behavior scales are designed with applicability for a wide age range but with primary emphasis on childhood and adolescence, some items may not be suitably worded or may not reflect a performance that is age-relevant. For example, an item may tap skills associated only with childhood (e.g., performing a specific activity or completing a task with adult assistance in an age-typical manner) or with adulthood (e.g., menstrual care for an adult or adolescent woman). Some scales contain provisions for alternative items or alternative performance of items. However, depending on the nature of these provisions, they may reduce the comparability of measures of the related skills from different adaptive behavior scales.

In other instances, scales may be constructed such that they are relevant to only certain age groups (e.g., the motor scale in the Vineland ABS), or different versions of the same scale may be used in different settings (e.g., school versus residential and community settings). For example, the Adaptive Behavior Assessment System (Harrison & Oakland, 2000a) is available in four forms: parent, teacher, adult self-report, and adult reported by others. The two versions of the AAMR ABS differ with respect to the age groups emphasized and the settings about which items are structured and weighted in item selection. For example, in relation to the AAMR school-age scales, "items were selected in part based on discrimination among institutionalized individuals and community dwelling individuals previously classified at different adaptive behavior levels, and among adaptive behavior levels in public school populations" (Lambert et al., 1993b).

Item Density

Adaptive behavior scales are structured to be comprehensive without being cumbersome (Adams, 2000). Consequently, several features must be balanced. A number of factors and descriptive categories of behavioral development must be represented adequately in order to ensure comprehensiveness and documentation of both strengths and limitations for clinical and diagnostic purposes. The number of items associated with each descriptive category must be sufficient to provide a scale and to be applicable across age ranges. A relatively wide age range must be represented. In balancing these factors, item density, that is, the inclusion of multiple items reflecting age-typical performance at a range of ages, must be maintained at a fairly uniform level. This means that within any one subscale of an adaptive behavior scale, for example, there may be only one or two items typical of performance for a 10-year-old. When subscale scores are aggregated into summary scores, this results in a meaningful number of age-relevant items, although the items sampled in each subscale are limited. For this reason, some manuals recommend that clinicians fully explore the nature of tasks that the focal person performs that may be age typical (e.g., Sparrow et al., 1984a). Nonetheless, it should always be recognized that items in adaptive behavior measures represent a *sampling* of items that have passed reliability and validity screens, rather than a complete characterization of adaptive behavior.

Reliability of Informant Judgments

Because adaptive behavior scales are typically completed through interview of informants or direct responses (marking of a protocol by the informant), the reliability and the validity of informant responses have been particular concerns. These concerns are heightened when informants have a stake in the outcome of the assessment (e.g., when responses may affect eligibility for services). Developers have addressed this issue through several strategies: (1) assessing the interrater and test-retest reliabilities of measures, (2) providing instructions to raters for coding items (e.g., Sparrow et al., 1984a), and (3) specifying training for clinicians and preparation of raters (e.g., Bruininks et al., 1996). Reliabilities are initially assessed at the item level and then at the scale and factor levels. Current measures evidence acceptable interrater and test-retest reliability, with consistency scores at levels of .90 and above (seldom at a level below .80) for clinical and normative subgroups, partitioned by age and clinical variables. Similarly, adequate internal consistency of subscales or domains is documented using split-half or alpha coefficients. Full details on standardization and reliabilities are provided in the manuals associated with the major adaptive behavior scales (Adams, 2000; Bruininks et al., 1996; Harrison & Oakland, 2000b; Lambert et al., 1993b; Sparrow et al., 1984b; see also Harrington, 1985). Additional discussion is provided in Chapter 3.

Validity of Informant Judgments

Validity can be categorized in terms of: (1) content validity (evidence of content relevance, representativeness, and technical quality); (2) substantive validity (theoretical rationale); (3) structural validity (the fidelity of the scoring structure); (4) generalization validity (generalization to the population and across populations); (5) external validity (applications to multitrait-multimethod comparison); and (6) consequential validity (bias, fairness, and justice; the social consequence of the assessment to the society—Messick, 1995). Technical manuals present analyses of data gathered in the process of test development that addresses content validity (in terms of representativeness and inferences from age norms), substantive validity (in that they present either a theoretical or empirically derived model of adaptive behavior to which the scale conforms), generalization validity (with respect to differing age or disability groups), external validity (in terms of concurrence with previous or contemporary adaptive behavior measures and intellectual measures), and consequential validity (in terms of evidence of bias or procedures utilized to reduce bias). As previously noted, primary concerns in the use of adaptive behavior scales in eligibility determination decisions center on informant bias.

Manuals for the major adaptive behavior scales encourage the use of multiple informants, for example, teachers and parents. This allows the rater to obtain a complete picture of the adaptive functioning of the person being assessed. It also allows for reconciliation of ratings among these informants. Both legislative action and judicial decisions at the federal level have focused on concerns that parents may misinform clinicians regarding their children's skills in order to obtain SSI benefits. Federal review of the SSI program has indicated that such deception is an uncommon occurrence.

Adequacy of Normative Samples

Another psychometric concern is whether the norming samples are adequate. Although normed on smaller samples than comprehensive intelligence tests use, current adaptive behavior measures typically have adequate norming samples in relation to both representation of people with and without mental retardation and representation of age groups in the population in relation to the age span of the measure.

- For the Adaptive Behavior Assessment System (Harrison & Oakland, 2000a), the norming groups for 5- to 21-year-olds included 1,670 (parent form) to 1,690 (teacher form) children; for 16- to 89-year-olds, the norming groups included 920 (rated by others) to 990 (self-report) adults without disabilities throughout the United States.
- For the Scales of Independent Behavior-R (Bruininks et al., 1996), the norming sample included 2,182 people ages 3 years 11 months to 90 years, with a sampling frame based on the general population of the United States stratified for gender, race, Hispanic origin, occupational status, occupational level, geographic region, and community size.
- For the AAMR Adaptive Behavior Scale-School scales (Lam-

bert et al., 1993a) the norming group included 2,074 students (ages 3-21) with mental retardation living in 40 states, and a sample of 1,254 students (ages 3-18) without mental retardation from 44 states.

- For the Vineland Adaptive Behavior Scales (Sparrow et al., 1984a) the standardization sample was representative of the U.S. population. It consisted of 3,000 children ages birth through 18 years 11 months of age, including 99 children in special or gifted education among the 2,500 who were of school age.
- For the Comprehensive Test of Adaptive Behavior-R (Adams, 2000), the norming sample represented four of five U.S. regions (excluding the West) and included a nonschool sample of 4,456 people with mental retardation ages 10 to 60+ years and a school sample of 2,094 children and adolescents with mental retardation ages 5 to 22, and a sample of 4,525 children and adolescents without mental retardation ages 5 to 22.

Sociocultural Biases

Bias refers to a consistent distortion of scores that is attributed to demographic factors, principally nonmodifiable personal characteristics such as age, gender, race, and ethnic or cultural membership. In the United States there have been significant concerns about the relationship between ethnicity or racial origin and performance on intelligence tests (Neisser et al., 1996). This has generalized to adaptive behavior measures. As the importance of adaptive behavior measures in classification of mental retardation has increased, this concern has been heightened as disproportionate numbers of minority children have been identified as having mental retardation, primarily because of low-income status and the overrepresentation of individuals with mental retardation among low-income people (Boyle et al., 1996).

Some (e.g.,Valdivia, 1999) have suggested that age norms are arbitrary and reflect white, middle-class childrearing standards, and that developmental attainments are affected by culturally different skills and expectations. The result is an overidentification of skill limitations among minority children. However, research that indicates a causal relationship between the childrearing practices of minority families in North America and developmental delay is very limited. Comparative research examining the relationship between minority status and pronounced delays that are not accounted for by socioeconomic factors is also limited. However, available data are sufficient to raise concerns that such issues should be studied further (Bryant et al., 1999; Craig & Tasse, 1999).

To some extent, inclusion of participants representative of the general population, including racial and ethnic minorities, in norming samples should mitigate against biases in scoring of adaptive behavior scales. To the extent that low income or very low income is more common among certain ethnic minority groups, however, differences in developmental trajectories for children may reflect differences in childrearing practices and stimulation that are associated with economic and social class and related levels of parental education (Hart, 2000; Hart & Risley, 1992; Walker et al., 1994).

Although research from the 1970s and 1980s found comparable performance on adaptive behavior scales among majority and minority ethnic groups (Bryant et al., 1999; Craig & Tasse, 1999), linguistic factors remain a concern. These include such considerations as interviewing informants in their primary language and dialect, and the comparability of translations of items in adaptive behavior scales to particular languages and dialects, including dialects in English (e.g., American and British). Translation is a concern because the comparability of translations of items has seldom been confirmed through back-translation from the translated content to the initial language, or through confirmatory analysis through further retranslation (Craig & Tasse, 1999). Noncomparability of items may alter norms due to item wording that requires a higher developmental level of performance in the translated item. Also, English language norms may be lower than the typical performance of a same-age child in another culture. Cross-cultural and cultural subgroup studies of adaptive behavior differences among ethnic, racial, or national groups are certainly needed, but evidence for substantial relationships between racial or ethnic group membership and performance on adaptive behavior scales, unmediated by socioeconomic differences, is very limited.

Nonetheless, culturally competent assessment practices require consideration of the developmental impacts of cultural practices or language differences among examiners, examinees, and informants that may affect the validity of the clinical information collected and interpreted. Under ideal circumstances, adaptive behavior measures should be administered in an examinee's or informant's primary language. Often, there may be no substitute for assistance by a translator familiar with the informant's dialect, even for examiners who are fluent in the informant's primary language. In instances in which the informant is bilingual, it may be appropriate to probe interview responses in both languages.

Adaptive Behavior Scales with Well-Known Properties

There are at least 200 published adaptive behavior instruments that have been used for diagnosis, research, program evaluation, administration, and individualized programming. Some of these scales were developed to serve only one of these purposes; however, several have attempted to include both the breadth required for diagnosis and the depth required for clinical use. Most tests fall short of accomplishing both purposes. Referring to the dual purpose of adaptive behavior scales, Spreat (1999) concluded that it is "unrealistic to think that the same test can be used for program evaluation, diagnosis, classification, and individual programming" (p. 106). Among the very large number of adaptive behavior scales on the market, very few have adequate norms and reliability to diagnose mental retardation in people with IQs in the questionable range (e.g., 60-80). Kamphaus (1987b) reported that the Vineland Adaptive Behavior Scales-Survey Form (Sparrow et al., 1984a) and the Scales of Independent Behavior (Bruininks et al., 1984) had adequate standardization samples. In a national survey of school psychologists, only three scales were found to be in wide use for diagnosis: the Vineland Adaptive Behavior Scales, the Scales of

Independent Behavior, and the AAMR Adaptive Behavior Scale-School Edition (Stinnett et al., 1994). The Adaptive Behavior Assessment System (Harrison & Oakland, 2000a) is quite new and relatively untested, but its psychometric properties and norms extend to age 89.

Each of these scales (except the Adaptive Behavior Assessment System) has been reviewed extensively and compared with others in detailed reports. Readers are referred to the test manuals and to Reschly (1990), Harrison and Robinson (1995), Thompson et al. (1999), Jacobson and Mulick (1996), Spector (1999), Hill (1999), *Test Critiques,* test reviews in the *Journal of Psychoeducational Assessment,* and the *Mental Measurements Yearbooks* for more detailed psychometric information about these and other measures. Although each scale described has both strengths and weaknesses, each has impressive psychometric characteristics and is highly recommended for use in eligibility determination and diagnosis. Decisions about which instrument to use depend on the age of the individual to be tested and available norms, available sources of information, the context in which the individual is known, and the training of the rater.

Vineland Adaptive Behavior Scales

The Vineland Adaptive Behavior Scales (VABS—Sparrow et al., 1984a) have their conceptual roots in the Vineland Social Maturity Scale (Doll, 1936b), although overlap between the original and the new scales is minimal (Kamphaus, 1987b). There are actually three scales, including a survey form (VABS-S) and an expanded form (VABS-E), which uses a conversation data gathering format during interviews with parents or guardians. A psychologist, social worker, or other professional who has appropriate training in interview techniques must complete these forms. Norms on children having no disability are available from birth to 18 years, 11 months, based on a standardization sample of 3,000 cases that were stratified by age, gender, ethnicity, parental education, geographic region, and community size consistent with U.S. census data. The proportion of children from homes with low socioeconomic status was lower than that in the census data. The expanded version is designed to meet the requirements of diagnosis and of planning/intervention, and is intentionally longer and more detailed in order to ascertain information on specific skill deficiencies. Data from reliability and validity studies of the survey form are very impressive, especially in light of the flexible conversational procedures used for obtaining information.

The third scale is a classroom form (VABS-C), appropriate for children ages 3-12, and can be completed by the teacher fairly quickly. It does not require specific or graduate training to complete. However, teachers have limited opportunities to observe all behaviors on the VABS-C and must necessarily provide estimates of behaviors that do not occur in the school context. A strength of this scale is that teachers are asked to record when they estimate behaviors, so the resulting threat to reliability and validity can be appraised.

AAMR Adaptive Behavior Scales

There are two versions of the Adaptive Behavior Scales (ABS)–a school version (ABS-S:2—Lambert et al., 1993a) and a residential and community version (ABS-Residential and Community, ABS-RC:2—Nihira et al., 1993). The ABS-S:2 is used to identify students who are significantly below their peers in adaptive functioning for diagnostic purposes. It also determines strengths and weaknesses, documents progress, and assesses the effects of intervention programs. Although it is linked to AAMR by name, the ABS does not provide subscale scores in the 10 adaptive skill areas listed in the 1992 AAMR definition of mental retardation. Stinnett (1997) matched ABS items to the 10 adaptive skill areas in the AAMR definition and found that some skill areas are addressed in depth by the ABS (social skills and self-care domains), while others have too few items to give reliable estimates (home living, health and safety, leisure). Nine behavior domains measure personal independence and personal responsibility in daily living, including prevocational/vocational activity. A second part of the ABS-S relates to social and maladaptive behavior.

The ABS-S was standardized on population samples of people with and without mental retardation. Standard scores, age-equivalent scores, and percentile rank scores can be converted from raw scores on the adaptive behavior subscales and three factor scores for ages 3-21. The standardization samples have been judged to be excellent, although the fact that the sample of people with mental retardation did not include people in the IQ range 71-75 is likely to overestimate adaptive behavior when using the mental retardation norms (Stinnett, 1997). Since other norms should be used for determining a diagnosis of mental retardation, according to the manual, this should not be a problem in the current SSA context. The ABS-S:2 has excellent interrater reliability.

The ABS-S:2 provides norms only through age 21 and includes some content specifically appropriate for school settings rather than adult environments. The residential and community version, ABS-RC:2, was developed to be appropriate for use with persons through 79 years of age. ABS-RC:2 norms are not available for adults with typical functioning, and most norm-referenced scores provide comparisons only with adults with developmental disabilities. The standardization sample consisted mostly (80 percent) of adults living in residential facilities, and the overall functioning level of the sample may be lower than if other community-dwelling adults had been included (Harrison, 1998). Because standard scores and percentile ranks do not indicate standing relative to people without developmental disabilities, and because the norming sample is probably not representative of the population of adults with developmental disabilities, the ABS-RC:2 may not fit the psychometric criteria used in determining a diagnosis of mental retardation according to AAMR requirements (American Association on Mental Retardation, 1992).

Scales of Independent Behavior

The Scales of Independent Behavior (SIB-R—Bruininks et al., 1984) is a component of the Woodcock-Johnson Psycho-Educational Battery. The SIB provides norms from infancy to adulthood (40+ years), contains 14 adaptive behavior subscales that fall into four major clusters, and provides an additional full-scale broad independence score.

The SIB-R manual addresses many of the issues that make the scoring interpretation of adaptive behavior scores challenging, including physical disability, the use of adaptive equipment, alternative communication methods, tasks no longer age appropriate, partial performance of multipart tasks, lack of opportunity due to environment or safety, and cognitive ability to understand social expectations for performing behaviors. In general, individuals are to be rated according to what they actually do (or would do if age appropriate), rather than giving "credit" for these considerations or denying credit if tasks are performed well with the assistance of adaptive equipment, medication, or special technology (Hill, 1999). However, if functional independence is to be considered "within the context of the environments and social expectations that affect his or her functioning" (Hill, 1999), interpreting scores without considering opportunity and societal expectations for a person with physical limitations could be problematic for a diagnosis of mental retardation.

Adaptive Behavior Assessment System

The Adaptive Behavior Assessment System (ABAS—Harrison & Oakland, 2000a) is the newest of the adaptive behavior measures that has sound psychometric properties. Although it had extensive field testing before publication, formal reviews are not yet available. It was developed to be consistent with the 10 AAMR adaptive skill domains, and, depending on the weight placed on using the AAMR definition for diagnosis by a clinician, this may be a relevant characteristic. Moreover, the ABAS is appropriate for use with children (age 5 and older) as well as adults. It includes two adult forms, including a self-report and a report by others, and norms that extend well into adulthood. It appears to have good potential for assessing adaptive behavior for diagnostic purposes. Average reliability coefficients of the adaptive skill areas across age groups range from .86 to .97, with the majority above .90 and corrected reliability coefficients of individuals with clinical diagnoses above .98. Norms for age birth to 5 years are expected to be available in 2002.

Battelle Developmental Inventory

The Batelle Developmental Inventory (BDI—Newborg et al., 1984) is a developmental scale, rather than an adaptive behavior scale, and is appropriate for children from birth to age 8 (Spector, 1999). It does not have the problems with floor effects in diagnosing developmental delays at the youngest ages that are present in other adaptive behavior scales. It contains broad domains similar to those found on adaptive behavior scales, which include: personal-social, adaptive, motor, communication, and cognitive. The BDI has well-documented reliability and validity, with norms based on a nationally representative sample of children (Harrington, 1985; Oehler-Stinnett, 1989). Several studies have shown significant and meaningful correlations between the BDI and other measures of cognitive, adaptive, language, and social functioning, with samples of children with and without disabilities (Bailey et al., 1998). The BDI is susceptible to "age discontinuities" (Boyd, 1989) or "differences in norm table layout" (Bracken, 1988) that are relatively common in measures of young children during this period of typically rapid development. This problem, and recommended strategies to avoid errors in diagnosis, are discussed in the section on norms.

Other Scales

The adaptive behavior scales described above have been consistently identified in research and practice reports as meeting criteria of technical excellence in measurement. Several other tests have been widely used and have many positive features but do not have the same reputation. Because clinicians are encouraged to utilize multiple measures in diagnosis, these other measures may be useful in providing supplemental or complementary information.

The Comprehensive Test of Adaptive Behavior (CTAB—Adams

& Hartleben, 1984) has been described as "fairly efficient and inexpensive," with "excellent reliability, solid validity, and adequate norms" (Reschly, 1990). It is reported to be appropriate for ages 5-22, yet it may not have a sufficient ceiling to discriminate performance levels among children above age 14 (Evans & Bradley-Johnson, 1988). A second limitation of this scale is that the standardization sample was limited to the state of Florida. Because Florida is a large and populous state with a culturally diverse population, it is likely that results can be generalized to the national population. Scores on the revised version of this measure, the CTAB-R, are based on a standardization sample that includes four of five regions of the United States (Adams, 2000).

The Adaptive Behavior Inventory (ABI—Brown & Leigh, 1986) was designed to "reflect the ability of school-age youngsters to meet age-appropriate socio-cultural expectations for personal responsibility" (Smith, 1989). It is appropriate for use with students ages 5 through 18 and is completed by the teacher. The ABI has a normative sample representative of all school-age children, including those with disabilities, and of a sample with mental retardation. The standardization sample was proportional in demographic characteristics to the 1980 census data. However, Smith (1989) notes that, at the low end of the normal intelligence norms, a few raw score points can dramatically change the adaptive behavior "quotient," and suggests that the norms on students with mental retardation are more useful. An attempt was made to select items that would avoid ceiling effects for the normal population and to ensure basal measures for the population with mental retardation. There is evidence that the ABI has adequate construct, content, and criterion-related validity, as well as internal reliability, but no data were provided on interrater reliability. Smith (1989) cited many problems with the norm tables but concluded the ABI could contribute some information to the determination of mental retardation.

The Independent Living Scales (ILS—Loeb, 1996) were designed to assess the degree to which older adults are capable of caring for themselves (i.e., functional competence). It requires an individual to demonstrate adaptive skills, rather than using a third-party informant or self-report to gather information on typical behavior. Reviews of the ILS have been generally negative, and it may not be suitable for disability determination purposes.

The Adaptive Behavior Evaluation Scale (ABES—McCarney, 1983) and the Parent Rating of Student Behavior (PRSB—McCarney, 1988) are used to identify mental retardation, learning disabilities, behavior disorders, vision or hearing impairments, and physical disabilities in students ages 5 to 21. Moran (2001) concluded that the information in the manual was not adequate to show how students with mental retardation differed from students with other disabilities. Norms are available to age 18 for the ABES and to age 12 for the parent scale. Reliability is good. High correlations with intelligence tests suggest it may be a duplication of this construct.

The Adaptive Behavior: Street Survival Skills Questionnaire (SSSQ—Linkenhoker & McCarron, 1983) was designed to assess adaptive behavior in youth from age 9 years and adults with mild to moderate mental retardation. The subscales are similar to general adaptive behavior scales, but there is a greater emphasis on skills required to function in community settings than on basic adaptive skills. It also differs from other adaptive behavior scales because it is administered as a test directly to the individual and, as such, does not measure typical performance in "real life." Haring (1992) found this to be an advantage in terms of its excellent reliability but noted that there were concerns about validity. Another concern was whether one may obtain a comprehensive picture of overall adaptation to the natural environment, because some skills could not be tested using the SSSQ's multiple-choice picture format. He suggested that the SSSQ could provide useful data when combined with the results of other comprehensive tests. To the extent that SSSQ data can predict entry or retention of competitive, gainful employment among people with mental retardation, it may have utility.

For the Social Skills and Vocational Success, Chadsey-Rusch (1992) described three measurement approaches to operationalize a definition of social skills, including (1) the perception of others in the workplace, especially employers, (2) the goals and perceptions of the target individual, and (3) performance of social behaviors in natural contexts. Perceptions of others are typically measured by sociometric ratings and behavior rating scales. The Social Skills Rating System, described below, is a behavior rating scale that was developed to provide this information for students. Sociometric ratings provide useful information but are impractical for diagnostic purposes, and the use of nonstandardized rating forms is not recommended for diagnosis of significant limitations in social skills. Direct measures from target individuals involve presenting them with hypothetical situations and conducting direct observations. It is unclear whether individuals with low-normal intelligence or mild mental retardation would be able to respond reliably to hypothetical situations.

The Social Skills Rating Scales (SSRS—Gresham & Elliott, 1987) is probably the best measure available of social skills adaptation in the school context. Although developed for school-age children, this scale may hold promise for adapted use with adults in work settings. In addition to rating skill performance, raters also specify whether each skill is critical to success in the environment in which the child is observed, i.e., school or classroom.

Table 4-1 shows the principal available adaptive behavior measures that are comprehensive in nature and their characterstics, including age range for use, age range of norm groups, date of publication, available versions, examiner requirements, appropriate scores for use in determining presence of adaptive behavior limitations, and assessed reliability of scores.

ASSESSMENT ISSUES IN ELIGIBILITY DETERMINATION

Relation of Principal Adaptive Behavior Scale Content to SSA Criteria

In Chapter 1 we provided the details of SSA's criteria for a disability determination of mental retardation in terms of both mental capacity and adaptive functioning. Adaptive behavior measures are useful in the identification of limitations concurrent with an IQ significantly below average. They also have utility in documenting delays or functional limitations consistent with marked impairment in motor development, activities of daily living, communication, social functioning, or personal functioning. These measures also may be validly used, with repeated or periodic administrations, for assessment of changes in status. Generally, however, adaptive behavior measures will be less effective in fine-grained analysis and classification of such problems as specific motor disorders or communication disorders and deficiencies in concentration, persistence, or pace.

SSA guidelines further clarify the intent and nature of activities of daily living and social functioning for adults, and personal functioning for younger and older children, closely paraphrased below:

- *Activities of daily living* include adaptive activities such as cleaning, shopping, cooking, taking public transportation, paying bills, maintaining a residence, caring appropriately for one's grooming and hygiene, using telephones and directories, and using a post office, etc. In the context of the individual's overall situation, the quality of these activities is judged by their independence, appropriateness, and effectiveness. It is necessary to define the extent to which the individual is capable of initiating and participating in activities independent of supervision or direction.
- The number of activities that are restricted does not represent a "marked" limitation in activities of daily living, but rather the overall degree of restriction or combination of restrictions must be judged.
- *Social functioning* refers to an individual's capacity to interact appropriately and communicate effectively with others. Social functioning includes the ability to get along with others, e.g., family members, friends, neighbors, grocery clerks, landlords, and bus drivers. A history of altercations, evictions, firings, fear of strangers, avoidance of interpersonal relationships, or social isolation may demonstrate impaired social functioning. Strength in social functioning may be docu-

TABLE 4-1 Principal Comprehensive Adaptive Behavior Measures and Their Characteristics

Adaptive Behavior Measure[a]	Age Range: Use	Age Range: Norms	Year Published	Versions
AAMR Adaptive Behavior Scale-Residential and Community	18-79 years	18.0 to 60+ years N = 4,103 people with DD	1993	Children’s version (see belo
AAMR Adaptive Behavior Scale-School	3-18 or 3-21 years	3.0-18.11 years N = 2,074 students with MR; N = 1,254 students w/o MR	1993	Adult version (see abov
Adaptive Behavior Assessment System	5-89 years	5-21 years; N = 1,670 &1,690; general population 16-89 years; N = 920 & 990; general population	2000	-Parent fo -Teacher -Adult for

d Their

r lished	Versions	Examiner Requirements[b]	Appropriate Scores	Principal Reliabilities
03	Children's version (see below)	Completion by a professional; or completion by a paraprofessional, with professional supervision (perhaps Class C, not specified)	-Personal self-sufficiency -Community self-sufficiency -Personal-social responsibility & 10 domain scores	Test-retest: (N = 45) -Factors: r = .93 to .98 -Domains: r = .88 to .99 Interrater: (N = 16) -Factors: r = .97 to .99 -Domains: r = .83 to .99
03	Adult version (see above)	Completion by a professional; or completion by a paraprofessional, with professional supervision (perhaps Class C, not specified)	-Personal self-sufficiency -Community self-sufficiency -Personal-social responsibility & 9 domain scores	Test-retest: (N = 45) -Factors: r = .72 to .79 -Domains: r = .75 to .95 Interrater: (N = 15) -Factors: r = .98 to .99 -Domains: r = .95 to .99
00	-Parent form -Teacher form -Adult form	Completion by a professional; or completion by a paraprofessional, with professional supervision (perhaps Class C, not specified)	-Global Adaptive Composite (GAC) -10 Domains: communication; community use; functional academics; home/school living; health & safety; leisure; self-care; self-direction; social; work	(Parent Form) Test-retest: (N = 102) -GAC: r = .96 -Domains: r = .83 to .94 Interrater: (N = 81) -GAC: r = .84 -Domains: r = .57 to .82

Continued on next page

TABLE 4-1 Continued

Adaptive Behavior Measure[a]	Age Range: Use	Age Range: Norms	Year Published	Versions
Comprehensive Test of Adaptive Behavior-Revised	Birth-60+ years	5-22 years; N = 2,094; students with MR 10-60+ years; N = 4,456; with MR 5-22 years: N = 4,525; students w/o MR	2000	-Normati Adaptive Behavior Checklist Revised (NABC-R compose a subset CTAB-R i -Also a parent/gu form of t CTAB-R
Scales of Independent Behavior-Revised	3 months-90 years	3 months-90 years; N = 2,182; general population	1996	-Short fo -Early developm form -Other re instrume
Vineland Adaptive Behavior Scales[c]	1-99 years	0.1 to 18.11 years N = 3,000 general population	1984	-Intervie survey fo -Expande form -Classroo edition

r lished	Versions	Examiner Requirements[b]	Appropriate Scores	Principal Reliabilities
0	-Normative Adaptive Behavior Checklist-Revised (NABC-R) is composed of a subset of CTAB-R items -Also a parent/guardian form of the CTAB-R	Completion of NABC-R by a parent or guardian Completion by a professional; or completion by a paraprofessional, with professional supervision (perhaps Class C, not specified)	-Total score -7 Domains: self-help; home living; independent living; social skills; sensory-motor; language/academics	(School sample) Test-retest: (N = 58) -Total: r = .99 -Domains: r = .98 to .99 Interrater: (N = 32) -Total: r = .99 -Domains: r = .95 to .99
6	-Short form -Early development form -Other related instruments	Completion by a professional; or completion by a paraprofessional, with professional supervision (possibly Class C for interpretation of scores)	-Broad Independence Score (BIS) -4 Cluster scores: motor skills; social interaction & communication skills; personal living skills; community living skills	(Children w/o MR) Test-retest: (N = 31) -BIS: r = .98 -Clusters: r = .96-.97 Interrater: (N = 26) -BIS: r = .95 -Clusters: r = .88-.93
4	-Interview survey form -Expanded form -Classroom edition	Class C; or completion by social worker or educator	-AB composite -Communication -Daily living -Motor -Socialization	(Interview survey form) Test-retest: (N = 484) -Composite r = .88 -Domains r = .81 to .86 Interrater: (N = 160) -Composite r = .74 -Domains r = .62-.78

Continued on next page

TABLE 4-1 Continued

---

[a] Comprehensive adaptive behavior measures are those that assess adaptive behavior through multiple items in multiple domains of functioning. Typically these measures are structured in terms of factors, domains, and subdomains or scales.

[b] Test publishers use criteria to determine whether professionals and others may purchase testing materials (e.g., protocols, kits, scoring manuals). The highest level of requirement is Class C, which means that a person has specific training and experience in psychometric assessment and meets other criteria typically consistent with those in the 1985 and more recent Standards for Educational and Psychological Testing.

[c] The Vineland Adaptive Behavior Scales are undergoing revision, and a new edition should be available within one to two years.

mented by an individual's ability to initiate social contacts with others, communicate clearly with others, interact, and actively participate in group activities. Cooperative behaviors, consideration for others, awareness of others' feelings, and social maturity also need to be considered. Social functioning in work situations may involve interactions with the public, responding appropriately to persons in authority, e.g., supervisors, or cooperative behaviors involving coworkers.

- A "marked" limitation is not represented by the number of areas in which social functioning is impaired, but rather by the overall degree of interference in a particular area or combination of areas of functioning.

- *Personal functioning in preschool children* pertains to self-care, i.e., personal needs, health, and safety (feeding, dressing, toileting, bathing; maintaining personal hygiene, proper nutrition, sleep, health habits; adhering to medication or therapy regimens; following safety precautions). Development of self-care skills is measured in terms of the child's increasing ability to help himself or herself and to cooperate with others in taking care of these needs. Impaired ability in this area is manifested by failure to develop such skills, failure to use them, or self-injurious actions. This function may be documented by a standardized test of adaptive behavior or by a careful description of the full range of self-care activities. These activities are often observed not only at home but also in preschool programs.

- *Personal functioning in adolescents pertains to self-care.* It is measured in the same terms as for younger children, the focus, however, being on the adolescent's ability to take care of his or her own personal needs, health, and safety without assistance. Impaired ability in this area is manifested by failure to take care of these needs or by self-injurious actions. This function may be documented by a standardized test of adaptive behavior or by careful descriptions of the full range of self-care activities.

The overall correspondence of several adaptive behavior measures to the content within the functional areas that are considered in ascertaining marked limitations is shown in Table 4-2. Each of the four adaptive behavior measures included in the table collects or assesses information regarding developmental status or performance in the areas of motor development, activities of daily living, communication, social functioning, and personal functioning. This table is a useful means to summarize and illustrate the detailed description of adaptive functioning that meets listing criteria, which are required to establish eligibility for SSI and DI.

Sensitivity of Scales at Ranges in Which Diagnostic Confirmation Is a Priority

Because adaptive behavior scales are targeted either specifically at children and adolescents or at groups ranging from children to young adults, there is a strong developmental component to their structures (Widaman et al., 1987). Such scales sample behaviors that are typically achieved at a range of ages and can indicate strengths and weaknesses in the ability to adapt. However, this also means that most scales are structured in steps that permit sampling of typical developmental tasks at each age. For any given age, it is unlikely that developmental tasks will be oversampled. In fact, as noted above, in the construction of adaptive behavior scales, such oversampling is typically avoided. Therefore, these instruments generally do not have firm cut-

TABLE 4-2 Correspondence Between SSI Classification Domains and Domains or Subdomains in Prominent Adaptive Behavior Measures

SSI Classification Domain	AAMR–ABS[a]	ABAS	SIB-R	VABS
Motor development	-Physical development	-Health & safety	-Motor skills (gross & fine)	-Motor skills (gross & fine)
Activities of daily living	-Independent functioning -Domestic activity	-Self-care -Home living	-Eating & meal preparation -Toileting -Dressing -Personal self-care -Domestic skills	-Personal daily living skills -Domestic daily living skills
Communication	-Language development -Numbers & time	-Communication -Functional academics	-Language comprehension -Language expression	-Receptive communication -Expressive communication -Written communication

Social functioning	-Social behavior -Conformity -Trustworthiness -Social engagement -Disturbing interpersonal behavior -Personal-social responsibility -Self-direction -Socialization	-Social -Self-direction	-Social interaction -Leisure -Internalized maladaptive behavior -Asocial maladaptive behavior -Externalized maladaptive behavior	-Interpersonal relationships -Play & leisure time -Coping skills -Maladaptive behavior
Personal functioning	-Independent functioning -Economic activity -Responsibility	-Community use -Work	-Home/community orientation -Time & punctuality -Money & value -Work skills	-Community daily living skills

NOTE: AAMR-ABS = American Association on Mental Retardation Adaptive Behavior Scales-School version (Lambert et al., 1993a); Residential and Community version (Nihira et al., 1993); ABAS = Adaptive Behavior Assessment System (Harrison & Oakland, 2000a); SIB-R = Scales of Independent Behavior-Revised (Bruininks et al., 2000); VABS = Vineland Adaptive Behavior Scales (Sparrow et al., 1984c) off points at each age at which adaptive behavior is considered to be substantially or markedly impaired. The determination of whether adaptive deficits are marked in character requires clinical interpretation informed in part by the data provided by the scoring of adaptive behavior measures.

The committee discussed the use of specific numeric cutoff points at which a marked deficit is present. There are few data on which to base such a decision. In fact, only one adaptive behavior test manual provides data that would be useful for answering this question. As a result, the committee commissioned Monte Carlo simulations to understand better the implications of requiring a specific numeric cutoff point. These data are presented in Chapter 5.

Here we review data from the only adaptive behavior test manual that could provide some guidance about the use of cutoff points for adaptive behavior assessment. Some caution is needed in interpreting these findings, because the study is small and unrepresentative. It may, however, be useful for identifying some of the issues likely to arise in setting a specific numeric cutoff point.

Using IQ as a parallel, it might seem that a reasonable cutoff score on an adaptive behavior scale could be a composite score or several scale scores of two standard deviations below the mean (i.e., –2 SD). However, depending on the functional domain and the measure, many people with a diagnosis of mild mental retardation do not have adaptive delays or limitations to this degree. Table 4-3, adapted from Harrison and Oakland (2000b), shows the percentage of adaptive behavior domain scores for a sample of children with mild mental retardation (N = 66) and controls without mental retardation matched for gender, age, and socioeconomic status (N = 66) that scored below the –2 SD standard on the teacher form of the ABAS. Because the sample for mild mental retardation shown in Table 4-3 was drawn from a school-age sample, some members of the group may have had IQs above 70 (i.e., less than 2 SDs below the mean). It should also be noted, however, that there is no gold standard against which cutoff scores could be appraised, or research from which a "true" proportion of people with observed IQs in the range of approximately 65-75 who

TABLE 4-3 Percentage of People Ages 5-18 Lower Than Two Standard Deviations Below the Mean on the Domains of the Adaptive Behavior Assessment System

	Mild Mental Retardation % ≤ –2 SDs	Matched Controls % ≤ –2 SDs	Ratio Mild MR/Matched
Functional academics	75.8	9.1	8.3
Self-care	51.5	13.6	3.8
Community use	63.6	18.2	3.5
Social	40.9	16.7	2.5
Self-direction	54.6	22.7	2.4
School living	40.9	16.7	2.4
Communication	66.2	28.8	2.3
Leisure	28.8	15.2	1.9
Health and safety	37.9	24.2	1.6
Composite or summary score	50.0	13.6	3.7
% with at least 2 scores ≤ 2 SDs	75.8	31.8	2.4

NOTE: When assessed as mean or total scores, row differences were significant at $p < .0001$ with t-tests for matched samples. SOURCE: Data from Harrison & Oakland (2000b, p. 89).

manifest adaptive limitations consistent with mental retardation may be directly projected. But it does seem reasonable, based on clinical experience, that any cutoff scores used should classify a majority of people in this IQ range as having adaptive limitations consistent with mental retardation.

Several important conclusions are suggested by this comparison, although the specifics might differ if another adaptive behavior scale with a somewhat different structure were used:

1. One-half of children with mild mental retardation (column 2) had a summary score falling in the marked limitation range (≤ 2 SDs), compared with about 14 percent of children without mental retarda-

tion (column 3). *One-half of children (and adolescents) with diagnosed mental retardation did not have summary scores falling in this range.*

2. About three-quarters of children with mild mental retardation had two or more domain scores falling in the marked limitation range, compared with slightly less than one-third of the children without mental retardation (last row).

3. There is considerable overlap in adaptive behavior attainment among children with mild mental retardation and matched peers. Even in those domains in which the greatest differences in attainment are evident, between 24 and 48 percent of children with mild mental retardation do not have adaptive limitations meeting a standard of marked limitation in those domains (although many may have such limitations in other domains).

4. Children with mild mental retardation were most likely to have adaptive behavior skills consistent with marked limitation (e.g., ≤ -2 SDs) in the domains of functional academics, communication, and community use.

5. Children without mild mental retardation were most likely to have adaptive behavior skills consistent with marked limitation in the domains of communication, health and safety, and self-direction.

6. Based on the ratio of marked deficits in the two groups (column 4), children with mild mental retardation were *much* more likely to have deficits in functional academics (especially), self-care, and community use and more likely to have deficits in social, self-direction, school living, and communication than children without mental retardation.

These findings are also evident in analyses of other adaptive behavior scale data (Jacobson, 1997) and client registry or population data sets (Jacobson, 1992; Larson et al., 2000; Lubin et al., 1982). Moreover, the findings suggest that eligibility criteria for SSI and DI applicants with IQs in the range of –2 to –2.66 SDs, by including the presence of marked limitations in activities of daily living, communication, social functioning, and personal functioning, will exclude a substantial number of people with IQs in this range. These score patterns also suggest that if a criterion of –1.0 to –1.5 SD was adopted as a cutoff point for each of several limitations in different domains, a much smaller percentage of people would be excluded from eligibility. Indeed, this latter approach is consistent with guidance already noted from the Social Security Administration.

General Clinical Practice and the Selection and Use of Adaptive Behavior Scales

In a 1990 survey, Archer et al. (1991) found that two adaptive behavior scales ranked 20th and 21st among 29 specific psychological assessment measures used by psychologists serving adolescents, and they were used "frequently" by fewer than 10 percent of practitioners. Adaptive behavior scales were seldom used as components of assessment batteries. A more recent study by Watkins et al. (1995) found that only 13 percent of respondents in the sample of clinical psychologists engaged in ability testing as part of their clinical practice, but 66 percent engaged in intellectual assessment. Adaptive behavior scales were not among the types of measures (or named measures) that respondents identified as important for the training of future practitioners. One particular adaptive behavior scale was ranked 17th in use among 38 scales, but less than 15 percent of respondents reported using this scale frequently. In their review of child assessment practices of psychologists, Kamphaus et al. (2000) mention the utility of adaptive behavior measures as components of multidimensional models of functioning and child psychopathology but did not include any adaptive behavior scales in an extensive listing of prominent measures that are currently used in psychological testing in schools. These findings are generally consistent with other findings regarding service utilization, showing, for example, that people with mental retardation, regardless of age, are less likely than others in need to receive psychological services in the community mental health systems, including assessment services.

In contrast to the studies above, which focus more on general clinical practice, direct surveys of school psychologists have disclosed that adaptive behavior scales are among the measures used most frequently. These studies also show that use of adaptive behavior scales has been growing over time (Hutton et al., 1992; Ochoa et al., 1996; Stinnett et al., 1994). Presumably, these findings reflect the importance of the schools as a setting in which the presence of mild mental retardation is frequently first identified and the use of adaptive behavior scales as a component of this practice. Nonetheless, available research on the clinical use of adaptive behavior scales for diagnosis and treatment-related purposes by either school psychologists or community clinical psychologists appears to be relatively sparse and does not focus on groups of people with mental retardation, as such (i.e., Clinger et al., 1988; McNamara et al., 1994; Pearson & Lachar, 1994; Roberts et al., 1993; Voelker et al., 1990; Wolber et al., 1997).

Taken together, these findings suggest that the primary cadre of psychologists with experience and expertise in the use of adaptive behavior measures, those who are most likely to use them in assessment and classification of mental retardation, consists of school-based practitioners. At the same time, it has been noted that use of adaptive behavior measures in the process of identifying children with mild mental retardation, especially, may be forgone in many schools because the academic failures or behavioral problems that prompt teacher referral of students, in and of themselves, may be considered valid indicators of limitations, deficits, or delays in adaptive behavior (McCullough & Rutenberg, 1988). However, children who do meet intellectual and functional criteria for mental retardation also are classified as having disabilities other than mental retardation in some schools and in some cases, and not necessarily consistently so (McCullough & Rutenberg, 1988). Finally, it has been suggested that adaptive behavior and social competence represent an important facet of adjustment in academic contexts, as important if not more so than intelligence (Forness et al., 1998). It is not evident that adaptive behavior scales efficiently differentiate among individuals with diverse disabilities that can be described as mild in degree (e.g., mild mental retardation, hyperactivity, specific learning disability) (Gresham & MacMillan, 1997; Zigler et al., 1984).

Although schools may be the service settings in which adaptive behavior measures are most likely to be used, the information derived from these assessments may not be considered meaningful for the purposes of classification by decision makers. It is possible that the use of these measures is diminishing in local educational agencies, with increasing emphasis on specific educational classification in many states. As a result, they may become less available for use in disability determinations. It also appears that community practitioners, aside from those associated with developmental disabilities clinics or centers or with community developmental disabilities services, may not be well versed in the use and interpretation of adaptive behavior measures or prepared to apply different measures in different situations for different purposes.

Clinical Considerations in the Selection and Use of Adaptive Behavior Scales

In the committee's view, adaptive behavior is an essential component of the mental retardation diagnostic construct, and all agencies contemplating mental retardation diagnoses should give consideration to adaptive behavior. This consensus rests on the accumulated wisdom in the field of mental retardation, including the fact that adaptability in meeting the demands of everyday living was fundamental to conceptions of mental retardation long before effective tests of intellectual functioning were developed. Adaptive behavior has been fundamental to conceptions of mental retardation at least since the early 19th century (Doll, 1936a, 1967).

Choice of Assessment Methods

In the committee's view, the use of standardized measures of adaptive behavior is potentially valuable in the overall assessment of adaptive behavior. The value of such instruments will depend on careful decisions about instrument choice, score interpretation, and consideration of other information that bears on adaptive status. The committee does not recommend any specific list of instruments, but choices should be guided by the reviews of the available instruments in this chapter and the research literature on existing and new instruments.

Instrument Choice

Informed judgments are required about using an adaptive behavior instrument in the evaluation of SSA eligibility based on a diagnosis of mental retardation. For most people, the use of currently available adaptive behavior instruments along with other information on adaptive functioning will improve decisions about mental retardation classification. The decision on which standardized instrument to use must be informed by knowledge of the following characteristics of clients, respondents, and instruments.

1. The instrument must be appropriate to the age of the client and the client's approximate functioning level. Currently available instruments typically do not provide sufficient coverage of some aspects of adaptive functioning for adolescents and adults who are functioning in the IQ range of 60 to 75, the range in which diagnostic decisions about mental retardation are most difficult. Careful analysis of the appropriateness of the item content and density of specific measures is crucial.

2. An appropriate respondent, knowledgeable about the examinee, is required for most adaptive behavior instruments. If an appropriate respondent is not available, use of the instrument in some other way (e.g., self-report, unless a self-report version of the protocol is available) violates basic standardization procedures, rendering normative comparisons invalid. There are some techniques that extend the range of appropriate respondents. For example, some instruments permit more than one respondent to answer different items, depending on which respondent is most knowledgeable about the behavior (Adams, 2000). Another instrument permits adult client self-report

(Harrison & Oakland, 2000a), a promising method that needs further study of its feasibility with clients having IQ scores in the range of 60 to 75 (Millham et al., 1978). Still another way to extend respondent usefulness has been to permit guessing on items involving behaviors that have not been observed (e.g., Harrison, 1984; Harrison & Oakland, 2000a). These results become increasingly unreliable and invalid as the number of guesses increases.

3. Scores from the instrument that are useful in diagnostic decisions must be provided and, in turn, interpretations need to be guided by the structure and organization of the adaptive behavior inventory. Diagnostic decisions about mental retardation nearly always involve normative comparisons using various derived scores. Standard score scales are preferred for these comparisons. Useful score scales and appropriate norms are vital features of adaptive behavior instruments used in diagnostic decisions.

4. The instrument's structure must guide interpretation. For example, the ABAS is organized into 9 or, depending on age, 10 adaptive skills areas, clearly reflecting the 1992 AAMR manual on classification (American Association on Mental Retardation, 1992), which also identified 10 adaptive skills areas. Factor analysis results do not support the existence of more than one overall ABAS general factor. Interpretation should focus on the composite score or, perhaps, implement the AAMR classification criterion of deficits in two or more adaptive skills areas. Use of 3 to 5 group factor scores, appropriate with the SIB, the VABS, and some other instruments, would not be appropriate with the ABAS.

5. Interpretation of the results of instruments must consider the possible influence of unintentional response sets as well as more deliberate efforts to raise or lower the adaptive behavior results in order to achieve certain outcomes. One important protection against inappropriate adaptive behavior decisions, which is due to respondents intentionally reporting invalidly low adaptive behavior performance, is consideration by the examiner of further information about everyday

performance of social roles and related activities that are consistent with presence of adaptive limitations.

Typical Performance Measurement and the Nature of Adaptive Behavior

Perhaps the most fundamental problem with regard to adaptive behavior measurement is the relationship of existing measures to the conceptions of the underlying construct. Is adaptive behavior a set of abilities and skills useful in coping with environmental demands that are mastered by the individual? Is adaptive behavior the consistency with which an individual performs various skills in coping with environmental demands? The former, mastery, suggests assessing what people can do, whereas the latter suggests assessing what people typically do.

Mastery assessments typically involve direct testing of a person with a set of tasks administered under standardized conditions that permit comparisons to norms or to absolute standards for performance. In the mastery approach, examinees are encouraged to do their best, and testing conditions are designed to elicit maximum performance. Traditional measures of achievement and intellectual functioning are examples of maximum performance tests.

Measures of typical performance involve an attempt to assess what an individual typically does or how a person usually feels (Cronbach, 1990). Individuals or third-party respondents are asked to indicate their usual feelings or behaviors, not their best or most positive feelings or behaviors. Inventories or checklists that attempt to assess vocational interests, emotional status, and personality traits nearly always are typical performance measures.

The adaptive behavior construct has both typical performance and maximum performance elements, a characteristic that complicates measurement operations. The typical adaptive behavior inventory inquires about both the individuals' skills or abilities (what they can do) and about what they usually do in various circumstances (performance of skills or typical performance). Various measures differ on the degree to which they attempt to assess the can-do and does-do characteristics of adaptive functioning (Adams, 2000; Sparrow et al., 1984a). This leads to some unique problems with these measures. The skills or abilities items may be readily assessed through direct measures of the individual with behavioral tasks, while performance or does-do features can be assessed only through extensive behavioral observations that often are impractical given the breadth of the adaptive behavior construct and the number of relevant settings. Therefore, the most common method of assessing adaptive behavior is through a third-party respondent, although many difficult measurement issues arise in such assessments.

The distinction between maximum performance and typical performance assessments might be illustrated with some common adaptive behavior items. Use of a telephone is a common item on communication subdomains of many adaptive behavior inventories. Looking up a phone number is a relevant item for that subdomain. This skill could be assessed by giving the individual a local phone directory, asking them to look up a number, and observing the results, yielding a measure of whether the individual can demonstrate this skill. Another method is to ask a third-party respondent (e.g., a parent or teacher) either very general questions about communication skills, a more specific question about telephone use, or a direct question like "Does Edouard find a number by looking it up in a telephone book?" As a general rule, skills can be measured directly with greater accuracy, reliability, and validity than performance. It is, however, performance or typical performance that most adaptive measures address. In addition to the usual considerations attendant on any measure of individuals, special considerations always exist with typical performance measures (Cronbach, 1990).

Some Differences Among Typical Performance Measures

Adaptive behavior measures differ regarding the use of typical or maximum performance methods. The most widely used measures use a typical performance approach involving third-party respondents (Bruininks et al., 1996; Harrison & Oakland, 2000b; Lambert et al., 1993b; Sparrow et al., 1984b), although several differences exist among the response formats for items in these measures. Adams (2000), in contrast, uses a mixture of typical performance with third-party respondents and maximum performance operations.

This chapter has discussed problems with item and score scale floors and ceilings, item sampling in relation to the behaviors that are most problematic at different developmental periods, item density, reliability of informants, and validity of informants. It is crucial that people conducting or interpreting adaptive assessments take these problems into account.

Structured Versus Unstructured Interviews

There are vast differences in how the third-party respondent reports on the adaptive behaviors of a client, particularly in the structure for the interview. Interview methods recommended for different measures vary from high to low structure. On one instrument, the items are shown to the respondent and the respondent is given responses from which to choose (e.g., Bruininks et al., 1996), while in another the interviewer is required to assess adaptive competencies through a general conversation with prompts such as "Tell me about Thomasina's language skills" (Sparrow et al., 1984a). The latter measure requires a more skilled interviewer, as well as a relatively verbal respondent who spontaneously offers sufficient information to permit the interviewer to determine scores on items, or evocation of relevant information through prompts for further details. Many of the parents of children suspected of meeting criteria for mild mental retardation may themselves experience cognitive difficulties and are relatively nonverbal (Ramey et al., 1996), making semi- or unstructured interviews much more difficult to conduct in a standardized manner.

Response Sets

Highly structured interviews have unique problems as well. A number of well-known, often unintentional, response sets are especially applicable to the more structured third-party respondent measures and, to a lesser extent, the less structured approaches (Anastasi & Urbina, 1997; Cronbach, 1990; Sattler, 1988, 1990). Response options such as "never," "sometimes," and "always" or Likert-type response formats using a five- or seven-point continuum with anchor statements like "very good," "good," "acceptable," "poor," and "very poor" are open to a variety of response sets. Some of the more common and relevant response sets are (a) social desirability, involving responses consistent with positive or desirable connotations of the items or behaviors, (b) acquiescence, involving the tendency to say "yes," "true," or, in cases in which knowledge is lacking or uncertain, "sometimes," and (c) halo effects, involving reporting higher adaptive behavior among persons who are more liked by the respondent.

RESEARCH AREAS

Adaptive Behavior: A Social-Cognitive Perspective

In recent years, existing operational definitions of adaptive behavior and techniques for measuring adaptive behavior have been criticized as being inadequate for determining the presence of significant limitations in individuals with mild mental retardation. In particular, criticism has focused on the inadequacy of existing techniques for measuring the social domain of adaptive behavior (MacMillan et al., in press). This attention to and concern about the assessment of social behavior is hardly surprising, given the prominent place that social behavior has historically occupied among the domains of day-to-day life that constitute adaptive behavior for individuals with mental retardation.

The social domain is particularly important to assess for individuals with mild mental retardation because prominent limitations that these individuals experience are often in the domain of interpersonal relationships, rather than in skill domains that are not predominantly social in nature (e.g., activities of daily living, motor skills). Interestingly, individuals with mild mental retardation often face their most significant obstacles to competitive employment and job retention arising not from task-related skills, but rather from limitations in their social functioning (Bullis & Foss, 1986; Butterworth & Strauch, 1994; Chadsey-Rusch, 1992; Foss & Bostwick, 1981; Greenspan & Shoultz, 1981; Salzberg et al., 1988; Salzberg, Likins et al., 1986).

A major reason why proper assessment of the social domain of adaptive behavior can be a challenge for eligibility examiners is that the limitations in social functioning in individuals with mild mental retardation are often difficult to quantify with available assessment methods. Often, these limitations take the form of poor judgment or social rule violations, rather than markedly maladaptive behavior (Greenspan & Granfield, 1992). As Switzky et al. (1996) put it, "relatively immature, as opposed to deviant, social behavior has always been the central distinguishing feature of this disability" (pp. 7-8). Furthermore, these limitations may be more noticeable in certain settings or circumstances than in others (Greenspan, 1999). For example, an individual with mild mental retardation may be familiar with the routines and may benefit from the supports that staff provide in the supervised apartment program in which he lives. However, he may have difficulty keeping a job because he has difficulties interacting with a changing cast of coworkers or customers who are unfamiliar with his social style.

Clearly, examiners need to be concerned about whether available instruments and practices for assessing adaptive behavior adequately document critical difficulties in social functioning that prevent individuals with mild mental retardation from fulfilling key roles and expectations in society. At present, a variety of assessment instruments have been employed in research and clinical settings that attempt to capture these individuals' social limitations. The available instruments include indirect assessments, such as rating scales that assess an individual's typical performance of social behaviors based on information from informants, including teachers, parents, and job coaches, as well as instruments that permit direct assessment of the processes that, according to theoretical models of successful social adaptation, under-lie the ability to perform situationally appropriate behavior. With regard to the direct assessment of processes, the overarching construct of social cognition has been put forth by developmentalists over the past four decades (e.g., Crick & Dodge, 1994; Dodge, 1986; McFall, 1982; Trower, 1982). Social cognition encompasses such constructs as social problem solving (Spivack & Shure, 1974), decision making (Hickson & Khemka, 1999), and social and emotional learning (Elias et al., 1997). Overall, the construct of social cognition represents the cognitive aspects of social functioning.

As social cognition has matured as a research discipline, researchers have developed methodologies for assessing social-cognitive processes and have demonstrated the usefulness of these methodologies for detecting the limitations that individuals with mild mental retardation exhibit in their ability to adapt to changing social situations. Currently, instruments and interviewing procedures for assessing social-cognitive processes can provide examiners with valuable supplemental information about the social functioning of individuals with mental retardation. This information may otherwise be lacking because of inadequacies in existing adaptive behavior measures. In addition, social-cognitive assessment can also be helpful in establishing that an individual's social problems are indeed a manifestation of an underlying cognitive impairment (in accordance with standard definitions of mental retardation), rather than primarily reflecting other factors, such as environmental constraints or motivational characteristics. Thus, social-cognitive assessment increases the likelihood of making accurate diagnostic and disability determination decisions by increasing the pool of information available to an examiner regarding an individual's functional limitations, while simultaneously reducing the risk of "false positive" decisions.

The potential utility of social-cognitive assessment techniques for the evaluation of individuals with mild mental retardation for eligibility determination purposes is evident in focusing on three major skill areas: social perception, the generation of strategies for resolving social problems, and consequential thinking. These three processes occupy a prominent place in most theoretical models of social cognition (e.g., Crick & Dodge, 1994; Gumpel, 1994; Leffert & Siperstein, in press; McFall, 1982).

Social perception refers to an individual's ability to interpret or "read" relevant social messages from others (Maheady et al., 1984). These messages, known as social cues, consist of verbal and nonverbal stimuli, such as physical actions, words, facial expressions, tone of voice, and body language, which tell about others' behaviors, feelings, and intentions. Individuals with mental retardation often demonstrate difficulties at the most basic level of recognizing specific types of social cues (e.g., recognizing a person's emotional state on the basis of his or her facial expression) (Adams & Markham, 1991; Gumpel & Wilson, 1996; Harris, 1977; Hobson et al., 1989). Research with children and adolescents with mental retardation has found that they also have difficulty integrating information from multiple cues in order to interpret a social situation (Brosgole et al., 1986; Doman, 1967; Gomez & Hazeldine, 1996; Leffert & Siperstein, 1996; Leffert et al., 2000; Maheady et al., 1984).

The assessment of social perception skills in individuals with mild mental retardation has involved a variety of instruments, with subsequent methodological refinements, which have been developed and employed over four decades with children, adolescents, and adults. For example, the Test of Social Inference (TSI—de Jung et al., 1973) employs the technique of presenting an individual with mild mental retardation with illustrations of common social situations and asking him or her, for each illustration, to tell the examiner what the picture is about. Of the various social perception assessment instruments that have been developed, the TSI is the instrument that has been used most widely to assess social perception skills in this population (de Jung et al., 1973; Matthias & Nettelbeck, 1992).

The second social-cognitive process is the generation of strategies for resolving social problems. Through the assessment of strategy generation, researchers have been able to evaluate an individual's fund of social knowledge (i.e., his or her repertoire of social strategies), as well as the ability to adapt to varied social situations by generating situationally appropriate strategies. Research with individuals with mental retardation has consistently documented limitations in their performance of both of these components of strategy generation.

Regarding strategy repertoires, for example, researchers have found that children and adults with mental retardation have a limited repertoire of appropriate social strategies to draw from (Herman & Shantz, 1983; Smith, 1986). Their repertoires often exclude certain types of socially adaptive strategies. For example, they rarely employ the strategy of attempting to work out a mutually acceptable compromise solution in instances when one's interests conflict with another person's wishes (Hickson & Khemka, 1999; Hickson et al., 1998; Jenkinson & Nelms, 1994; Smith, 1986).

Regarding the ability to adjust one's social strategies to fit the needs of a particular social situation, children with mild mental retardation often fail to use information from the specific social cues present in the social situation to guide their search for appropriate strategies (Leffert et al., 2000). As a result, they often rely on generic, one-size-fits-all strategies, such as appealing to an authority, rather than adjusting their strategies in accordance with situational demands. At the workplace, for example, overdependence on the strategy of appealing to a supervisor, regardless of the nature of the problem, can be a problem in creating tension with peers or in creating a perception that the person cannot function autonomously in a work setting.

Unlike the area of social perception, there is no single instrument for assessing strategy generation in individuals with mental retardation. Rather, there is a standard clinical methodology that consists of presenting the individual with a hypothetical situation in the form of a story and asking "What would you do if this happened to you?" The Social Problem-Solving Test (Castles & Glass, 1986) is an example of an assessment instrument employing this methodology that was specifically designed for use with individuals with mental retardation. The technique of asking open-ended questions that tap the process of generating strategies for solving problems is widely used in standard assessments of intellectual functioning, such as the Wechsler scales.

The third social-cognitive process, consequential thinking (i.e., reasoning regarding the consequences of carrying out different social strategies), is also a critical process in the adaptive behavior of individuals with mild mental retardation. Recent studies with these individuals have documented limitations in their reasoning about the consequences of strategies that make it difficult for them to select a social strategy that is appropriate for a given social situation. Rather than weighing the likely outcomes of enacting a particular strategy in a given situation, these individuals tend to render global judgments (e.g., whether a strategy is generally a good thing or bad thing to do) when evaluating behavioral options. Consequently, they have difficulty evaluating which potential strategy is best to enact in a particular situation (Hickson et al., 1998; Jenkinson & Nelms, 1994).

Methodologies for assessing consequential reasoning have existed for several decades. A widely used technique, which was first employed by Spivack and Shure (1974), is to present the individual with a story describing a social problem, as well as a particular strategy that the protagonist might use to resolve the problem. The person is then asked, "What will happen after" the protagonist carries out this strategy? Another, more open-ended assessment technique is to present a social problem and then to ask the interviewee to relate everything that is going through the protagonist's mind as he or she tries to decide what to do about resolving the problem (e.g., Hickson et al., 1998; Jenkinson & Nelms, 1994). A third approach is to ask individuals with mild mental retardation to select the best strategy for resolving a social problem from among several alternatives. This is the approach that has been utilized in the Test of Interpersonal Competence for Employment (TICE—Foss et al., 1986). The TICE, which is commercially available, consists of two subtests that assess the individual's ability to evaluate strategies in relation to the situational demands of two distinct social interaction contexts common in work settings: interaction with coworkers and interaction with supervisors. Auty and colleagues (1987) have found positive correlations between subtest scores on the

TICE and supervisor-rated work skills, self-reported job satisfaction, and worker productivity among adults with mild mental retardation.

Social-cognitive assessments have already demonstrated their usefulness as a supplement to standard adaptive behavior rating scales. They have provided valuable information that has informed decision making about interventions to improve the social functioning of individuals with mental retardation. Specifically, instruments such as the TSI, the Social Problem-Solving Test (Castles & Glass, 1986), and the TICE have been successfully employed with this population for the purposes of determining where to begin instruction in social skills and documenting the improvements that have resulted from instructional interventions. These assessment instruments, which have been useful in instructional contexts, can also be valuable for the evaluation of an individual's eligibility for SSA services.

The social-cognitive processes and the approaches that are used to measure them can also inform and enrich the interviews that examiners conduct with individuals with mild mental retardation and other informants. Table 4-4 presents examples of questions that can guide examiners in eliciting information regarding the three social-cognitive processes reviewed here. By eliciting information about an individual's performance of these processes, the examiner can increase the likelihood of detecting impairments in social functioning that often characterize this population. Additional measures of social cognition or social skills and issues relevant to social skills assessment of people with mental retardation have been discussed further (Bell-Dolan & Allan, 1998; Blacher, 1982; Blake & Andrasik, 1986; Jackson et al., 1981; Matson et al., 1983; Meyer et al., 1990; Monti, 1983; Smith & Greenberg, 1979; Van Hasselt et al., 1981).

Overall, as a supplement to standardized adaptive behavior assessment scales, social-cognitive assessment has the potential to contribute to the improvement of SSI and DI eligibility determination practices by enriching the pool of relevant information that is available for resolving uncertainty in decisions regarding impairment in the social domain. In this fashion, the assessment of social-cognitive processes can

TABLE 4-4 Question Guide for the Assessment of Social-Cognitive Processes

Social perception: encoding of social cues	Does the person recognize that a social problem has occurred?
	Can the person accurately describe the problem?
	Does the person attend to and recognize the specific social cues (e.g., facial expression, tone of voice, body language, sequence of actions) present in social situations that indicate others' emotions and intentions?
Interpretation of social cues	Does the person accurately interpret others' emotions and intentions on the basis of the available cues?
Strategy generation	Is the person familiar with and able to think of a variety of strategies that are potentially appropriate for resolving social problems?
	Does the person modify his or her strategy from one situation to another based on the type of social problem, the other person's intentions, and related considerations?
Strategy evaluation and selection	Can the person anticipate the consequences of carrying out different strategies for resolving particular social problems in a given social context?
	In doing so, does the person consider long-term and more abstract consequences of a carrying out a particular strategy (e.g., impact on maintaining interpersonal relationships or keeping one's job) as well as the immediate and more tangible consequences (e.g., personal gratification, escaping from an immediate problem)?

become part of a more comprehensive approach to individual assessment that includes information from informants about day-to-day behavior, as well as direct assessment of skills of a social-cognitive nature. The information obtained from this type of evaluation can clarify the nature and extent of the limitations those individuals with mild mental retardation experience in adapting their behavior to meet the social demands and expectations of the school, workplace, and recreational and residential settings.

Alternative Assessment Instruments

There are a small number of well-normed adaptive behavior scales that are especially suitable for use in initial determinations for children and youth with possible mental retardation. There is a much larger number of scales that do not have extensive norms but may nonetheless be suitable as a means of gathering and summarizing information that can be assessed on a clinical basis. In addition to summarizing adaptive behavior status for the purposes of diagnosis and establishing SSI and DI eligibility, some adaptive behavior scales, such as the AAMR Adaptive Behavior Scales (both school and residential or community versions) and the Scales of Independent Behavior permit the recording of maladaptive behavior. A record of maladaptive behavior may permit an individual to be qualified for SSI by virtue of concurrent IQ in the range of –2 to –2.66 SD and presence of another mental (or behavioral) disorder (Jacobson, 1990; Jacobson & Janicki, 1983).

Whereas situational and functional assessment are appropriate in intervention design for maladaptive behavior or behavior disorders or problems, norm-referenced or criterion-referenced instruments are appropriate for initial assessment (Reschly, 1992). Possibly the most thoroughly researched and well understood instrument to assess both prosocial and problem behavior among children generally is the Achenbach Child Behavior Checklist (CBC—see Achenbach & Edelbrock, 1978). The CBC differentiates between internalizing (e.g., withdrawal, diminished interests, depressive affect) and externalizing (e.g., aggression against peers or adults, property destruction) behavior problems that may be consistent with a mental disorder (see also Borthwick-Duffy et al., 1997; Fidler et al., 2000; Hodapp et al., 1997). Norms for several versions of the CBC, which are age-graded, are available for both clinical and nonclinical child and youth populations.

In addition to the CBC, there are other instruments available to assess overt behavior, affect, or verbal statements consistent with the presence of mental or behavioral disorders among children and youth with mild mental retardation. Generally, these instruments do not have well-established norms but rather have been assessed for their sensitivity with diagnosed cases (e.g., Reiss & Valenti-Hein, 1994), and some scales are more suitable for youth than for children: the Assessment of Dual Diagnosis (Matson & Bamburg, 1998); the Psychopathology Instrument for Mentally Retarded Adults (Balboni et al., 2000; Linaker, 1991; Sturmey & Ley, 1990; Watson et al., 1988); and the Reiss Screen for Maladaptive Behavior (Havercamp & Reiss, 1997; Prout, 1993; Sturmey & Bertman, 1994) For practitioners skilled in clinical interviewing, a field-tested adaptation of a structured clinical interview is available. This scale is the Psychiatric Assessment Schedule for Adults with Developmental Disability or PASS-ADD (Moss et al., 1996, 1998, 2000; Prosser et al., 1998). All of the measures above have demonstrated concordance with psychiatric diagnosis. They can contribute to accurate and appropriate clinical diagnosis of concurrent conditions that may meet listing requirements for SSI and DI as an alternative to assessment of adaptive behavior.

Finally, as this chapter is being written, the World Health Organization (WHO) has completed development of ICIDH-2, the International Classification of Functioning, Disability, and Health (World Health Organization, 2000; see also Post et al., 1999), a functionally based nomenclature. This nomenclature has dimensions of impairments of body functions, impairments of body structures, activity limitations and participation restrictions, and environmental factors. For the purposes of assessment focused on ascertainment of mental and physical conditions, the most salient measurement dimensions of ICIDH-2 are body functions (e.g., mental processes), and activity limitations and participation restrictions. In conjunction with the ICIDH-2, WHO has developed the World Health Organization Disability Assessment Schedule II (WHODAS II), which, in its most extensive form, contains 36 items tapping domains of: (1) understanding and communicating, (2) getting around, (3) self-care, (4) getting along with others, (5) household and work activities, and (6) participation in society. The utility of the WHODAS II remains to be established as a means to consolidate adaptive or maladaptive behavior information of value in informing the SSI and DI eligibility determination process.

CONCLUSIONS AND RECOMMENDATIONS

Review of the extensive literature on adaptive behavior and its assessment suggests that adaptive behavior is best viewed as a multidimensional construct. That is, current science suggests that there are various domains of behavior that form the construct of adaptive behavior. Factor analyses of existing measures finds consistent domains of functioning. These domains vary by age, consistent with the development of adaptive behavior. The committee therefore, makes two major recommendations to SSA:

> **Recommendation: Standardized adaptive behavior instruments should be used to determine limitations in adaptive functioning. In general, the cutoff scores for adaptive behavior should be one standard deviation below the mean in two adaptive behavior areas or one and one-half standard deviations below the mean in one adaptive behavior area.**
>
> - **Adaptive behavior measures should be used whenever possible, but only when there is an instrument that matches the client's characteristics and when an appropriate third-party respondent is available.**
> - **A client can be determined to have a significant limitation in adaptive functioning even with scores that do not meet the above criteria IF there is compelling evidence of adaptive behavior deficits that significantly impair performance of expected behaviors.**
>
> **Recommendation: Revisions should be made in the adaptive behavior areas or domains emphasized in SSA regulations to con-**

form to factor analysis results. The following areas by age should be adopted by SSA:

- **Infancy/early childhood (approximate ages birth to 4): motor/mobility, social, communication, daily living skills (self-help).**
- **Childhood (approximate ages 5 to 17): motor/mobility, social, communication/functional academic skills, daily living skills.**
- **Adolescence/adulthood (approximate ages 18 and older): motor/mobility, social, communication/practical cognitive skills, daily living skills, work skills/work-related behaviors.**

Current science also suggests that several measures of adaptive behavior tap into these domains. These measures have excellent psychometric properties, with reliabilities of about .90. Also, current measures also evidence strong validity, as described in the chapter. The committee has identified several measures that would be useful in disability determination for mental retardation.

Unlike intelligence tests, which measure maximum performance, adaptive behavior assessment focuses on what the individual typically does. Assessments work best when they document: (a) quantitative level of performance, (b) fluency of performance (e.g., qualitative criterion performance), (c) the extent to which the individual has failed to acquire skills or failed to perform skills already learned, and (d) the inability of the individual to perform skills through lack of opportunity. The committee's review of the scientific and practice literature also reveals that adaptive behavior is a broadly focused construct. The focus is on the ability of the individual to function independently, with minimal external supports, by adjusting his or her behavior in a self-guided fashion to meet varied situational demands and expectations.

Our review of the practice literature reveals that adaptive behavior scales are in wide use by some groups of clinicians. In our judgment, good clinical practice requires that data from standardized adaptive behavior scales be combined with other clinical or behaviorally oriented information in determining the presence or the absence of adaptive behavior deficits. These other bits of data could include a review of developmental and social history, direct observation of the individual's behavior, verbal reports from interviews, and the use of the other structured and semistructured interviews. Best-practice guidelines require that clinicians using adaptive behavior measures employ those that are culturally compatible and have suitably contemporary and age-related norms.

Finally, the committee has identified a number of research areas, focusing on which would improve the measurement of adaptive behavior for mental retardation diagnosis. These areas include social-cognitive and social skill assessment—with a specific focus on social cognitive processes of social perception, strategy generation, and consequential thinking—and vocational and work-related skills assessment with prognostic value. In addition, there is a strong need to fund studies examining the nature and distribution of adaptive behavior deficits among individuals with mental retardation in general and those with mild mental retardation more specifically.

Chapter 5

The Relationship of Intelligence and Adaptive Behavior

Determining whether a person has mental retardation involves complex decisions that integrate information on current intellectual functioning and adaptive behavior. Information about each of these core dimensions is always incomplete and dependent on imperfect measures of the underlying constructs. Judgment is therefore necessary when making decisions about how best to assess intellectual and adaptive functioning and in interpreting the results; this chapter provides guidance for those judgments. However, the guidance cannot take the form of absolute decision rules that replace judgment about the appropriateness and meaning of evaluation results. For this reason, high standards and much preparation are needed for the professional personnel making diagnostic decisions, including thorough knowledge of mental retardation as a diagnostic construct.

This chapter discusses the diagnostic implications of the preceding chapters on intellectual functioning and adaptive behavior, as well as a review of the literature on the relationship between measures of intellectual functioning and measures of adaptive behavior. That fundamental relationship has significant implications for the discussion of how diagnostic decisions are made by combining information across multiple domains of functioning, from multiple sources, and from multiple methods of gathering information. A principle of convergent validity will emerge in this discussion as critical to a sound diagnosis of mental retardation.

DIAGNOSTIC CONSTRUCT OF MENTAL RETARDATION

Diagnostic constructs have two key components: conceptual definitions and classification criteria. Both are critical to understanding the meaning of the diagnostic construct. The four conceptual definitions of mental retardation discussed in this report (see Chapter 1) do not suggest explicit classification criteria. Although there is controversy regarding some features of these diagnostic systems (MacMillan et al., 1993, 1995; Reiss, 1994), the conceptual definitions differ little among the diverse organizations involved. It should also be noted that many other organizations and agencies establish conceptual definitions and classification criteria for mental retardation, including the Social Security Administration (SSA) and state departments of education.

Broad consensus exists throughout the developed world about the basic features of the conceptual definition of mental retardation: it involves significant limitations in the core dimensions of intellectual functioning and adaptive behavior. Most national and worldwide diagnostic systems use the term “mental retardation,” and nearly all suggest that deficits in adaptive behavior arise because of limited intellectual functioning. The greatest variations in conceptual definitions and terminology occur in the legal requirements for classification of students as mentally retarded in the special education system of each state's department of education (Denning et al., 2000; Patrick & Reschly, 1982; Utley et al., 1987). Across the 50 states and the District of Columbia, different terminology is used (e.g., mental retardation, mental disability, significantly limited intellectual capacity), along with widely varying classification criteria. Nearly all states, however, define a disability based on deficits in the dimensions of intellectual and adaptive functioning.

Controversies regarding mental retardation diagnostic systems arise most often regarding classification criteria, that is, how the conceptual definition of mental retardation is operationalized. Classification criteria vary significantly regarding the cutoff scores that are adopted to determine which cases meet or do not meet diagnostic eligibility criteria. Higher cutoff scores, of course, increase the population with a diagnosis of mental retardation, and lower cutoff scores decrease it. A little-appreciated influence is the joint effect of IQ and adaptive behavior cutoff scores on diagnostic decisions. As is shown later, the combined effects of different cutoff scores can drastically alter the number of people who can be considered for a diagnosis of mental retardation. Classification criteria also vary regarding the use of composite and part scores as well as the number of part scores that may be used from measures of adaptive and intellectual functioning.

Classification Criteria for Intellectual Functioning

The cutoff scores for measures of general intellectual functioning are better established than the cutoff scores for measures of adaptive behavior. There is broad consensus in the major diagnostic systems that performance on the intellectual dimension must be approximately two or more standard deviations below the population mean, which translates into an IQ score of 70 or less on measures with a mean of 100 and a standard deviation (SD) of 15. The degree of flexibility around the cutoff score of 70 varies among diagnostic systems; some allow a range bounded by one standard error of measurement, which

TABLE 5-1 Proportions of People with Scores At and Below Different Cutoff Scores on a Normally Distributed Characteristic

Cutoff Score	Percentage Meeting or Exceeding Score
Below 70	2.28
70 and below	2.68
Below 75	4.75
75 and below	5.48
Below 80	9.18
80 and below	10.20
Below 85	15.87
85 and below	17.62

translates to about 4 points for measures with reliabilities above r = .90. Other systems make a general statement that IQ can be approximately 70 to 75.

The difference between a cutoff score at or below IQ ≤ 75 and a cutoff score at or below IQ ≤ 70 is dramatic, as shown in Table 5-1. Twice the proportion of people have scores at or below 75 (5.48 percent) than have scores at or below 70 (2.68 percent). In other words, a seemingly trivial change of five points on the intellectual dimension doubles the number of people from the given population that are potentially eligible for consideration on that dimension. Before raising an alarm with this statistical information, however, it should be noted that no prevalence study of people identified as having mental retardation has ever approached the level of 5 percent of the general population, at least in part because of the necessity of a concurrent deficit in adaptive behavior. More commonly, investigations have yielded a prevalence of 1 to 1.5 percent.

The broad consensus that exists on the classification criteria for the intellectual dimension does not exist in the special education rules adopted by the states. Current state criteria on a cutoff IQ score for the intellectual dimension vary from a low of about 69 to a high of 80

(Denning et al., 2000). Furthermore, Patrick and Reschly (1982) found that the stringency of the IQ criterion is not always related to the prevalence of students classified as having mental retardation and placed in special education. State-to-state variations in special education rules regarding conceptual definitions and classification criteria for mental retardation lead to large differences in prevalence and many inconsistencies in the diagnosis of mental retardation between such agencies as SSA and the public schools. The use of school data is discussed later in this chapter.

Classification Criteria for Adaptive Behavior

In most diagnostic systems, the classification criteria for adaptive behavior are not developed as well or as clearly as those for intellectual functioning. Two elements are particularly relevant: the degree of difference from normal or average performance that is required to determine that a limitation in adaptive functioning exists—that is, the cutoff score—and the number of domains or areas in which limitations may be observed. Each of these elements has a significant influence on the number of people who might be considered for a diagnosis of mental retardation.

Cutoff Scores

As noted earlier, there is far less agreement on the appropriate cutoff score(s) for adaptive behavior measures than there is for measures of intellectual functioning. Precise cutoff scores generally have not been specified in diagnostic systems, primarily because of the lack of confidence in adaptive behavior measures and the availability of multiple instruments that may be used interchangeably or somewhat idiosyncratically. A selection of quotations suggests the wide range of views:

- "If an adequate standardized instrument were available for the measurement of adaptive behavior, the upper limit of Level - I could presumably be set, as with the Measured Intelligence dimension, at greater

> than minus one Standard Deviation from the population mean" (Heber, 1961, p. 61).
>
> - "If more precise instruments were available for the measurement of Adaptive Behavior, and general norms could be precisely stipulated, the upper limit could presumably be set at minus two standard deviations from the population mean" (Grossman, 1973, p. 19).
>
> - "It seems impractical at this time to suggest fine gradations that can be achieved with accuracy, and, in the final analysis, clinical judgment is needed to arrive at an estimate of adaptive behavior level. . . . Standardized scales, supplemented by clinical judgments whenever possible, should be applied in making diagnoses" (Grossman, 1983, p. 46).
>
> - "Despite increased emphasis on adaptive skills in the definition, there has been virtually no support for the use of a single global score or age equivalent index to operationalize adaptive skill limitations. There are a number of reasons why a global score and precise cutoff point would not be productive" (American Association on Mental Retardation, 1992, p. 42).
>
> - "The second criterion for diagnosing a person as having mental retardation is that the individual have limitations in two or more adaptive skills. This part of the diagnosis is more substantive and subjective and requires clinical judgment that takes into account environmental demands and potential support systems" (American Association on Mental Retardation, 1992, p. 49).

The American Psychological Association Division 33 (Editorial Board, 1996) mental retardation diagnostic system is the one exception to the general trend in diagnostic systems of avoiding precise specification of adaptive behavior cutoff scores to define mental retardation eligibility. The Division 33 scheme is explicit in recommending the use of a "comprehensive, individual measure of adaptive behavior" (p. 13) and in specifying precise cutoff scores:

> For adaptive behavior measures, the criterion of significance is a summary index score that is two or more standard deviations below the mean for the appropriate norming sample or that is within the range of adaptive behavior associated with the IQ range sample in instrument norms (Editorial Board, 1996, p. 13).

In its next paragraph, the Division 33 discussion allows for part scores, but the criterion for eligibility is that "two or more of these scores lie two or more standard deviations below the mean" (p. 13). The Division 33 system also adopted the now rather standard criterion of intellectual functioning at two or more standard deviations below the mean. The consequences of these requirements on the number of people currently considered for a diagnosis of mental retardation or on the number of persons considered in the future have not been addressed. The results of previous studies had suggested that a stringent criterion for adaptive behavior plus the usual criterion for intellectual functioning led to a sharply reduced number of people eligible to be considered for a mental retardation diagnosis (Heflinger et al., 1987; Reschly, 1981a). However, these studies used a particular measure of adaptive behavior that had a very low correlation with measures of intellectual functioning.

The classification criteria governing diagnosis of mental retardation for special education services by state departments of education generally do not provide guidance regarding the use of adaptive behavior composites, part scores, or cutoff scores to determine eligibility. It is not surprising that the use of an adaptive functioning criterion in the schools is inconsistent and unpredictable (Reschly & Ward, 1991). Moreover, enormous variations exist across the states and, in some instances, across local school districts within states.

Adaptive Behavior Domains

Diagnostic systems are either silent on the appropriate number of adaptive behavior domains, or they adopt widely varying schemes. The most recent classification system of the American Association on Mental Retardation (AAMR) specifies 10 adaptive skills areas without any explanation of how that number was determined or why some domains were included and others excluded (American Association on Mental Retardation, 1992). In contrast, the American Psychological Association Division 33 diagnostic system is generally consistent with the factor analytic literature (see Chapter 4 and later discussion in this chapter) in specifying the use of a composite score, recognizing that a general adaptive behavior factor exists, and the use of a limited number of part scores.

The appropriate number of domains in the assessment of adaptive behavior depends on the instrument, age level, and other considerations. Some useful guidance on the number of meaningful domains is provided by factor analytic studies; however, different factor methods yield different results, so such studies are rarely definitive. Most adaptive behavior scales yield a general factor, regardless of the number of domain or subdomain scores, if the analytic method permits the emergence of such a factor (e.g., Harrison & Oakland, 2000a; McGrew & Bruininks, 1989). Typically, one or more group factors also emerge, particularly if (a) confirmatory factor analytic procedures are applied, (b) items reflect diverse areas of functioning, (c) sufficient floors and ceilings are provided, (d) broad age ranges are included, and (e) individuals from the moderate and mild levels of mental retardation, as well as people with borderline and normal levels of functioning, are included in the sample. Across all ages, McGrew and Bruininks suggested the possibility of four or five group factors. This literature is discussed in detail in Chapter 4.

Beyond its theoretical importance, the appropriate number of adaptive behavior domains has a very practical significance: it can have a tremendous influence on the number of people who may be diagnosed as having mental retardation and therefore are eligible for Supplemental Security Income (SSI) and Disability Income (DI) benefits. At the request of the committee, Thompson (2001) ran a series of Monte Carlo simulations to address this effect. She found that the number of adaptive behavior domains on which deficits must be shown had a marked effect on identification rates, with more individuals being identified as having mental retardation if only a single adaptive behavior domain had to meet a defined cutoff score than if two or more domains had to meet a cutoff score. Furthermore, the number of domains on which deficits could be measured had a modest but significant effect on identification rates: more individuals will be diagnosed as having mental retardation if deficits can be found in 1 or 2 out of 9 or 10 domains, than if deficits are found in 1 or 2 of only 4 domains of adaptive behavior. In other words, it is easier to qualify for a diagnosis of mental retardation if there are more domains in which deficits can be shown.

It is important to remember that expectations about adaptive behavior and competence vary by sociocultural group, settings, and age levels. The conceptual definition and descriptions of adaptive behavior in the 1983 AAMR manual (Grossman, 1983) have been particularly instructive in this regard because in it different competencies were associated with broad age ranges (e.g., preschool, childhood, adolescence, adult). The committee considered the interaction of age-based expectations and adaptive behavior domains, including the current domains identified in the SSA listings (see Table 5-2). Clearly, the current SSA scheme recognizes different domains at different age levels and is similar in most respects to the adaptive domains discussed in Chapter 4. In other respects, however, the SSA domains are inconsistent with findings from factor analytic results, the descriptions of adaptive behavior in authoritative sources, and the content of current adaptive behavior inventories. This led to the committee's recommendation, presented in Chapter 4, for revising the SSA adaptive behavior domains (see right side of Table 5-2).

The SSA domain of concentration, persistence, and pace is not assessed by most adaptive behavior instruments, although these skills could be part of the work attitudes and skills domain recommended for adults. The current SSA domains do not include some that are prominent in current conceptions and measures of adaptive behaviors, particularly the self-help and communication domains in the preschool years. For childhood and adolescence, the SSA scheme does not include motor/mobility, communication/functional academics, or daily living skills, although the latter may be covered by SSA in the domain of personal functioning. And, the communication/functional academics and work attitudes and skills domains are missing from the SSA adult domains. The adaptive behavior areas specified in Table 5-2 are

TABLE 5-2 Adaptive Behavior Domains in Current Social Security Administration Regulations and Committee Recommendations

Age	Social Security Administration Domains	Age	Committee Recommendations
Birth-2	1. Motor (fine/gross) 2. Social	Birth-4	1. Motor 2. Social 3. Self-help 4. Communication
3-17	1. Personal 2. Social 3. Concentration/ persistence/pace	5-17	1. Motor/mobility 2. Social 3. Communication/ functional academics 4 Daily living skills
18+	1. Daily living 2. Social 3. Concentration/ persistence/pace	18+	1. Motor/mobility 2. Social 3. Communication/ practical cognitive skills 4. Daily living skills 5. Work

the ones recommended by the committee to be adopted by SSA to guide decisions about diagnoses of mental retardation (the recommendation itself appears in Chapter 4).

Most current adaptive behavior measures have domains that are similar to the domains recommended by the committee; however, no scale is perfectly matched to these domains. Moreover, essential content, such as functional academic skills involving basic literacy, temporal relationships, and quantitative concepts, appears on most scales, but in different domains. For example, the area of functional academic skills is a separate domain in the Adaptive Behavior Assessment Scales (ABAS—Harrison & Oakland, 2000a) and the Comprehensive Test of Adaptive Behavior (CTAB—Adams, 2000), but it is spread over at least two domains in the Vineland Adaptive Behavior Scales (VABS—Sparrow et al., 1984a) and the Scales of Independent Behavior-Revised (SIB-R—Bruininks et al., 1996). Valuable information from each of these instruments on the functional academic skills area is available, but direct translation of the available scores to a decision about performance in this area is difficult. Similar relationships exist between the available scores from instruments and the recommended areas in Table 5-2. Decisions about performance in each of these areas therefore need to be based on the results of adaptive behavior instruments, to the extent that one or more instruments are appropriate for a given client, and a broad variety of other information.

RELATIONSHIP OF ADAPTIVE BEHAVIOR AND INTELLECTUAL FUNCTIONING

The relationship between measures of adaptive behavior and intellectual functioning is highly variable and has multiple influences. Correlations between adaptive and intellectual functioning have varied in published studies from near zero (no relationship) to nearly 1.0 (perfect relationship). The strength of this relationship is important because it influences diagnostic decisions significantly.

Variables That Influence Correlations

A comprehensive review by Meyers et al. (1979) summarized data on the correlations of measures of adaptive and intellectual functioning. This section is informed by that review as well as by data published in test manuals over the past 20 years. Newer adaptive behavior scales generally conform to the generalizations made by Meyers et al. in 1979.

Scale Content

A major influence on the relationship of adaptive and intellectual functioning is the content of the measures, particularly the adaptive behavior measure. Very low correlations are obtained with adaptive behavior scales that do not include content on practical, everyday cognitive skills. For example, scores on the Adaptive Behavior Inventory for Children (ABIC) (Mercer & Lewis, 1978) have an extremely low correlation with scores on the Wechsler Intelligence Scale for Children-Revised (WISC-R), according to the manual and other studies (Kazimour & Reschly, 1981). The effect of using the ABIC in mental retardation classification decisions was to eliminate the mild level (IQ 55 to 70 or 75) because virtually no one with an IQ in that range had an adaptive behavior score that was more than –1 or –2 SD below the mean (Heflinger et al., 1987; Reschly, 1981a).

The composite and domain scores on adaptive behavior measures that include practical, everyday cognitive skills show higher relationships with intellectual functioning; for example, the correlations with intellectual measures for the communication domain on the VABS is about 0.4 (Sparrow et al., 1984a), and the functional academic skills area of the ABAS (Harrison & Oakland, 2000a) is about 0.5. In contrast, the VABS domains of daily living skills, socialization, and motor skills have correlations with IQ that vary from about .20 to about .35. The relatively low correlations that many adaptive behavior measures have with IQ tests mean that many individuals low on one of the measures may not be low on the other, an outcome that has important implications for mental retardation diagnoses. Correlations may also be low because of ceilings on adaptive behavior measures or because of the attenuation of the correlation between adaptive behavior and intelligence among people with IQs above 100.

Competencies Versus Perceptions

Adaptive behavior measures typically reflect the judgment of a respondent about a client's performance, which also influences correlations of adaptive behavior measures with IQ scores. Generally, adaptive behavior items that do not reference specific behaviors are less likely to correlate highly with IQ test results. For example, an item that asks, "Does Egbert read common signs?" with response choices such as "Often, Sometimes, Never" is less likely to yield precise information on skills than items that reference specific behaviors, like "Can Esmerelda read a rest room sign and act accordingly?" The latter focuses on a "can do" skill rather than a respondent's perception of the client's participation in some activity. Scores on scales with more behaviorally specific skills generally have higher correlations with measures of intellectual functioning (Adams, 2000; Harrison & Oakland, 2000a).

Sample Variance

All correlational studies depend on sample variability, and the relationship of adaptive and intellectual functioning measures is no exception. Research studies that include participants who score very differently on both measures will generally produce higher correlations, other things being equal. Constraints on sample variability in a study ensures lower correlations.

Ceiling and Floor Problems

Most adaptive behavior measures have an insufficient number of items at the highest levels of performance for clients or study participants; this is called a ceiling problem. Insufficient numbers of items at the lower levels of performance create floor problems. Both intellectual and adaptive behavior measures often have ceiling and floor problems, most often at the extremes of ability or near the bottom and top ages covered by the instrument (see Chapters 3 and 4 for a more complete discussion of these issues). Generally, for individuals with low intellectual functioning, problems with intellectual measures are likely to be insufficient floors. For adolescents and adults with IQs from 60 to 85, or who have intellectual skills in the normal range, the opposite problem occurs with adaptive behavior measures, which usually do not have sufficiently high ceilings. Absence of a sufficient number of ceiling items produces very low IQ by adaptive behavior correlations for adults on most measures, although the ABAS (Harrison & Oakland, 2000a) suffers less from this problem than do other adaptive behavior scales.

Level of Mental Retardation

In their review of research, Meyers et al. (1979) found that for young children functioning far below average, the results of intellectual functioning and adaptive behavior measures were nearly identical. In fact, for extremely low levels of functioning on both types of test, nearly the same items are used on measures of adaptive and intellectual functioning, providing a ready explanation for the nearly perfect relationship. It is reasonably safe to say that at the severe and profound levels of mental retardation (Grossman, 1983), particularly with young children, little difference exists between the constructs of adaptive and intellectual functioning.

The relationship of adaptive and intellectual functioning measures is less certain with individuals functioning in the moderate and mild levels of mental retardation as defined by IQ. In the VABS manual (Sparrow et al., 1984a), correlations between various intellectual measures and VABS domain and composite scores are reported for adults with mental retardation who were in residential and nonresidential living arrangements. The correlations for the VABS composite score and the Wechsler or Stanford-Binet intellectual measures were generally slightly higher (r = .4 to .5) than similar correlations for normal children (r = .32). The samples of adults had mean IQs that varied from about 25 to 50, suggesting that these VABS correlations were derived from samples of people with moderate to severe mental retardation. Few persons in these samples were in the range of mild mental retardation, the ability range that is most relevant to the committee's deliberations.

Harrison and Oakland (2000a) reported the relationship of the ABAS composite to Stanford-Binet Fourth Edition IQs for a "clinical" sample composed of children with autism (15 percent of the sample) and mental retardation (85 percent of the sample). The mean IQ for this sample was about 51, suggesting that most of the participants had intellectual ability scores indicating the moderate level of mental retardation. This ABAS correlation with IQ is virtually the same for people without mental retardation and for those with autism or mental retardation, but, again, this correlation does not bear directly on the relationship of adaptive behavior and intellectual functioning for people in the critical IQ range of 60 to 75.

The preponderance of the evidence suggests that the correlations between measures of IQ and adaptive behavior are higher for individuals with severe or profound levels of mental retardation. The lower the IQ level within the range from about 20 to 40, the higher the correlation to adaptive behavior scores. The magnitude of this relationship for people with the moderate level of mental retardation is less certain, but it appears to be slightly higher than for people without this disability. Sufficient evidence is not available regarding the intellectual functioning and adaptive behavior relationship for people in the mild range of mental retardation, although extrapolating the findings just cited for the profound, severe, and moderate levels leads to the conclusion that the relationship is either no higher or only very slightly higher than for people without cognitive disabilities. It also leads us to conclude that the best and most accurate guide to the IQ and adaptive behavior relationship in the IQ 60 to 75 range is the correlation for people without cognitive disabilities, unless there are other specific results for people with mild mental retardation. None of the recently published adaptive measures has supplied information for this particular population.

It should be noted that high correlations of IQ and adaptive behavior have been reported for one adaptive behavior measure and that they may be spurious. Bruininks et al. (1996) reported that for children ages 5 to 12, the correlation between scores for adaptive behavior on the SIB-R and IQ was .78, but only .20 for adolescents and adults ages 13 to 90. The overall correlation between these two measures across all ages, 3 to 90, was .82. These unusually high correlations are probably due to contamination of the correlations with changes in developmental level. For that reason, the SIB-R correlations reported in the manual may not be accurate estimates of the relationship of the SIB-R scores and intellectual functioning.

Classification Agreement: IQ and Adaptive Behavior

A critical issue related directly to the committee's recommendation on combining IQ and adaptive behavior data is the classification agreement between measures of intellectual functioning and adaptive behavior. Classification agreement studies form part of the basis for determining appropriate cutoff scores on adaptive and intellectual functioning measures. The ideal data set would be large samples of people functioning in the range of borderline to the high end of moderate mental retardation. Diagnosis of mental retardation could be studied using different cutoff scores (–1 SD, –1.66 SD and –2 SD) on both measures simultaneously. The results would be informative about the implications of establishing different assessment requirements and the application of different eligibility criteria. However, actual data of this kind generally are not available; the committee therefore commissioned two sets of Monte Carlo simulations to assist in our analysis.

Classification Agreement: Existing Data

Few data exist on the classification agreement between recently published adaptive behavior measures and intelligence tests. As discussed at length in Chapter 4, only one test manual contains data that are relevant to this issue (Harrison & Oakland, 2000a, p. 89, Table 5.31). The data in Table 4-3 show that nearly 25 percent of the people with a preexisting diagnosis of mild mental retardation did not score 2 or more SDs below the mean on two adaptive behavior domains, demonstrating that this is not an appropriate criterion for a diagnosis of mental retardation.

These results should not be generalized to other adaptive measures. It should be pointed out that it is easier to meet a limitations criterion if, as in this case, more adaptive behavior subareas are used.

That is, 1 or 2 deficits out or 9 or 10 subareas is an easier criterion to meet than 1 or 2 deficits out of 4 or 5 subareas.

Although the committee searched, no further studies of the classification agreement between recently published adaptive behavior measures and IQ were found. The authors of other adaptive behavior measures have not provided data of this nature, even though the current and previous editions of the Joint Association Test Standards (Joint Committee on Standards for Educational and Psychological Testing of the American Educational Research Association, the American Psychological Association, and the National Council on Measurement in Education, 1985, 1999) require that test publishers provide validity data supporting the recommended uses of a test.

Simulations of IQ and Adaptive Behavior Classification Agreement

Because of the dearth of data examining the relationship between various levels of IQ and scores on adaptive behavior measures, particularly for individuals in the mild mental retardation range, the committee used Monte Carlo models to project the proportion of people who would be expected to have IQ and adaptive behavior domain scores in specific ranges. These Monte Carlo models used correlations of adaptive behavior domain scores with one another and with IQ (Thompson, 2001). The simulations used the best of the currently available adaptive behavior measures, applying assumptions that enhanced the likelihood of classification agreement between IQ and adaptive behavior. For example, high-end estimates of the relationship between adaptive and intellectual functioning were applied as well as internal consistency reliability estimates, rather than stability coefficients, making classification agreement slightly more likely. Projections of the proportions of adult cases meeting IQ and adaptive behavior cutoff scores were developed from the Monte Carlo analyses and are summarized in Tables 5-3 and 5-4. These projections were based on test manual data on the correlations of domain scores with one another and with IQ. The Monte Carlo results answer the question: Given an IQ in the range of mental retardation, how many cases will

TABLE 5-3 Monte Carlo Simulations of the Proportions of Children and Adults Eligible to be Considered for a Mental Retardation Diagnosis Using the VABS and SIB-R at Varying IQ and Adaptive Behavior Cutoff Scores

Adaptive Behavior[a]		IQ Cutoff Score		
No. of Deficits	Cutoff	IQ[b] <60	IQ[c] 60-70	IQ[d] 60-75
1	≤ 70	28-46	20-30	18-26
1	≤ 77.5	51-71	41-55	37-49
1	≤ 85	72-88	64-78	60-74
2	≤ 70	9-22	6-12	5-9
2	≤ 77.5	23-45	17-30	15-25
2	≤ 85	46-71	38-55	35-50

NOTE: VABS = Vineland Adaptive Behavior Scales. SIB-R = Scales of Independent Behavior-Revised.

[a] Adaptive behavior status refers to any one or any two of four part scores from the VABS or SIB-R.

[b] All persons with IQs below 60 are presumed eligible to be considered for a diagnosis of mental retardation according to current Social Security Administration criteria. The proportions in this column show the proportions of individuals with IQs below 60 who also have adaptive behavior scores below different cutoff scores.

[c] This column provides data on individuals with IQs of 60 through 70. Persons with IQs below 60 or above 70 are not included.

[d] This column presents data on individuals with IQs from 60 through 75. The proportions reflecting classification agreement are lower in column five than four because a smaller proportion of cases in the 71-75 IQ interval have low adaptive behavior scores compared with cases in the 60-70 interval.

have adaptive behavior scores that meet the different cutoff scores? For example, looking at the results in the first row, third column, it appears that for people with IQs of less than 60, between 28 and 46 percent meet the standard of at least one adaptive behavior part score at or below the cutoff of 70.

In Table 5-3, findings are presented jointly for two adaptive behavior scales that are well standardized and supply a composite score and four part scores, the VABS (Sparrow et al., 1984a) and the SIB-R

TABLE 5-4 Monte Carlo Simulations of the Proportions of Children and Adults Eligible to Be Considered for a Mental Retardation Diagnosis Using the ABAS at Varying IQ and Adaptive Behavior Cutoff Scores

Adaptive Behavior[a]		IQ Cutoff Score		
No. of Deficits	Cutoff	IQ[b] <60	IQ[c] 60-70	IQ[d] 60-75
1	≤ 70	46-57	34-40	30-35
1	≤ 77.5	70-80	59-66	53-60
1	≤ 85	87-93	79-86	75-82
2	≤ 70	31-37	21-23	18-20
2	≤ 77.5	56-64	43-48	38-43
2	≤ 85	78-85	67-74	63-69

Note: ABAS = Adaptive Behavior Assessment Scales.

[a] Adaptive behavior status refers to any 1 or any 2 of the 9 ABAS adaptive skills areas for children or 10 adaptive skills areas for adults.

[b] All persons with IQs below 60 are presumed eligible to be considered for a diagnosis of mental retardation according to current Social Security Administration criteria. The proportions in this column show the proportions of individuals with IQs below 60 who also have adaptive behavior scores below different cutoff scores.

[c] This column provides data on individuals with IQs of 60 through 70. Persons with IQs below 60 or above 70 are not included.

[d] This column presents data on individuals with IQs from 60 through 75. The proportions reflecting classification agreement are lower in column five than column four because a smaller proportion of cases in the 71-75 IQ interval have low adaptive behavior scores compared with cases in the 60-70 interval.

(Bruininks et al., 1996). The part scores conform generally to the adaptive behavior domains recommended in Table 5-2 and to the factor analytic results discussed in Chapter 4. The four VABS part scores included in the Monte Carlo analyses were motor, independent living, communication, and social. The SIB-R part scores were motor skills, personal living skills, community living skills, and social interaction/communication skills. Examination of the proportions of cases in the simulations with various combinations of low IQ and low adaptive behavior scores provides an estimate of the effects of altering the cutoff scores for both measures.

Six adaptive cutoff scores were used in one domain at –2.0 SD (≤ 70), –1.5 SD (≤ 77.5), and –1.0 SD (≤ 85), and in two domains at the same cutoffs of -2.0 SD, -1.5 SD, and –1.0 SD (Table 5-3, columns 1 and 2). Classification agreement with three IQ criteria was simulated, IQ < 60, IQ ≥ 60 but ≤ 70, and IQ ≥ 60 but ≤ 75 (Table 5-3, columns 3-5). Proportions of cases in the simulations that met various combinations of IQ and adaptive behavior cutoff scores are shown in the table.

The results in column 3 of Table 5-3 indicate that many people with IQs of less than 60 do not meet the cutoff score requirements for a deficit in adaptive behavior. As noted earlier, SSA defines all persons with an IQ of less than 60 as presumptively eligible to be considered for a diagnosis of mental retardation, and the committee recommends continuing that practice. The imperfect classification agreement between IQ and adaptive behavior at low IQ levels indicates that caution must be used in proposing the adoption of adaptive behavior cutoff scores. Consider individuals with IQs of less than 60 and the most lenient adaptive behavior cutoff score, deficits in one domain at or below 85 (see row three): only 72 to 88 percent of cases met the dual IQ and adaptive behavior criteria. Adoption of any adaptive behavior cutoff score even with persons presumptively eligible due to IQs below 60 could result in 12 to 18 percent of those currently eligible to be considered for a diagnosis of mental retardation becoming ineligible.

Examination of other combinations of IQ and adaptive behavior scores further supports the use of caution in setting an adaptive behavior cutoff score. In column 4 of Table 5-3, an IQ between 60 and 70 and a stringent criterion of two adaptive behavior domains at or below 70 yields a hit rate of only 6 to 12 percent, a level that is far below the stipulation that most people with IQs at or below 70 should be eligible to be considered for a diagnosis of mental retardation. Even the most lenient criterion included in the simulation, a deficit in one adaptive behavior area at a cutoff at or below 85, resulted in only 64 to 78 percent of cases meeting the dual criteria of an IQ score between 60 and 70 and an adaptive behavior score at or below 85. Other combinations also are instructive. At the criterion of two adaptive behavior domains at or below 85, only 38 to 55 percent of persons with IQs between 60 and 70 met the dual classification criteria (see last row in column 4).

The two adaptive behavior inventories differed regarding the rate of classification agreement with IQ in the Monte Carlo simulations in Table 5-3. Higher classification agreement occurred with the SIB-R children simulation and the lowest with the VABS children simulation. Using these scales, the classification agreement for adults probably should not be done because both scales have ceiling problems for individuals with mild mental retardation or borderline functioning who are in the late adolescent or adult years. Finally, it bears repeating that these simulations adopted assumptions that probably enhanced the degree of classification agreement and therefore are likely to overestimate the degree of classification agreement in actual practice.

Thompson (2001) conducted additional Monte Carlo simulations with the ABAS (Harrison & Oakland, 2000a). The combined results for children and adults are presented in Table 5-4. The number of ABAS adaptive skills areas is 9 for children and 10 for adults. It is easier to achieve classification agreement when more adaptive skills areas are included. Generally, it should be expected that the proportions indicating classification agreement will be higher in Table 5-4 than Table 5-3 due to the greater number of adaptive skills areas (9 or 10 areas in Table 5-4 versus 4 areas in Table 5-3) and the higher correlations for the ABAS than the VABS or the SIB-R. Nevertheless, a significant number of cases in the Monte Carlo simulation with IQs in the range of mental retardation do not have significant adaptive behavior deficits using the ABAS.

The uncertainty regarding the effects of different adaptive behavior cutoff scores is further increased by comparing the results in Tables 4-3 and 5-4. Both tables use data from the ABAS. In Table 4-3, real data are reported for a “convenience sample” of children and youth with mild mental retardation whose characteristics were not fully described. For that group, the authors reported that 76 percent had two or more adaptive skills area scores at or below 70. The ABAS Monte Carlo analyses using the cutoff of adaptive behavior at or below 70 revealed far lower proportions meeting the criterion of two or more adaptive behavior areas in the range of mental retardation. Having no way to reconcile these inconsistent results regarding the ABAS, the committee had to take the uncertainty caused by these results into consideration when recommending an adaptive behavior cutoff score.

Conclusions

Caution in the adoption of precise adaptive behavior cutoff scores is warranted by the limited evidence on the classification agreement between IQ and adaptive behavior measures at varying cutoff scores. Monte Carlo simulations, conducted to estimate the probable effects of varying adaptive cutoff scores, yielded results indicating the classification agreement often was rather low using the best of the currently available adaptive behavior measures. The inconsistencies between ABAS actual data, albeit from a limited sample, and the ABAS simulations provide further support for caution in recommending precise cutoff scores.

It is not possible to simulate score distributions for people who would actually be referred for benefits; such a distribution may be somewhat different and possibly reflect more pronounced functional limitations. It may also be the case that a simulation, as in the current instance, must be based on parameters for all people in a norming sample for whom information is available, rather than on people with a more restricted IQ range. Such considerations will affect the degree to which the simulation reflects actual circumstances and the functional characteristics of cases reviewed for benefits eligibility.

The committee's formal recommendation about the use of adaptive behavior scales, presented in Chapter 4, is predicated on the dual goals of providing reasonable guidance for decision making and avoiding the massive declassification effects of an excessively stringent cutoff. The cutoff scores recommended represent the committee's long deliberations about the relative effects or more or less stringent criteria. In the committee's judgment, the same cutoff score could not and should not be used for IQ and adaptive behavior due to the near certainty of massive declassification effects. We explicitly rejected the use of a –2 SD adaptive behavior cutoff score. The committee was, however, reluctant to allow an arbitrary adaptive behavior cutoff score to be used or to adopt the most lenient of the various scores that were investigated.

In order to fulfill our charge of providing more guidance to SSA regarding adaptive behavior and our obligation to use the best available science in making recommendations, the committee decided to recommend some discretion regarding the interpretation of the results of formal measures of adaptive behavior. That is, a formal assessment of adaptive behavior should be provided in all cases. However, a person may be diagnosed as having mental retardation even if the adaptive behavior results do not meet the cutoff criteria if there is compelling evidence of adaptive behavior deficits that significantly impair performance of expected behaviors. It is the obligation of those gathering and interpreting adaptive behavior information to make a compelling case if it is warranted.

SSI AND DI ELIGIBILITY DECISIONS

Only professionals who have extensive knowledge about mental retardation and its assessment and who are using extensive information about an individual from multiple sources and settings should formulate eligibility decisions, because they have such profound consequences for the lives of clients and their families. Since there are four possible outcomes in any diagnostic setting and two of them are errors (Coombs et al., 1970; Swets et al., 2000), diagnoses must be made extremely carefully. A *true positive* or “hit” occurs when mental retardation is present and the diagnosis is correct. Similarly, a *true negative* or “correct rejection” occurs when mental retardation is not present and the diagnosis indicates that. Anything else is an error: a diagnosis of mental retardation when it is actually absent is a *false positive* or “false alarm.” Finally, a diagnosis of no mental retardation when it is actually present is a *false negative* or “miss.”

If the diagnostic decision is incorrect, whether it confirms or denies eligibility, it is likely to be related to negative consequences for individuals, their families, and society. Incorrectly confirming eligibility results is a waste of scarce public resources and may diminish the client's motivation and opportunities to engage in normal activities of work and economic self-support. Incorrectly denying eligibility has equally harsh consequences for individuals and families involving, in extreme cases, lack of access to the basic necessities of life. Decisions with consequences of this magnitude must be made by knowledgeable persons using the best information available and applying a principle of convergent validity. They must also understand the types of errors that are likely to occur in situations as complex and challenging as determination of mental retardation.

Common Judgment Errors

A *confirmation bias* can occur when a decision maker seeks information that confirms an already existing hypothesis or judgment (Evans, 1989). The best course of action when examining a hypothesis is to seek evidence that tests the idea by seeking disconfirming evidence. However, "it is a common observation in psychological research that individuals tend to selectively search for evidence to support their views at the expense of seeking contrary evidence" (Bunn, 1992, p. 253).

This tendency to look for confirming, rather than disconfirming, evidence has clear implications for eligibility decisions. If examiners focus primarily on seeking supporting evidence, then errors are almost inevitable. Furthermore, examiners are unlikely to know that they have made errors or why, since the evidence they have gathered will tend to support their decisions. Thus, the confirmation bias can lead to a pattern of self-perpetuating errors. Active countermeasures are needed to reduce the impact of judgment biases, like seeking disconfirming evidence in a systematic fashion or using a structured examination process instead of an unstructured one.

In many judgment situations, initial impressions have been found to heavily influence the final decision. In a classic paper, Asch (1946) described the *primacy effect,* in which later evidence is interpreted in the light of earlier evidence; the early evidence actually causes a change of meaning in the later evidence. Stewart (1965) was one of the first to propose attention decrement as an explanation for primacy. He reasoned that changing a task so that subjects were forced to attend to later evidence should diminish the primacy effect. He found this to be true, but only when subjects responded at the end of the task sequence. If responses are made to each new piece of evidence as it becomes available, the effect is reversed; evidence appearing later in the sequence actually had a greater impact—a *recency effect*. Many studies have shown that primacy effects can be reversed to recency effects using a variety of attentional manipulations. When left to our own resources, however, we tend to emphasize whatever we know first about a person—a primacy effect. In the context of eligibility determination, the same information can lead to different impressions depending on the selection and order of the assessments. Thus, the order of the impression formation can be significant in determining the outcome of the process.

Comprehensive Evaluation

Eligibility decisions must be based on a comprehensive evaluation of the person and the environment in which he or she lives and works. Opportunities as well as demands in that environment must be considered along with the client's status on broad domains of human functioning. Physical, emotional, adaptive, and cognitive functioning must be considered, and current SSA guidelines are consistent with this principle.

The tendency of clinicians engaged in evaluations of mental retardation to focus almost exclusively on intellectual and adaptive functioning, ignoring other important information on overall mental health and emotional adjustment, has been called "diagnostic overshadowing" (Reiss et al., 1982; Reiss & Szyszko, 1983). Information about low intelligence and poor adaptive functioning has been known to overshadow other clinical information that suggests the presence of other mental or emotional disorders that meet formal diagnostic criteria. A number of studies conducted since 1980 confirms that type of bias on the part of the clinician and the frequent existence of psychopathology accompanying mental retardation (e.g., White et al., 1995). Many persons with mental retardation are eligible for dual diagnoses, a simultaneous diagnosis of mental retardation and some other physical or mental disorder, such as depression, conduct disorder, or sensory impairment. It is crucial that clinicians evaluating individuals for the diagnosis of mental retardation also look for other emotional, mental, and physical disorders that may complicate adaptive and intellectual functioning and confer eligibility in other SSA categories (see Chapter 6).

Consideration of Other Information on Intellectual and Adaptive Functioning

There is clear consensus in the human services professions that a broad variety of information must be collected and evaluated regarding the individual in addition to and, in many cases, independent of the results of standardized tests and inventories. No single test or inventory score should be the sole basis for a significant decision. Moreover, for some clients, formal instruments are inappropriate due to a poor match between the client's characteristics and the nature of the test or inventory requirements. None of the authors of major testing instruments claims that other information is irrelevant or that information from other sources confirming or disconfirming the results of a standardized instrument should be ignored. It is useful to consider different methods of data collection, different sources of information, and performance in different settings.

Methods for Collecting Data

Four methods of collecting assessment information have been described in the literature: direct testing of clients, observation of behavior in one or more settings, review of records reflecting previous and current performance, and interviews with the client and knowledgeable others. Disability determination examiners should be familiar with and utilize each of the methods in developing a well-informed perspective on the functioning levels of clients considered for a mental retardation diagnosis.

Direct testing of clients' adaptive behavior is a rarely used technique, even though direct testing of general intellectual functioning is the norm. Yet, at least some direct testing is appropriate with adaptive functioning. For example, functional academic skills, such as basic literacy, understanding temporal relationships, and quantitative concepts, are crucial to adaptive functioning for children, adolescents, and adults. Adults who cannot tell time or meet time-related work obligations are at a significant disadvantage in coping with everyday demands. Information from third-party respondents on these skills may or may not be accurate, especially as these skills relate to everyday functioning. Some adaptive behavior measures suggest establishing conditions under which behaviors can be "tested" if third-party respondents are unable to report their actual observations of the behavior (Adams, 2000).

Interviews with third-party respondents by using standardized adaptive inventories is the most common method for collecting adaptive behavior information. For children, the third party is most often a parent or a teacher. While third-party interview is not used extensively in intellectual assessment, the results of IQ tests should be further evaluated through interviews with the client and significant others to determine if the observed test performance is consistent with day-to-day functioning. In addition to third-party respondents, interviews with clients and other parties are components of a comprehensive adaptive behavior assessment. One adaptive behavior instrument has been normed with adults using a self-report format; however, no data were provided in the manual contrasting the self-report and third-party respondent results for persons with mild mental retardation (Harrison & Oakland, 2000a). The accuracy of individuals with mild mental retardation in reporting their own adaptive behavior on this instrument is not yet clear; nevertheless, unstructured and structured interviews with clients are necessary for determining if adaptive behavior deficits exist.

Observations of clients' intellectual functioning and adaptive behaviors can be systematic and structured or informal and anecdotal. Both techniques yield valuable information for understanding overall functioning. Opportunities for systematic observation of clients' adaptive behaviors are limited by resource constraints and the near impossibility of conducting highly structured observations in all relevant settings and at the times that are appropriate. Moreover, many adaptive behaviors that are crucial to adequate functioning do not occur frequently, making systematic observation even more difficult. Therefore, even informal and anecdotal observations from different people and across different settings are valuable to an overall decision about adaptive behavior and should be obtained to the extent feasible.

Review of records is another data collection method with strong applicability to the determination of intellectual and adaptive behavior deficits. School records are especially useful if evaluators understand the nature of mental retardation, classification practices in schools, and subtle indicators of low functioning in classrooms and schools. However, school records indicating either a diagnosis of mental retardation or the absence of one cannot be used as a definitive indication of intellectual and adaptive behavior status. Records from agencies other than schools can also be useful in determining adaptive behavior deficits. Medical, social service, and legal sources may yield further information that is useful in making judgments about deficits. Further discussion of the use of records from schools and other agencies in order to make diagnostic decisions appears below.

Sources of Information

Judgments about intellectual and adaptive functioning should be based on multiple sources of information including, at a minimum, the individual client and significant others such as (depending on age) parents, teachers, peers, neighbors, and family members. The kind and amount of information gathered from different people will vary significantly across clients. In some cases, judgments must be based primarily on an interview with a single third-party respondent and on observations or interviews with the client, while in other cases there will be multiple sources of information. SSA disability examiners have to make judgments about the sufficiency of the information in deciding whether to actively seek additional sources of information.

Settings

The client's functioning across different settings is also relevant to decisions about intellectual and adaptive behavior deficits. The settings that are most relevant depend on the client's developmental level. For preschoolers, the relevant setting is the home and, depending on the client, day care or preschool settings. For children between ages of about 5 and 18, the school and home settings are crucial for nearly all clients, as are skills in meeting expectations as they age for roles in the neighborhood and the community. Deficits that are apparent only in a single setting generally should not be the basis for a determination of an adaptive behavior deficit. Diligence in collecting and examining information from multiple settings is very important.

The literature refers to so-called six-hour retarded children (President's Committee on Mental Retardation, 1970), described as having performance deficits only in school settings and coping adequately in home and community settings. "Six-hour retarded children" were assumed by many to blend into the normal adult population without significant adaptive limitations. Studies of young adults who clearly met this conception yielded a very different picture (Koegel & Edgerton, 1984). Contrary to the assumptions, as young adults, these children had enormous difficulties in coping with everyday demands and avoiding being exploited by others due to their functional limitations in practical cognitive and other adaptive skills.

Current classification and placement practices make it less likely that children will be identified in this way in schools (see later discussion), but many of them do have significant problems coping with the everyday demands involving social relations with peers and negotiating the cognitive demands required for personal and social adaptation outside school. Careful analysis of how children and adolescents perform in school, home, and other settings must be made in order to come to the most accurate diagnosis.

Examiner Qualifications

Individuals making mental retardation diagnoses must meet high standards regarding professional preparation and relevant experience. SSA disability examiners should have in-depth and up-to-date knowledge in the following areas: mental retardation theory, research, treatment, and best practices; mental retardation diagnostic construct; measurement of intellectual and adaptive functioning; assessment principles and best practices; mental disorders theory, research, and best practices; purposes and practices of multiple agencies, such as schools, law enforcement, and health care; knowledge of human development; and assessment of the individual's strengths and limitations in the context of multiple environments, including family, work, and community.

Convergent Validity

Convergent validity is an application of the concept of the multitrait/multimethod examination of the validity of measures of psychological constructs (Campbell & Fiske, 1959). The committee recommends the principle of convergent validity as a means for SSA examiners to make sense of all the information evaluated for diagnostic decisions about mental retardation. In clinical practice, information is collected and evaluated over broad domains of functioning, using multiple methods of gathering data, multiple sources for that information, and multiple settings (e.g., Gresham, 1991). If the information is generally consistent with a particular diagnostic decision, such as mental retardation, that decision is made based on this confirmation. If, however, there are several elements in the overall array of information that are inconsistent with such a diagnosis, the decision is not confirmed or—and this is important—further investigation is undertaken to explain discrepancies.

Inconsistent Information

Many individuals with legitimate diagnoses of mental retardation will present clinically with one or more elements of information that are inconsistent with the diagnosis. Many times, on further examination, the inconsistent information will derive from a mistake or distortion of measures of adaptive or intellectual functioning—for example, the existence of high scores from an adaptive behavior inventory with little or no ceiling for a young adult. In other cases, a particular respondent or performance in a specific setting may not be consistent with the diagnosis of mental retardation. Inconsistent information must be investigated thoroughly in order to avoid the harmful consequences of false positive or false negative decisions. Further investigation may take many different forms, including more extensive interviews with the client and significant others, additional assessments of adaptive or intellectual functioning, or follow-up contacts with personnel in other agencies who may or may not have made a diagnosis of mental retardation.

Information from Other Agencies

As noted previously in this report, mental retardation diagnoses are made for many purposes by many different agencies, including schools, law enforcement, and health care agencies. Different agencies use different diagnostic systems and classification criteria, making diagnostic disagreements among agencies perfectly legitimate. Moreover, the standards used by some agencies, public schools in particular, differ markedly from state to state and, occasionally, from district to district within states.

Several sources of information confirm the increasing reluctance of school officials to make a diagnosis of mental retardation. First, the

TABLE 5-5 Changes in Prevalence of Learning Disabilities and Mental Retardation over the 1976-1999 Period in Public School Special Education Programs

	1976-77	1998-99	Change	% Change
Learning disabilities	797,213	2,861,333	2,064,120	260%
Mental retardation	969,547	613,207	–356,340	–36%

NOTE: All data are from the 1978 and 2000 Annual Reports to Congress by the Office of Special Education Programs, U.S. Department of Education.

prevalence of mental retardation in schools has declined substantially over the last 25 years. It is extremely unlikely that all of the decline is attributable to a truly lower prevalence of mild mental retardation. The decline in mild mental retardation is paralleled by a corresponding and substantially greater increase in learning disabilities (see Table 5-5). According to the Office of Special Education Programs child count data (U.S. Department of Education, 2000), prevalence of learning disabilities has increased by 260 percent while mental retardation prevalence has declined by 37 percent since the 1976-1977 academic year, when these data were first collected. These changes are even more impressive because they occurred during a period when children and youth with moderate, severe, and profound mental retardation gained access to the public schools for the first time in many states and districts. Although the child count data do not differentiate levels of mental retardation, it is highly likely that the decline in mild mental retardation has been even greater than the overall decline in mental retardation, simply because those with more severe mental retardation are more obviously impaired in many areas and are therefore more likely to be correctly diagnosed with mental retardation.

The declining prevalence of mental retardation in the public schools is even more complex because it varies significantly across the states. The mental retardation prevalence among states varied by a factor of three in the 1998-1999 data; that is, the highest prevalence reported by any state was three times higher than the lowest state-reported prevalence. Moreover, 28 states reported public school prevalence of less than 1 percent of the student population, meaning that some students with mild mental retardation are either not being placed in special education or are in special education because of other diagnoses, such as learning disability.

Studies in California confirm both the reluctance of school professionals to confer the diagnosis of mental retardation and the willingness to use other diagnoses for children with characteristics that meet mental retardation classification criteria (MacMillan, Gresham et al., 1996). This leads to concerns about the continued viability of the mild mental retardation diagnostic construct in special education and schools (MacMillan & Reschly, 1996; MacMillan, Siperstein, & Gresham, 1996). The California studies indicate that staffing teams simply refused to diagnose students as having mild mental retardation even when IQ, achievement, and adaptive behavior data clearly pointed to that diagnosis. The degree to which these results generalize to other states is unknown, although the large decline in children diagnosed as having mild mental retardation suggests that the reluctance among California school psychologists and special educators to use this diagnosis may exist in many other places as well.

The presence or absence of a diagnosis of mental retardation from another agency, especially from public schools, should neither confirm nor disconfirm an SSA diagnosis of mental retardation. Information from other agencies should be evaluated by SSA examiners, but should not be regarded as definitive. Although the official diagnoses used in other agencies are often not applicable to SSA eligibility determination, such information as direct measures of skills and records reflecting overall adjustment can be highly useful. Attempts should be made to obtain records from other agencies with that information interpreted by persons familiar with the functioning of the agency. Information from schools that is particularly relevant to mental retardation diagnoses includes measures of skills such as standardized test results, teacher-assigned grades, history of retention in grade, curriculum track pursued, and participation in special education programs, regardless of diagnosis.

CONCLUSIONS AND RECOMMENDATIONS

Broad consensus exists about the appropriate cutoff criteria for intellectual ability in mental retardation diagnosis: A cutoff score of approximately two standard deviations below the mean is well accepted in most settings. A similar consensus does not exist, however, regarding the appropriate cutoff for adaptive behavior. The use of a stringent adaptive behavior cutoff like that used for intellectual functioning would sharply reduce the number of people with IQs below 70 eligible to be considered for a diagnosis of mental retardation. On the basis of the committee's knowledge of individuals with mental retardation as well as the relevant research literature, this outcome is undesirable. We, therefore, propose formal adaptive behavior assessment as part of a comprehensive evaluation for individuals with or suspected of having mental retardation and cutoff scores that are more lenient than those widely used for intellectual functioning.

Diagnostic decision making in mental retardation needs to be based on a comprehensive evaluation that uses multiple methods of collecting data from multiple sources across multiple settings. We support a principle of convergent validity as a means to interpret a broad variety of information. As discussed in Chapter 3, in the rare case in which a composite IQ is suspected to be spurious, the composite score should be ignored and either an appropriate part score (as described in Chapter 3) or other methods should be used to confirm or disconfirm a diagnosis of mental retardation. Diagnostic decisions should always be based on the preponderance of evidence, not just one numerical score.

Finally, the need for more research, particularly on the measurement of adaptive behavior, is crucial to improving decisions about mental retardation eligibility. Adaptive behavior assessment is not as well developed as intellectual assessment, although improvements over the last decade have occurred.

On the basis of its review, the committee makes the following recommendations.

Recommendation: A diagnosis of mental retardation must be based on high-quality assessments of intellectual and adaptive functioning that meet the following criteria:

- **A broad variety of information on adaptive behavior and intelligence should be collected, including data on performance in different settings, from different sources, and using varying methods.**
- **Comprehensive, multifactored measures of intelligence and adaptive behavior should be used in mental retardation eligibility determination. Brief, unidimensional measures or short forms of comprehensive tests should not be used.**
- **The principle of convergent validity shall be applied in eligibility decisions about mental retardation eligibility. Information that is inconsistent with a diagnosis of mental retardation should be recognized, evaluated, and explained in the overall diagnostic decision.**
- **Assessments must be conducted by people with appropriate education and experience for the kind of instrument used and the nature of the eligibility decision to be made. People conducting intellectual assessments must meet publishers' requirements for Class C instruments.**
- **Measures of adaptive and intellectual functioning should be carefully selected and interpreted in order to minimize the negative effects of low validity, low reliability, floor and ceiling effects, and steep item gradients.**
- **The norms for measures of adaptive and intellectual functioning must be suitably contemporary. Use of outdated norms or previous editions of recently restandardized measures is not**

acceptable. The norms for intellectual measures should be no older than 12 years because of the deterioration of the normative standards over time.

- **Decisions about mental retardation eligibility should be made by people with appropriate preparation in the areas of mental retardation and other disabilities and disorders, measurement of intellectual and adaptive functioning, knowledge of human development, and the influence of context on behavior.**

The committee's recommendations reflect concerns about the quality of the available evidence as well as the necessity to provide reasonable guidance to people making eligibility decisions regarding a diagnosis of mental retardation. The committee concludes that more research on the measurement of adaptive behavior with children and adults is urgently needed, including investigation of classification agreement. The following recommendation reflects these concerns.

Recommendation: Federal agencies, including the Social Security Administration (SSA), should fund studies to evaluate the accuracy of program eligibility decisions and foster research on adults with mental retardation, including their adaptive behavior. The research funding should include investigations of multimethod techniques for the assessment of job-related skills, social adaptation, health, and well-being. In addition, relevant epidemiological studies and research on the accuracy of diagnosis of mild mental retardation are essential to inform policy and decision making.

- **SSA should evaluate the consequences of implementing the committee's recommendations in the context of public policies and economic conditions, reporting findings to the public within five years.**
- **Since improved accuracy in eligibility determination depends more on improved measures of the key dimensions of mental**

retardation than on adjusting cutoff scores, the committee recommends research on improving measures, especially adaptive behavior assessment, and on methods to combine information on adaptive and intellectual functioning in making eligibility decisions based on a diagnosis of mental retardation.

- **SSA should make available for use by legitimate researchers tapes of Supplemental Security Income and Disability Insurance program utilization, comparable to public-use tapes available for Medicaid program utilization.**
- **SSA should link its data on individual benefit awards to other agency data on health care and service costs for those same beneficiaries.**
- **SSA should examine data on eligibility determination procedures across its 10 districts, in order to discover if implementation of classification policies is consistent or varies regionally.**

Chapter 6

Differential Diagnosis

Differential diagnosis refers to the process by which a clinically significant problem or set of symptoms is evaluated and distinguished from other conditions, usually associated with similar clinical features. In the case of mental retardation, many neurodevelopmental and psychiatric disorders can mimic or accompany the mental retardation diagnosis. These other disorders may also be independent grounds for Supplemental Security Income (SSI) and Disability Income (DI) benefits eligibility. It is often challenging for the Social Security Administration (SSA) disability determination specialist to distinguish between mental retardation and other neurodevelopmental and psychiatric disorders with similar signs and symptoms. In addition, when these other disorders are present in addition to mental retardation, they may make assessment of intellectual and adaptive functioning even more complicated than it frequently is. In such cases, it is imperative that clinical evaluations take these complex factors into consideration and draw conclusions based on objective data and best clinical practice.

This chapter focuses on the challenges likely to be encountered in determining eligibility for SSI and DI benefits through a diagnosis of mental retardation. As outlined in earlier chapters, especially Chapter 2, the accuracy of the diagnosis of mental retardation is crucial both to SSA and to the client: to the former so that it can carry out its mandate, and to the latter for essential income support and as a gateway to other needed entitlements. While earlier chapters dealt with issues related to defining mental retardation and to the diagnostic criteria for this condition, here we consider distinguishing mental retardation from other conditions. In addition, we review circumstances that can complicate the process of assessment and discuss pertinent issues in evaluating intellectual and functional status across the life span.

The prevalence of mental retardation is highest among school-age populations (e.g., Kiely, 1987; Larson et al., 2001) and, by definition, its symptoms must be present before adulthood, though actual diagnosis may not take place until after schooling ends. Depending on their state of residence, between 5.2 and 11.5 percent of all children attending public schools are currently classified as having a disability (not necessarily mental retardation) of sufficient severity to receive special services (U.S. Department of Education, 1994, Table AA25). However, eligibility for special services in school does not necessarily confer eligibility for SSI benefits. Most individuals with a diagnosis of mental retardation who receive special education services become gainfully employed once they leave the school system. There is, however, an increased probability that individuals in this population may need financial support at some time during their adult lives.

Mental retardation is a prevalent disability among the population receiving special education services, but substantial numbers of these individuals will have some other condition associated with varying degrees of disability, including developmental, learning, sensory, motor, or psychiatric disability. These conditions produce a wide range of impairments, and to a greater or lesser degree can share signs or symptoms with mental retardation. It is important to determine the nature of each individual's limitations in order to develop plans for education, habilitation, and vocational training, as well as for making SSI eligibility determinations.

CLINICAL CONSIDERATIONS AND ELIGIBILITY DETERMINATION

To arrive at a correct diagnosis or case classification, clinicians evaluate the client's signs and symptoms and compare them to those of other similar conditions. Additional information is then gathered to rule out possibilities systematically until a valid diagnosis can be made. Symptoms refer to the concerns expressed by the affected individual or his or her representatives, while "signs" are objective findings from a physical examination, laboratory tests, psychological assessment, or educational evaluation.

In the case of mental retardation, presenting signs and symptoms are usually evident and associated with underlying intellectual and adaptive impairments. Nevertheless, differential diagnosis can be complicated by at least four other factors. First, the severity and nature of impairments can vary substantially, with each person having his or her own strengths and weaknesses in performance. Thus, clinical presentation may be different among affected individuals, some of whom have abilities that are only subtly different outside the classroom setting from those of people without mental retardation. Second, the specific manifestations of mental retardation change with development throughout the life span. In less severe cases, it is difficult to be confident about a diagnosis made during infancy or early childhood, yet assessment is rarely conducted for clients after they leave school. Third, other conditions can produce signs and symptoms very similar to those of mental retardation, especially before elementary school begins. These include other developmental, learning, sensory, and motor impairments. In adolescence or young adulthood, psychiatric impairments may become increasingly evident, further complicating determinations of mental retardation or even exacerbating its effects. Finally, there are a number of other conditions that can complicate assessment and make it difficult to determine whether a diagnosis of mental retardation actually is an underestimation of a person's true capabilities. These issues are all reviewed in this chapter, with particular attention to their significance in the context of eligibility for SSI and DI.

Eligibility for Support

It is sometimes difficult to distinguish between mental retardation and other conditions currently covered by SSI, the so-called categorical disorders: autism, learning disability, borderline intellectual functioning, and organic mental disorders marked by specific behavioral phenotypes. It is particularly important to make these distinctions when selecting treatments and developing plans for education, habilitation, and vocational training. Diagnostic information is used in making very different types of decisions in clinical, educational, vocational, and SSI contexts. In educational settings, it is critical to determine if a condition other than mental retardation (like a specific learning disability or sensory, motor, or psychiatric impairment) is causing or contributing to poor performance. These differential diagnoses have direct implications for developing individualized plans of treatment and instruction to encourage learning. In clinical settings, arriving at the correct differential diagnosis determines appropriate and effective treatment and prevention strategies.

In the SSI context, the explicit focus is on eligibility for support—that is, current employability for adults and marked or severe limitations in developmental skills for children. For SSI eligibility, the underlying cause or origin of intellectual or functional impairment is not important if the individual meets the income criteria and intellectual disability definitions for mental retardation. Eligibility is dependent on the severity of impairment regardless of its source. Therefore, the primary goal of a diagnosis differentiating between mental retardation and some other condition is not relevant when the severity of disability clearly meets or exceeds eligibility criteria within SSA guidelines. This would include situations in which either the total IQ test score is less than 60 or the individual has another condition that is so disabling that it interferes with the determination of true intelligence, such as blindness, deafness, uncontrolled seizures, cerebral palsy, or some other disorder. Similarly, once eligibility for SSI benefits for individuals with mental retardation is determined, the decision process does not need to consider additional comorbid disorders, and the process of differential diagnosis need not continue further.

Differential diagnosis is a different matter for individuals having less severe disabilities that put them at the border of SSI eligibility criteria. For individuals with mild mental retardation, it is vitally important to generate a complete description of signs and symptoms, including diagnosis of comorbid conditions, as well as an accurate assessment of intellectual and functional status. Furthermore, the presence of additional disabilities in such cases can impair functioning to such an extent that competitive employment is unlikely or even impossible. For cases having less severe clinical presentations, carefully conducted evaluation procedures that include standardized assessments of intellectual development and adaptive function (see Chapters 3 and 4) can usually provide the accurate estimates of capabilities needed in arriving at a valid differential diagnosis, except perhaps for very young children.

Currently, SSA defines mental retardation by quantitative rather than qualitative performance criteria relative to age-referenced expectations. This is incorporated into SSA regulations, which include extended discussions of assessment methods and criteria that vary with age and with respect to broad domains of functioning. SSA has deliberately built some flexibility into its procedures for determining SSI eligibility at various chronological ages (Social Security Administration, 2000), but the clinician must recognize the strengths and limitations of assessment tools at each stage of development. Evaluators must be sensitive to the diagnostic precision of the instruments employed, aware of the types of errors that imprecision may introduce, and be able to access sources of additional information that can reduce the likelihood of misclassification. These issues are discussed at length in Chapters 2 and 3.

Individuals can qualify for SSI in two ways. The first and most obvious is by having a condition listed by the SSA as eligible by definition. Two lists define conditions that meet the standard for severity; one applies to both children and adults and the other list is for use with children under 18 years of age. These lists are intended to objectify and facilitate the determination process, but problems can arise because, as pointed out in earlier chapters, the clinical definitions of mental retardation provided in authoritative references (e.g., *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*—DSM-IV—American Psychiatric Association, 1994), are not the same as the criteria provided in the SSA listing. Because of this disparity, some individuals with a clinical diagnosis of mental retardation do not meet SSI eligibility criteria. SSI eligibility is intentionally limited to only those individuals whose disability is severe enough to impose substantial restrictions on their ability to work, either now or in the future.

Individuals can qualify for SSI support in a second way: they are eligible if their status is functionally equivalent to a listed condition. In the case of mental retardation, functional equivalence for children is assessed in six domains: (1) acquiring and using information, (2) attending and completing tasks, (3) interacting and relating with others, (4) moving about and manipulating objects, (5) caring for oneself, and (6) health and physical well-being (Social Security Administration, 2000). To meet functional equivalence criteria, the child must have either marked impairment (equal to or greater than 2 but less than 3 standard deviations below the mean on a standardized, norm-referenced assessment) in two domains, or extreme impairment (equal to or greater than 3 standard deviations below the mean) in a single domain. While this second path to eligibility offers children having unrecognized mental retardation an opportunity to obtain needed support, this opportunity rests to a substantial degree on the judgment of the examiner. Therefore, the use of uniform, valid procedures and criteria are essential for cases falling into this category, as well as for similar situations with adults.

Issues Associated with Life Span Development

Age-related considerations are crucial in making a diagnosis of mental retardation because key signs and symptoms, as well as appropriate assessment methods, are quite different across the life span. Abilities and socially appropriate expectations change dramatically with development, and these factors can complicate the evaluation process in distinct, age-specific ways. Because adaptive behavior changes and tends to increase with age (e.g., Hundert et al., 1997), most measures of adaptive behavior are structured so that items are presented in a developmental sequence. Declines in adaptive behavior can be associated with advanced aging, as it is in the general population, or at an earlier chronological age in the instance of some disorders, like Down syndrome (Kapell et al., 1998; Zigman et al., 1996), or in the presence of severe to profound mental retardation. Moreover, although there are systematic differences in the typical adaptive performance of people with mild mental retardation and their peers with more severe degrees of mental retardation, there is also overlap in adaptive behavior skills among people assessed with differing degrees of intellectual disability (Janicki & Jacobson, 1982). For example, it is possible for someone with mild mental retardation to have certain adaptive behavioral skills that are less advanced than those of another person with moderate mental retardation. These differences may reflect different experiences, opportunities, and participation in services. The following sections use four stages of maturation to identify and discuss key factors in making the differential diagnosis of mental retardation and determining SSI eligibility.

INFANTS, TODDLERS, AND PRESCHOOLERS

The limitations of current assessment technologies are a major problem in evaluating disability in very young children. Advances are being made rapidly, especially in identification of genetic conditions likely to increase risk, but diagnosis of mental retardation remains difficult except in cases with clear biomedical manifestations or severely delayed development. This is especially true until the age when language delays can be documented, or even later when academic achievement is regularly evaluated. For children under age 3, a developmental quotient (developmental age/chronological age × 100) may serve as a surrogate for standard IQ, such that quotients of less than 67 define marked impairment and quotients less than 50 define extreme impairment (Social Security Administration, 2000). This approach is currently considered to be the best available diagnostic practice, but it has limited predictive validity and precision. The distinction between generalized and specific delays, both of which can have an impact on function to a degree that will quality for SSI, often has to be postponed until later ages.

In preschoolers, language impairment is the most commonly identified specific developmental disorder. In principle, specific language impairment is distinguished from mental retardation by the relative preservation of nonlanguage skills. However, both conditions are frequently associated with social and behavioral deficits, and both are strong predictors of poor academic achievement at older ages (McArthur et al., 2000). Therefore, specific language impairment can be difficult to distinguish from mild mental retardation, as the following case illustrates.

Allen is a 3-year-old who has been referred for delayed language development. He was born at full term following an uncomplicated pregnancy. He is a healthy child whose early developmental milestones were achieved on time or only slightly delayed with the exception of language. He started speaking single words at 18 months of age, began using baby talk at 21 months and did not start speaking in phrases until recently. Allen waved "bye-bye" before his first birthday, knew some of his body *parts by 24 months, and follows some multistep commands. He likes to play and roughhouse with his 4-year-old brother. His hearing has been tested and found to be normal. His rate of language development has been approximately two-thirds of normal expectations. On formal evaluation he was found to be functioning in the low borderline range of intelligence but he showed clear unevenness in his results. Language abilities were in the deficient range, while problem solving abilities were in the borderline range. With this developmental profile, it is difficult to distinguish between specific language impairment or mild mental retardation. Early intervention should target development of language and general intellectual functioning, with frequent follow-up evaluations (at approximately six-month intervals) to monitor progress and refine diagnosis.*

In Allen's case, uncertainty regarding differential diagnosis can be resolved in a year or two when he can be given a detailed assessment that focuses on both language development and intelligence, but in many similar cases, diagnostic evaluations will be delayed until early in the child's schooling.

In contrast to the severe constraints on the ability to diagnose mild mental retardation in young children, a great deal is now known about the conditions associated with increased risk for mental retardation. Babies born prematurely, very small for their gestational age, or exposed in utero to alcohol or other teratogenic agents are at increased risk of mental retardation. Babies are also at risk if their mothers are in their young teens, have a history of poor prenatal care, have less than a high school education, or are living in poverty (e.g., Sameroff, 1986). With particular relevance to SSI determination, many of these risk factors are also associated with economic disadvantage. However, while large numbers of young children are at risk, the majority of them do not have mental retardation. Current research is focusing on refining the ability to identify those children who will, in fact, develop atypically.

Accumulating evidence suggests that intervention may be most effective if it is initiated as early as possible, although there is still debate in the professional and scientific community on this point (Bailey et al., 1999; Ramey & Ramey, 1999). Therefore, public health policy needs to weigh the relative value of providing earlier supports to an expanded subpopulation, only a subset of whom will prove to have a substantial disability, against the effect of delaying supports and treatments for a more limited number of individuals who may then need more intensive interventions over a longer period of time. A prime example of this dilemma has been highlighted in the field of autism, in which early, intensive treatment has produced a range of positive outcomes; for some children, the results of treatment have been dramatic (Lovaas, 1987; National Research Council, 2001; Rogers, 1998). Recent neuropsychological and neurobiological evidence also supports the hypothesis of a window of opportunity for intervention effectiveness in very young children at risk for mental retardation and other developmental disabilities (Bailey et al., 1999).

There are several categorical disorders that are likely to be encountered while evaluating very young children with suspected mental retardation. The disability examiner must determine whether intellectual or adaptive deficits are due to retardation or aspects of these other conditions.

Autism and Pervasive Developmental Disorders

John is a 4-year-old with delayed language and social skills development. Parents report that John is "in his own world" and has limited spontaneous speech, although he frequently repeats back exactly what is said to him. He has other unusual behaviors. These include toe walking, hand flapping, and excessive spinning of the wheels on his toy truck. John has excellent puzzle-solving skills, but he does poorly with changes in routine and seems most content when he is entertaining or playing by himself for long periods of time. During formal testing, John often seemed inattentive and his obtained total test score (IQ) of 55 on the Stanford-Binet, Fourth Edition, would seem to be an underestimate of his true intelligence.

Autism, currently classified as a pervasive developmental disorder with symptoms that are present by 3 years of age, classically involves disturbances of: (1) social interaction, (2) communication, and (3) behavior. Autism differs from mental retardation in that intellectual deficits may or may not be apparent, but social, communication, and behavior disorders are always noted. Difficulties with social interactions include avoidance of eye contact during social activities, difficulty developing peer relations, excessive self-directed play, or avoidance of shared activities. Communication deficits may include unusual speech prosody (e.g., atypical intonation, cadence, or inflection; immediate or delayed echolalia; perseveration; lack of spontaneous speech; or difficulty with pragmatic language initiation). Behavior difficulties noted in autism include repetitive stereotypies (e.g., hand flapping, toe walking, rocking); extreme difficulties with changes in routine; and limited imaginative play or preoccupation with parts of toys (e.g., spinning the wheels of a truck rather than imaginative play with the whole truck) (American Psychiatric Association, 1994; Stone, 1997; Volkmar et al., 1994, 1997, 1999; Wolraich et al., 1996).

A recent review of 23 surveys focusing on autism and including over 4 million individuals reported a median prevalence of 5.2 per 10,000 population (95 percent confidence interval 4.5-5.5 per 10,000; Fombonne, 1999). Prevalence rates appear to have increased in recent years, presumably due to improved early identification and changes in case definition, but additional research is needed to confirm this. On average, there are 3.8 times as many males as females with autism (Fombonne, 1999).

Distinguishing whether mental retardation is present in individuals with autism is challenging, due in part to limitations in communication skills and social interaction deficits of the affected individuals (Arnold et al., 2000), as well as the diverse etiological and symptomatic expressions of the disorder. Nevertheless, Fombonne (1999) recently reported that an estimated 80.3 percent of 1,533 children with autism also had mental retardation. Estimates indicated that approximately 29 percent had mild to moderate mental retardation and 42 percent had severe to profound mental retardation. While the overall proportion of male to female cases with autism was reported to be 6:1 among the subgroup with higher cognitive functioning (n = 800), this ratio dropped to 1.7:1 among the subgroup with moderate to severe mental retardation (n = 748). These data suggest that there may be gender differences in the underlying causes of disability among these two subpopulations.

Clearly, a child who has significant adaptive behavior deficits but a full-scale IQ in the 76 to 85 range would not be eligible for SSI support because of mental retardation but would be eligible instead because of his or her diagnosis of autism. Once a diagnosis of autism or mental retardation is made, no further diagnostic information is needed for the disability determination process.

Pervasive developmental disorder, not otherwise specified (PDD NOS) shares many characteristics with autism, including marked impairment of communication, social interaction, and stereotypic behaviors. However, individuals with this disorder do not meet specific diagnostic criteria for autism, Rett's disorder, childhood disintegrative disorder, Asperger's disorder, schizophrenia, or avoidant personality disorder. While clear deficits exist in social, communication, and other skill areas, they are generally less severe than those seen in classical autism (Fombonne, 1999; Volkmar et al., 1999). Prevalence estimates range from 1.9-16.3 per 10,000, with a median base rate of approximately 8.8 per 10,000. This indicates that pervasive developmental disorder occurs more than twice as frequently as autism (Brask, 1970; Burd et al., 1987; Cialdella & Mamelle, 1989; Fombonne, 1997, 1999; Fombonne & du Mazaubrun, 1992; Hoshino et al., 1982; Lotter, 1966).

There have been very few systematic outcome studies of children with pervasive developmental disorder or of its comorbidity with mental retardation. Available data suggest that the prognosis for individuals with this disorder is better than for those with autism (Smith et al., 2000), but difficulties persist in the social, communication, and behavioral domains. These children may also be at higher risk for emotional disabilities in later life (Towbin, 1997; Volkmar et al., 1999). As with autism and the other disorders reviewed in this section, pervasive developmental disorder may coexist with mental retardation. The process of differential diagnosis helps determine the correlates and consequences of various syndromes.

Autism spectrum disorders (autistic disorder, pervasive development disorder-not otherwise specified, Asperger's disorder, and childhood disintegrative disorder) are listed among the conditions qualifying for SSA disability benefits. Therefore, a diagnosis of one of these conditions or of mental retardation is sufficient for eligibility for SSI if income standards are met. It must be remembered that symptoms can be subtle in some cases of autism spectrum disorders, and when such cases are identified, the additional impact of intellectual limitations can be quite substantial. Therefore, it is imperative to consider the combination of multiple mild impairments that together may result in significant overall disability.

Cerebral Palsy

Tina is an 18-month-old who was born 12 weeks prematurely. Birth weight, height, and head circumference were appropriate for her gestational age. Tina had an intraventricular hemorrhage (bleeding from the interior of the brain) when she was 1 week of age. Parents report that Tina has not developed motor skills like her older sister. Her legs are stiff and she cannot sit up by herself, although she is able to roll and creep on her belly to move around. She seems very happy, has begun to babble, and seems to be trying to say a few words, although her speech is not fully intelligible. Tina's performance on the Bayley Scales of Infant Development indicated a developmental quotient of 70, but her long-term prognosis, and whether or not she will be able to live and work independently in the community, are unclear.

Cerebral palsy is a disorder of the central nervous system that affects the posture, movement, and muscle tone of the individual. It is caused by a static insult to the brain. Children with cerebral palsy have persistent motor delays and associated cognitive, behavioral, neurosensory, and orthopedic abnormalities, which may change with age. Cerebral palsy is classified as: (1) spastic cerebral palsy (65 percent), including diplegia (30 percent), hemiplegia (30 percent), and quadriplegia (5 percent); (2) dyskinetic (choreoathetoid) (19 percent); and (3) ataxic (10 percent) (Palmer & Hoon, 1995). The prevalence of cerebral palsy is estimated at 2 per 1,000 live births. An increase in cerebral palsy among low birth weight infants has been documented during the last decade, attributable to the decreased mortality in this group (Palmer & Hoon, 1995).

The estimated overall prevalence of mental retardation among children with cerebral palsy varies from 30 to 77 percent. Hemiplegia and diplegia are associated with relatively higher cognitive functioning compared with other types of cerebral palsy (Palmer & Hoon, 1995). In a recent study of 12,709 children with cerebral palsy, 61 percent were found to have no, mild, or moderate mental retardation, while 39 percent had severe to profound mental retardation (Strauss et al., 1998). Differential diagnosis is important for determining eligibility for those individuals whose cerebral palsy is very mild and who have borderline intellectual functioning or mild mental retardation (i.e., full-scale IQs ranging from 60 to 80). The motor deficits may make performance of some tasks on intelligence tests difficult or impossible. Accommodations in the testing procedures and motor-free intelligence tests may be helpful in determining these individuals' intellectual potential (see Chapter 3). For individuals with severe cerebral palsy, motor impairments meet disability criteria for SSI eligibility.

Seizure Disorders

Jimmy is a 5-year-old who was developing normally until he reached 4 years of age, when he was noted to have deterioration of expressive and receptive language skills and onset of behaviors similar to children with autism. Jimmy was noted to have decreased social interaction, limited initiation of speech, and staring spells.

Seizures are alterations in motor, behavior, consciousness, or sensory function caused by repetitive, simultaneous electrical activity of the nerve cells in the brain (Vining & Freeman, 1996). Epilepsy is diagnosed when the individual has recurrent seizures. Nonfebrile seizures occur in 0.5-2 percent of children, while epilepsy occurs in 0.5 to 1 percent of children. There are two large categories of seizures, namely, generalized seizures and partial seizures. Generalized seizures include grand mal (tonic-clonic) seizures and petite mal (absence seizures), while partial seizures arise from focal areas of the brain and their manifestations usually are associated with the area of brain disturbance. Seizures may end with a postictal period in which the individual is partially or nonresponsive and may have accompanying neuromotor deficits, such as weakness of the arm.

Seizure disorders are more frequent in children with other developmental disabilities (Vining & Freeman, 1996). Children with mental retardation have a 5-10 percent incidence of epilepsy (Hauser & Hesdorffer, 1990). Of the children with mental retardation who have other neurological disorders, 40-50 percent of children with cerebral palsy and 70 percent of children with postnatal brain injury have epilepsy (Goulden et al., 1991). West syndrome (infantile spasms), Lennox-Gastaut syndrome, and Aicardi syndrome are epilepsy syndromes frequently associated with a high risk for mental retardation. Neonatal seizures are usually associated with central nervous system abnormalities as well as increased risk of mental retardation, cerebral palsy, and epilepsy (Painter & Gaus, 1991).

Landau-Kleffner syndrome, or acquired epileptic aphasia, is a rare presentation of epilepsy in childhood in which children show progressive loss of speech and epilepsy. These children may be difficult to diagnose because of their previously normal development and their slow progressive loss of receptive and expressive language skills, making psychological testing both challenging and variable over time. Children with Landau-Kleffner syndrome may be misdiagnosed as having mental retardation or autism. An electroencephalogram (EEG) that is markedly abnormal over the left temporal region of the brain helps to make this diagnosis. Unless entirely controlled through medication, seizure disorders may affect learning capacity from preschool through high school and, ultimately, employability.

Hearing Impairment

Hearing loss in children is often a silent and hidden disability that can lead to speech, language, and cognitive delays in the absence of appropriate intervention. One in 1,000 children are born with profound hearing loss, while 2 in 1,000 will acquire deafness during early childhood. And 95 percent of acquired hearing loss is attributable to middle ear infections (Northern & Downs, 1991). Hereditary hearing loss occurs in 1 in 2,000 to 6,000 children, caused by over 70 syndromes, with about 90 percent of hearing loss inherited as an autosomal recessive disorder, often with siblings involved (Konigsmark & Gorlin, 1976).

A report from the Institute of Medicine (1974) concluded that a hearing loss of 15 dB or greater (500-2,000 Hz) represents a significant impairment. By this definition, 2.2 percent of children ages 4-11 were found to have bilateral hearing losses. Losses of 15 dB may result in language delays probably due to the inability to hear the unvoiced consonants: s, p, t, k, th, f, and sh. Thus, assessment of developmental skills may be challenging in some children, even with mild hearing deficits, due to lack of constancy of auditory cues, confusion of acoustic parameters in rapid speech (e.g., understanding directions), difficulty with perceiving speech sounds, and poor mastery of grammatical rules and meanings (Northern & Downs, 1991). A child is considered deaf when there is a 70 dB or greater hearing loss. Even in these cases, if hearing loss is an isolated disability and no other factors limit cognitive development, successful academic experiences, including college and independent adult employment (Batshaw & Perret, 1992), can be expected. However, many individuals with hearing impairment experience developmental delays, and differential diagnosis is necessary to determine whether poor preschool achievement or suspected cognitive impairment is due to mental retardation or to hearing loss.

Vision Impairment

The incidence of blindness in childhood, estimated to be about 1 in 3,000 (Foster, 1988), is defined as visual acuity in both eyes of less than 20/200 or a visual field of less than 20° despite correction with eyeglasses. Causes of blindness include congenital infections and malformations, injury, prematurity, postnatal infections, vitamin deficiencies, and tumors. Approximately half of all children with blindness have other developmental disabilities (Warburg et al., 1979). Children with early onset blindness frequently have developmental delays in gross and fine motor functioning, adaptive skills, and language (Dekker & Koole, 1992; Frailberg, 1977; Pring, 1984; Teplin, 1983). These children also often exhibit unusual behaviors and self-stimulation (Good & Hoyt, 1989).

Cognitive testing is difficult in children with blindness, especially before they reach age 5, when verbal subtests of individually administered intelligence tests (e.g., Wechsler Intelligence Scale for Children-III—WISC-III) can provide valid estimates of intellectual development in the absence of language impairment. While children with visual impairment may have difficulty with abstract concepts, they often do well with numerical computation (Batshaw & Perret, 1992). As with hearing loss, differential diagnosis is important for determining the underlying reasons for any delays in cognitive development and poor preschool achievement.

SCHOOL-AGE CHILDREN

The U.S. Department of Education (1999) estimates that there are currently over 5 million children receiving special education under the Individuals with Disabilities Education Act (IDEA). Methods appropriate for evaluating school-age children (K-12), diagnosing mental retardation, and determining SSI disability status have been in place for many years. However, the current social and political climate tends to limit both the number of mental retardation diagnoses and placements in designated special education programs because these actions may be considered to be stigmatizing.

As mentioned earlier, there is wide variation in state special education policies regarding placement and classification of individual students. Nationally, the numbers of students identified as having a learning disability have been increasing while numbers of students identified as having mental retardation have been decreasing (Lyon, 1996), and this has serious implications for the process of determining eligibility for SSI support. Many children who in the past might have been diagnosed with mental retardation are now being classified otherwise (e.g., learning disability) and therefore have a diagnosis that is not currently included among SSA's listing of categorical disabilities. In addition, many students with less severe impairments who in past decades would have received special education services may now complete their schooling in regular classroom programs. Thus many young adults with mental retardation may not be on SSI rolls, and after completing their schooling may not know of the benefits available to them.

The conditions complicating the diagnosis of mental retardation in young children reviewed earlier continue to be a concern throughout childhood. However, in contrast to the situation during early development, the demands of schooling provide the opportunity for recognizing academic difficulties. Clinicians and educators typically have access to methods for identifying most complicating conditions; they should therefore be able to determine the factors that are contributing to poor performance in the classroom or during assessments. Nevertheless, disability examiners must now face the task of making differential diagnosis among learning disability, mild mental retardation, attention deficit hyperactivity disorder, borderline intellectual impairment, and SSI-eligible mental retardation. This is the topic of the next section.

Mild Mental Retardation Versus Learning Disability

Clear criteria for differentiating learning disabilities from mental retardation are provided in both the DSM-IV and the *International*

Classification of Diseases and Related Health Problems, Tenth Revision (ICD-10) (World Health Organization, 1993). Children with specific learning disabilities show academic underachievement relative to their composite IQ. Specific learning disabilities may occur in basic reading (decoding), reading comprehension, mathematics reasoning, mathematics calculation, written expression, oral expression, and listening comprehension. Severe forms of specific learning disability may result in illiteracy, failure to understand the most basic of mathematical concepts, or inability to write. These disabilities can interfere with the individual's ability to maintain gainful employment, and estimates indicate that they may affect 4 percent or more of the overall population (U.S. Department of Education, 1999).

While impairments limited to a specific cognitive domain are characteristic of learning disabilities, mental retardation is associated with a substantially broader profile of deficits. There are, however, rare cases of individuals with a diagnosis of both mental retardation and comorbid learning disability. This occurs when there is documented performance in a specific domain that is substantially below the level typical of the individual's composite IQ. Learning disability and mental retardation are currently usually considered to be mutually exclusive diagnoses, although this is changing with learning disability being increasingly thought of as low achievement, rather than in terms of a discrepancy between achievement and ability. In addition, mental retardation and learning disabilities are associated with assessment profiles that are quite distinct and unambiguous for school-age individuals and adults. In fact, an IQ in the range required for a diagnosis of mental retardation is listed among the exclusionary criteria for learning disabilities. It should be noted that the term "mental retardation" has been replaced with "learning disability" throughout Great Britain, and therefore the differential diagnostic terminology is different. Nevertheless, the operational differential diagnostic criteria employed in Great Britain to distinguish specific developmental disorders of speech and language, for example, from learning disability would be comparable to DSM-IV and ICD-10 standards.

Clearly, differential diagnostic criteria are available, objectively stated, and widely disseminated. However, advocacy and administrative policies in public school districts also influence how children are labeled, and "learning disability" is often preferred to "mental retardation." This can reduce stigmatization, but it could also complicate SSI eligibility determination. Rather than focusing on classification provided in school records, the actual test results should be the focus of the determination process, even in cases in which children were not diagnosed with mental retardation in their school-based classification.

This is a major issue that requires broad policy attention and vigilance (e.g., MacMillan, Gresham et al., 1996). It is tacitly acknowledged in the SSI disability evaluation guidelines, in which reliance on objective evidence and variation in policies regarding special education placement are discussed (Social Security Administration, 2000). It is also illustrated by the substantial variability that exists from state to state in the numbers of school-age children classified with mental retardation (Frankenberger & Fronzaglio, 1991; U.S. Department of Education, 1994). As mentioned earlier, the question has been systematically examined in California, where school-based classifications of learning disability are made in some cases that, following accepted clinical standards, could be validly diagnosed with mental retardation (e.g., MacMillan et al., 1998).

Given this reality and the absence of learning disabilities among SSA's listing of qualifying conditions, a considerable number of people who truly have less severe mental retardation, and both need and are legally eligible for support, may not be judged to meet eligibility requirements by some examiners unless the procedures for defining functional equivalence are clearly stated and uniformly implemented. Mild intellectual impairment is not, however, a condition that automatically prevents employment. Indeed, these individuals are the most likely candidates among the population of individuals with mental retardation to be gainfully employed. Nevertheless, there is a clearly increased risk that they will have difficulty achieving and maintaining employment after high school (Sitlington, 1996).

Melissa is an 8-year-old referred for academic underachievement and *increasing moodiness. She is currently clumsy, her attention is variable, and she frequently daydreams. Her early developmental history indicated delay in learning to tie her shoes and articulation difficulties requiring speech therapy. She had difficulties with prereading skills during kindergarten. Despite tutoring and 1.5 hours per night of homework, she did not attain the skills needed to pass 1st grade. While repeating 1st grade, she is still having trouble with her reading lessons. Melissa guesses at words and does not exhibit a phonetic approach to reading, and despite a standard score in reading of 67, she scores in the borderline range on the WISC-III (composite IQ = 71). She demonstrates poor attention as measured by a continuous performance task, and she shows evidence of a graphomotor disorder that limits her ability to write. There is some suspicion of possible mental retardation, but specific language impairment seems more likely.*

Melissa would probably not meet SSI eligibility criteria based on these findings, but additional evaluations might provide sufficient documentation of substantial impairments and perhaps mental retardation. In any case, she seems to be falling further and further behind her peers, and if she does not qualify for support now, her impairments may meet eligibility criteria in the future.

Borderline Intellectual Functioning

A substantial number of children experiencing academic difficulties will be found to have composite IQs of 71 to 85 with relatively homogeneous performance profiles consistent with their overall level of intelligence. In the absence of substantial adaptive deficits, these cases will fail to meet clinical diagnostic criteria for either mental retardation or learning disabilities. According to the DSM-IV, a classification of borderline intellectual functioning may be appropriate in these cases as long as the composite IQ falls between 71 and 85. However, there is no alternative SSI classification for individuals with composite IQs of 71-85 without associated functional limitations, and therefore these individuals do not meet SSI eligibility criteria for mental retardation. Borderline intellectual functioning is not a disability in and of itself. However, it can coexist with another condition (or conditions) and contribute to the performance difficulties of children and adults associated with pervasive developmental disorders, cerebral palsy, or sensory impairment. Thus, its presence should be considered a contributing factor when other disabling conditions are also present.

HIGH SCHOOL YEARS AND ADOLESCENCE

Adolescence is a period of particular developmental significance, encompassing the transition into young adulthood (Corbett & Barton, 1992; Marshak et al., 1999; McCullough & Rutenberg, 1988). Most people consider adolescence to end at around age 18 (roughly corresponding to expected graduation from high school), but individuals with mental retardation experience developmental delays that can extend their adolescence, as well as their high school education, through age 22 (Individuals with Disabilities Education Act, 1990, 1997). The vast majority of individuals have had the presence of any developmental disabilities identified by the time they reach adolescence. Nevertheless, this is a time of heightened opportunities but also of new risks. Sherrod et al. (1993) note that the rate of change during adolescence may create "special vulnerabilities and opportunities for intervention" (p. 218). For individuals with mental retardation or borderline intellectual functioning, the developmental changes typical of adolescence (emancipation, self-concept, use of leisure time, sexuality, vocational preparation, independent living) will be, to varying degrees, affected by their inability to achieve cognitive, communicative, social, and behavioral milestones. The stresses of normal adolescence may be amplified substantially for these individuals as they progress toward young adulthood (Marshak et al., 1999), or they may persist to ages later than those typically associated with the chronological end of adolescence.

People with certain mental disorders besides mental retardation are categorically eligible for SSI and DI benefits. Individuals with mental retardation are at increased risk for such psychiatric illnesses and severe behavioral or emotional dysfunction (Goestason, 1985; Koller et al., 1982; Menolascino & Fleisher, 1991; Nezu et al., 1992; Pfeiffer & Baker, 1994; Rutter et al., 1976). They also are vulnerable to the full range of psychopathology that occurs in the general population (Campbell & Malone, 1991; Eaton & Menolascino, 1982; Galligan, 1990). Psychiatric impairment and extreme maladaptive behavior can complicate the diagnosis of mental retardation. There has been little study of the association between risk of specific psychiatric disorders and age in individuals with mental retardation, but later adolescence through young adulthood generally represents a high risk period for many disorders (DSM-IV). Furthermore, this increased risk appears to be similar across Hispanic and non-Hispanic white Americans (Karno et al, 1987; Vernon & Roberts, 1982). Thus, behavioral or psychiatric conditions can contribute to the degree of disability, and with respect to SSI eligibility determination, this can be a particularly important factor for individuals with mild mental retardation.

Prevalence estimates for psychiatric or behavior problems vary widely for the population with mental retardation (Dykens, in press), depending on the methods used to ascertain cases (e.g., hospital patients, epidemiological samples, state systems) and the methods used to measure psychopathology (e.g., behavioral checklists, psychiatric interviews). Among referrals to a psychiatric clinic, for example, as many as 87 percent of children with mental retardation have been reported to meet criteria for one or more psychiatric disorders (Philips & Williams, 1975), but this clearly represents a subpopulation at elevated risk. Larger and more representative samples consisting of all consumers served by the California (N = 78,603) and New York (N = 30, 578) statewide mental retardation service systems found that 10 to 15 percent of individuals had significant psychopathology (Borthwick-Duffy & Eyman, 1990; Jacobson, 1982). Other studies using epidemiological approaches have reported rates of psychopathology ranging from 30 to 40 percent (Einfeld & Tonge, 1996; Reiss, 1990; Rutter et al., 1976). Clearly, a substantial subpopulation of individuals with mental retardation have significant psychiatric problems, but rates are much higher when salient behavioral disorders are included (Jacobson, 1999). The exact numbers of people with less severe intellectual impairments or mild mental retardation and comorbid psychiatric disorders are hard to estimate. Some advocates assert that these individuals are disproportionately represented among the homeless and chronically unemployed or in the criminal justice systems, while others argue that they are remarkably resilient and manage to meet basic needs.

The heightened risk of psychiatric disorder has been variously attributed to the intellectual, medical, physical, or sensory disabilities associated with mental retardation. These include difficulty in processing information, poor social judgment, poor ability to communicate, inability to self-regulate, and inability to reason beyond the here and now (Menolascino & Fleisher, 1991). Although individuals with mental retardation experience the same range of psychiatric difficulties as other individuals, the risk for specific problems varies across the population. Persons with severe retardation, for example, are more likely to be diagnosed with behavior disorders (Borthwick-Duffy, 1994; Einfeld & Tonge, 1996; Jacobson, 1982), while those with less severe intellectual impairments are more likely to receive conventional psychiatric diagnoses (Einfeld, 1992; Matson et al., 1988).

In addition, it is often difficult to separate the diagnosis of a behavior or mental health disorder from a behavioral characteristic linked to specific genetic (e.g., Prader-Willi syndrome) or nongenetic (e.g., fetal alcohol syndrome) conditions associated with mental retardation (McIntyre et al., 2002). Such variability among individuals bearing the label of “dual diagnosis” underscores the importance of obtaining appropriate psychiatric consultation and screening for at-risk individuals, and not attributing their aberrant behaviors directly to intellectual status. In some cases, successful treatment of dually diagnosed individuals can produce substantial improvements in performance. In fact, over time, effective treatments may be of sufficient benefit to change eligibility status, especially for individuals with less severe symptoms. Therefore, disability examiners must determine from the available assessment data if there is a realistic likelihood that treatment of secondary disorders will produce this type of improvement, and for redetermination purposes, the time course over which possible improvements would be likely to occur.

As adolescents make the transition from school to work settings, intellectual and adaptive behavior demands can increase, as do the consequences of maladaptive behavior. There is evidence that suitable supports are difficult to find for individuals with a diagnosis of both intellectual/cognitive and psychiatric impairments (Baker et al., 1993, 1996; Szymanski, 1987). These individuals are more likely to fail in community, residential, and work placements (Borthwick-Duffy & Eyman, 1990; Campbell & Malone, 1991; Galligan, 1990; Shoham-Vardi et al., 1996), indicating that individuals with mental retardation and psychiatric disorders or other behavior problems will fare considerably less well once they are adults, as demonstrated by the following case.

Daniel is an 18-year-old nearing the completion of high school. Daniel has been diagnosed with mental retardation almost since his first day of school and has been receiving special education services since preschool. The last time Daniel was tested, he had a composite IQ of 72 and was reading at the 2nd-grade level. Daniel is quite verbal but has clear deficits in communication and social awareness. Over the past few years, Daniel has been teased and picked on by his peers and has had problems controlling his anger in response to the teasing. Because his language and social skills are limited, he has difficulty communicating his frustration and resorts to inappropriate expression of his feelings by hitting, pushing, or impulsive rages. When referred for additional assessment, Daniel's scores on a widely used screening instrument indicated that he may have a psychiatric impairment and that further evaluation is needed. His transition teacher has had difficulty finding job placements in the community for him because of his impulsive and, at times, explosive behavior patterns. Although Daniel is approaching his formal transition from secondary school, a postschool plan has not been solidified.

The future employability of young adults with mental retardation is clearly influenced by the preparation they receive during formal schooling. Since 1990, the IDEA has mandated that schools develop a transition plan for all eligible students, which by definition includes all individuals with mental retardation regardless of severity. In developing an individualized transition plan, school personnel complete a vocational assessment and interview of the young adult or his parents to determine preferences and skills as well as the profile of disability (Grigal et al., 1997; Miner & Bates, 1997; Wehman, 1996). This process may also present an opportunity to obtain definitive diagnostic assessments that utilize community standards for classification of disability conditions. The transition process includes generating a plan that is intended to prepare students as effectively as possible for future employment success during adulthood. However, the success of these efforts will be influenced by a number of individual characteristics, including age, gender, physical health, cognitive functioning, adaptive behavior, and maladaptive behavior or psychiatric status of the participating individual (Blacher, 2001).

The same conditions that make differential diagnosis of mild mental retardation difficult can have a negative influence on transition, as can maladaptive behavior and psychiatric problems (Blacher, 2001). Ironically, more mildly affected individuals may not receive as much transition programming and preparation for employability as their more severely disabled counterparts because they are in regular classrooms, often struggling with the more traditional academic curriculum (Sitlington, 1996). Furthermore, untreated or inappropriately treated psychiatric disorders can severely limit employability, community integration, and general quality of life. Unlike their fellow students in regular classrooms, individuals with mild mental retardation have fewer cognitive resources to compensate for their psychiatric difficulties, and after they exit formal schooling they can lose touch with mental health or mental retardation service systems.

ADULTS

Wehman (1996) has indicated that approximately 200,000 individuals who had participated in special education exit the public school system each year. This highlights the fact that there is a substantial population of adults who, at least at some time, had a recognized intellectual impairment with a substantial impact on their achievement. Virtually everyone with more severe mental retardation will have been identified appropriately by the time they reach adulthood. In these cases, differential diagnosis should no longer be an issue for determination of SSI eligibility. However, the situation is very different for adults with milder forms of mental retardation and for younger adults.

Young adults who have recently turned 22 and left public high school should have recently been assessed to develop a transition plan to identify their work-related skills and to arrange for the involvement of another social service system once they depart public school. For example, regional centers were created in California to provide continuing services and supports. This system provides opportunities for state agencies responsible for follow-through services with these adults to evaluate each individual's status vis-à-vis SSI criteria for eligibility and establish a record that could be useful for SSA staff should an application for support be made at a future date.

For older adults with mild mental retardation, who have long since left high school, the situation is quite different. Many have probably vanished from the service delivery system, in part because the formal structure of educational settings that encourages systematic assessment is no longer available. Evaluation of academic skills and any other formal testing is unlikely to occur in the normal course of adult activities, and therefore there will be no opportunity to document intellectual impairment. This is especially problematic for older adults who come from families eligible for SSI support based on economic considerations, given their tendency to have limited access to services, as in the following case.

Marlene is a 36-year-old woman with mild mental retardation living in the community with her parents. Marlene completed high school 18 years ago after struggling through regular and special education classes. Transition activities were not yet part of the school curriculum, and therefore she did not receive any vocational, social, or daily living skills training. Once Marlene exited school, she remained at home with her parents *and received no outside assistance as she sought employment. Since then, Marlene's parents, despite their own economic hardships, have continued to take care of her everyday needs, so they never were concerned about her lack of independence. Marlene has floated in and out of sheltered workshop settings but never found her niche. Her parents are now reaching ages when they are increasingly concerned about who will care for their daughter when they are no longer able to meet her needs. Marlene does not have any marketable job skills, and although her parents are very worried, they are doing little at the present time to plan for their daughter's future.*

Adults with mental retardation who become economically independent of their families may become eligible for SSI supports themselves. If they have IQs below 60, these individuals would qualify for SSI support virtually without exception, as long as they have problems obtaining or maintaining employment. In these cases, the issues relevant to differential diagnosis are comparable to those mentioned for the school-age population with severe impairments. However, for more mildly affected individuals, there is a reasonable likelihood that they will have received a diagnosis other than mild mental retardation earlier on, and so they will not have established their eligibility for benefits under the mental retardation categorical listing. Especially careful assessment and consideration of functional equivalence will be required of the disability determination specialist to determine the eligibility of such an individual.

Considerations of psychiatric impairment or a dual diagnosis, as delineated fully in the section on adolescents, apply to adults with mental retardation as well. However, there are several additional challenges, particularly for older adults who left high school prior to 1990 and the initiation of the concept of transition planning. One is finding adults with disabilities once they have left the school system and no longer bear a school-based classification (e.g., mild mental retardation, learning disability). The second challenge is to ensure that these individuals get a complete reassessment in order to determine their eligibility for SSI. This is imperative, given evolving definitions of mental retardation and developmental disability, as well as classification practices (MacMillan et al., 1998). Finally, adult sons and daughters of non-English-speaking parents may not routinely come to the attention of disability examiners. If these individuals have not participated in special education during their school years, they may be invisible to the service systems until some unanticipated event forces the need for direct support.

ADDITIONAL CONSIDERATIONS

Cultural Issues

In developing the differential diagnoses for mental retardation, it is important to ensure that assessment instruments are culturally appropriate for the individual being evaluated (Sattler, 1988; Witt et al., 1998) and to clarify the relationship between culture and achievement as much as is feasible (Gallimore & Goldenberg, 2001). In addition to the influence of the majority culture and the English language, some standardized intelligence and adaptive behavior tests require social knowledge that may not be relevant in a given minority culture (Golden, 1990). Recent attempts to develop standardized tests of intelligence and adaptive behavior that are not culturally biased are described in Chapters 3 and 4.

In developing a differential diagnosis for mental retardation, particularly mild retardation, examiners need to determine whether evident psychopathological conditions are related to a specific cultural process. For example, Asian families may believe that emotional problems like anxiety bring shame and guilt to the family, and so may not disclose them (Cuellar & Paniagua, 2000). Paniagua (1998) has pointed out that, in some Hispanic cultures, mental or physical problems in a child may be thought to be the result of an extremely frightening or startling experience. These, as well as other attributions about the behaviors of their children with mental retardation (Chavira et al., 2000), can obscure the presence of a dual diagnosis or lead to a diagnosis other than mental retardation.

Clearly, examiners should be sensitive to the beliefs, attitudes, and practices related to a specific culture or religion that might be compatible or incompatible with assessment methods, employment interventions, and prioritization of goals (McAdoo, 1995; Rogers-Dulan & Blacher, 1995). In Hispanic families, for example, cultural values of family cohesion may conflict with policies promoting community inclusion and independent employment of persons with mental retardation, especially in integrated settings (Blacher et al., 1997).

Specific Disorders Associated with Mental Retardation

A complete differential diagnosis of mental retardation should address all possible disorders that share similar signs and symptoms with mental retardation. Consideration of the cause of an individual's intellectual disability may be important for treatment and counseling. Classification systems have emphasized both function (behavior) and etiology (biomedical) dimensions since the early 1960s (e.g., Heber, 1961). However, the importance of establishing the cause of disability has not been a major focus of this chapter because it is not one of SSA's priorities for this study. Since discussion of specific disorders associated with mental retardation may assist SSA eligibility determination specialists in evaluating individuals applying for SSI under the diagnosis of mental retardation, they are described briefly here.

Many of the issues that complicate functional and cognitive assessment or limit functional diagnostic precision, especially in infants, toddlers, and preschoolers, have already been discussed. These are serious concerns that could, at least in some cases, be addressed by refocusing attention to the underlying causes of mental retardation. There is currently an explosion of knowledge in the biomedical sciences, especially genetics and neuroscience, and this is allowing a constantly expanding diagnostic capability. Innovative technologies are now available to determine if conditions frequently associated with mental retardation are or are not present, and while their presence does not always result in substantial disability, failure to achieve economic self-sufficiency may be the rule rather than the exception. Therefore, as in the case of Down syndrome, SSA might consider including conditions with a high probability of causing mental retardation or some other substantial disability in its list of conditions eligible for SSI benefits. Of course, some affected individuals will achieve economic self-sufficiency and will therefore be ineligible, but if an individual with mental retardation encounters difficulties in obtaining or maintaining employment, it seems reasonable to presume that the underlying impairment contributed significantly to the situation.

Mental retardation can be associated with genetic as well as nongenetic causes. Therefore, we have provided a brief overview of these broad categories of conditions and illustrated the issues that can arise in SSA eligibility determination for two representative conditions, Prader-Willi syndrome (genetic) and fetal alcohol syndrome (nongenetic).

Genetic Conditions

There are over 750 known genetic causes of mental retardation. Collectively, these may account for as many as one-third of all cases of mental retardation (Matilainen et al., 1995; Opitz, 1996). Historically, identifiable genetic disorders were thought to cause severe to profound mental retardation, but more recent estimates suggest that 10 to 50 percent of mild mental retardation may also have a specific genetic origin (Rutter et al., 1996). Nonspecific genetic factors associated with individual differences in intelligence, another topic of considerable contemporary debate (e.g., Neisser, 1998), are not discussed here.

Among the relatively few disorders that have been studied to date, variability in impact on cognitive and adaptive functioning has been noted. Thus, IQ levels in individuals with disorders like fragile X syndrome, Williams syndrome, velo-cardio-facial syndrome and Prader-Willi syndrome can range widely (see Dykens et al., 2000 for a review). Like their counterparts with severe disabilities, however, high-functioning persons with these and other syndromes are likely to have characteristics that compromise their employability (e.g., intellectual impairments, specific maladaptive behaviors, or psychiatric disorders).

Prader-Willi syndrome is perhaps best known for its association with eating disorders. Whereas affected babies invariably have pronounced feeding difficulties, they typically develop hyperphagia and food-seeking behaviors (e.g., foraging and hoarding) during early childhood (Dykens & Cassidy, 1999). Without lifelong dietary management, affected individuals invariably become morbidly obese and weight-associated complications remain the leading cause of death for people with this syndrome.

While the vast majority of people with Prader-Willi syndrome are invariably preoccupied with eating and food, some also show a host of nonfood obsessions and compulsive behaviors, suggesting increased risk of full-blown obsessive-compulsive disorder in this population (Dykens et al., 1996). Persons with Prader-Willi syndrome also have adaptive skills that typically fall considerably below their measured IQ (Dykens et al., 1992). Compared with others with mental retardation, children and adults with Prader-Willi syndrome show high rates of temper tantrums, aggression, stubbornness, underactivity, excessive daytime sleepiness, and emotional lability (Dykens & Kasari, 1997; Dykens & Smith, 1998; State et al., 1999). Approximately 32 percent of people with the disorder have IQs of 70 or higher. Nevertheless, higher IQ individuals are extremely likely to have poor adaptive behavior and show the same behavioral and food-related vulnerabilities as their peers with lower IQs. Regardless of their IQs, persons with Prader-Willi syndrome share the same needs for intensive support. In the employment arena, most adults with Prader-Willi syndrome are not employed competitively, and low job retention is invariably associated with aggression, tantrums, and compulsive food-seeking (Dykens, 2002). In light of these findings, several states have now passed legislation that deems any person with a confirmed genetic diagnosis of Prader-Willi syndrome eligible for lifelong state services *regardless* of his or her IQ. This may be an effective model for the SSA classification process to adopt.

Nongenetic Conditions

Mental retardation can also be associated with a variety of causes that are nongenetic in origin. Among them are: (a) infections and parasitic diseases (e.g., congenital rubella syndrome, encephalitis), (b) deficiency diseases (e.g., congenital iodine deficiency syndrome), (c) congenital malformations (e.g., hydrocephalus), (d) neuroteratogenic exposure (e.g., fetal alcohol syndrome), (e) mechanical injury (e.g., perinatal hypoxia, traumatic brain injury), and (f) environmental deprivation. Of course, this is not an exhaustive listing, and for this discussion it is important to note only that each of these conditions can vary in its functional impact. In some cases, there may be minimal or no long-term consequences of the condition, while profound and permanent impairments may be observed in others. When the presence of one of these conditions is documented in an individual's medical history, it can make an important contribution to the disability determination process. In many cases, an etiological diagnosis can provide a valid basis for determining if economic hardship is associated with the presence of a disability or some other circumstance. The case of prenatal alcohol exposure can serve to illustrate this point.

Fetal alcohol syndrome and alcohol-related neurodevelopmental disorders result from maternal consumption of alcohol during pregnancy. These disorders represent one of the leading preventable causes of birth defects, mental retardation, and neurodevelopmental disorders. Fetal alcohol syndrome is estimated to occur in 5.2 per 10,000 live births in the United States (Centers for Disease Control and Prevention, 1997; Cordero et al., 1994). While some individuals may be more vulnerable, perhaps due to their genetic makeup, the immediate cause is clearly a nongenetic factor.

Fetal alcohol syndrome results in characteristic facial features, including microcephaly, short palpebral fissures, thin upper lip, flat philtrum and midface, as well as intellectual impairment together with behavioral, emotional, and tone abnormalities. In one study (Streissguth et al., 1996), up to 94 percent of children and young adults with the syndrome had mental health problems, including 61 percent with attention deficit hyperactivity disorder, over 50 percent with depression, and 29 percent with psychotic symptoms. Streissguth et al. also reported a mean IQ for their sample of children and young adults of 79, while adaptive skills were more than two standard deviations below the reference population mean.

Prenatal alcohol exposure can produce a spectrum of disorders including: (a) fetal alcohol syndrome, (b) fetal alcohol syndrome without confirmed exposure to alcohol, (c) partial fetal alcohol syndrome with confirmed exposure, (d) alcohol-related birth defects, or (e) alcohol-related neurodevelopmental disorder. The severity of impairment can also vary widely from case to case, with some individuals having severe and easily recognizable signs and symptoms and others having more subtle effects that can be difficult to recognize. In addition, when diagnosis occurs after infancy or early childhood, it can be difficult to document the linkage between symptoms and prenatal alcohol exposure convincingly, especially in individuals with milder impairments. Despite these difficulties, once signs and symptoms are evident and a diagnosis is made, it is likely that fetal alcohol syndrome (or one of the other conditions associated with prenatal alcohol exposure) will be a major contributing factor to substantial difficulties in school achievement, development of social skills, and maintenance of employment. Of course, mental retardation can be evident in many individual cases, but even in borderline situations, the other behavioral and functional impairments associated with prenatal alcohol exposure can safely be assumed to limit opportunities for achieving economic self-sufficiency. In these cases, SSA might consider the etiological diagnosis in arriving at its determination of eligibility for benefits.

CONCLUSIONS AND RECOMMENDATIONS

In reviewing the literature on the differential diagnosis of mental retardation from other disorders with similar signs and symptoms, the committee concludes that it is necessary to use multiple sources of objective data, including but not limited to, IQ test results, assessments of adaptive behaviors, academic test results, and structured interviews, to determine whether an individual meets diagnostic criteria for mental retardation. Many neurodevelopmental and psychiatric disorders, such as autism, pervasive developmental disorder, attention deficit hyperactivity disorder, cerebral palsy, sensory impairments, and seizure disorders, are associated with mental retardation and complicate its diagnosis. Many of these disorders are currently defined as making the client automatically eligible for SSI benefits. Individuals who have mild mental retardation may be eligible for SSI, on the basis of concurrent neurodevelopmental or psychiatric disorders.

The committee also concludes that accurate evaluation of infants and very young children requires special attention to selecting and using appropriate evaluation instruments, conditions of test administration, experience of the evaluator, and concurrent medical and psychosocial risk factors.

The committee concludes that learning disability in school-age children can be differentiated from mental retardation by examining data from cognitive, adaptive, and academic testing. Children who have learning disabilities rather than mental retardation have significantly lower academic performance than their peers, but they do not meet criteria for mental retardation—i.e., they have composite IQs over 70 and performance difficulties focused in specific domains of cognitive and academic content. Like learning disability in school-age children, borderline intellectual functioning may be differentiated from mental retardation by reviewing cognitive and adaptive functioning evaluations. Children with composite IQ scores of 71-75 without the adaptive behavior deficits required for a diagnosis of mental retardation can be classified as having borderline intellectual functioning. It is the combination of composite IQs and adaptive behavior deficits as described in Chapter 5 that provides the best operational definition for mental retardation.

Research data show that adolescents and young adults with mental retardation are at high risk for developing a range of behavior disorders or psychiatric impairments. In that case, they are more accurately described as having a dual diagnosis. Behavioral or mental health disorders can further complicate the differential diagnosis of mental retardation, frequently depressing scores on tests of cognitive or adaptive functioning. In these cases, reevaluation of the individual is warranted after treatment of the mental health disorder.

Research data also make it clear that cultural and socioeconomic background has profound effects on the assessment and the interpretation of behavior, particularly by proxy respondents. The disability examiner must take this into account as he or she prepares the client's assessment, using other sources of data as needed for an accurate diagnosis.

On the basis of the data reviewed in this and previous chapters, the committee makes the following recommendations related to distinguishing mental retardation from other conditions with which it shares signs and symptoms.

Recommendation: Social Security Disability Determination Specialists may differentiate other conditions from mental retardation by using intelligence and adaptive behavior test criteria as outlined in the committee's recommendations. Data, including school test results, intelligence and adaptive behavior test results, and psychiatric and medical test results, from community-based agencies (such as schools, hospitals, or clinics) can be used to inform the determination of SSA eligibility for the diagnosis of mental retardation, but the diagnoses given by community agencies should not be used.

- **Social Security Disability Determination Specialists may differentiate individuals with borderline intellectual functioning and learning disability from those with mental retardation by reviewing cognitive and adaptive behavior test results and determining whether the individual meets diagnostic criteria for mental retardation as recommended in Chapters 3 and 4.**
- **Social Security Disability Determination Specialists do not**

need to determine the presence or absence of mental retardation in individuals who are eligible for SSI due to other neurodevelopmental or psychiatric disabilities (e.g., autism, pervasive developmental disorder, attention deficit hyperactivity disorder, genetic syndromes, intrauterine exposure to alcohol or environmental toxins, sensory impairments, seizure disorders, or severe emotional-behavioral disorders).

- **Objective data on intellectual and adaptive functioning to determine mental retardation should be collected for individuals with mild neurodevelopmental or psychiatric disabilities who might have impairments that are consistent with or functionally equivalent to mental retardation.**

References

Achenbach, T.M., and Edelbrock, C.S. (1978). The classification of child psychopathology: A review and analysis of empirical efforts. *Psychological Bulletin, 85*(6), 1275-1301.

Ackerman-Ross, S., and Khanna, P. (1989). The relationship of high quality day care to middle-class 3-year-old's language performance. *Early Childhood Research Quarterly, 4*(1), 97-116.

Adams, G.L. (2000). *CTAB-R and NABC-R technical manual*. Seattle, WA: Educational Achievement Systems.

Adams, G.L., and Hartleben, J. (1984). *Comprehensive test of adaptive behavior*. Columbia, OH: C.E. Merrill.

Adams, K., and Markham, R. (1991). Recognition of affective facial expressions by children and adolescents with and without mental retardation. *American Journal on Mental Retardation, 96*(1), 21-28.

Allen, R.M., Cortazzo, A.D., and Adamo, C. (1970). Factors in an adaptive behavior check list for use with retardates. *Training School Bulletin, 67*(3), 144-157.

American Association on Mental Retardation. (1992). *Mental retardation: Definition, classification, and systems of support* (9th ed.). Washington, DC: American Association on Mental Retardation.

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

Anastasi, A., and Urbina, S. (1997). *Psychological testing* (7th ed.). Upper Saddle River, NJ: Prentice-Hall.

Angoff, W.H. (1993). Perspectives on differential item functioning methodology. In P.W. Holland, and H. Wainer (eds.), *Differential item functioning* (pp. 3-23). Hillsdale, NJ: Lawrence Erlbaum Associates.

Archer, R.P., Maruish, M., Imhof, E.A., and Piotrowski, C. (1991). Psychological test usage with adolescent clients: 1990 survey findings. *Professional Psychology: Research and Practice, 22*(3), 247-252.

Arnold, L.E., Aman, M.G., Martin, A., Collier-Crespin, A., Vitiello, B., Tierney, E., Asarnow, R.F., Bell-Bradshaw, F., Freeman, B.J., Gates-Ulanet, P., Klin, A., McCracken, J.T., McDougle, C.J., McGough, J.J., Posey, D.J., Scahill, L., Swiezy, N.B., Ritz, L., and Volkmar, F.R. (2000). Assessment of multisite randomized clinical trials of patients with autistic disorder: The Autism RUPP Network. *Journal of Autism and Developmental Disorders, 30*(2), 99-111.

Arvey, R.D. (1986). General ability in employment: A discussion. *Journal of Vocational Behavior, 29*(3), 415-420.

Asch, S.E. (1946). Forming impressions of personality. *Journal of Abnormal and Social Psychology, 41*, 258-290.

Auty, W.P., Goodman, J., and Foss, G. (1987). The relationship between interpersonal competence and work adjustment. *Vocational Evaluation and Work Adjustment Bulletin, 20*(2), 49-52.

Bailey, D.B., Aytch, L.S., Odom, S.L., Symons, F., and Wolery, M. (1999). Early intervention as we know it. *Mental Retardation and Development Disabilities Research Reviews, 5*(1), 11-20.

Bailey, D.B., Hatton, D.D., and Skinner, M. (1998). Early developmental trajectories of males with fragile X syndrome. *American Journal on Mental Retardation, 103*(1), 29-39.

Baker, B.L., Blacher, J., and Pfeiffer, S.I. (1993). Family involvement in residential treatment of children with psychiatric disorder and mental retardation. *Hospital and Community Psychiatry, 44*(6), 561-566.

Baker, B.L., Blacher, J., and Pfeiffer, S.I. (1996). Family involvement in residential treatment. *American Journal on Mental Retardation, 101*(1), 1-14.

Balboni, G., Battagliese, G., and Pedrabissi, L. (2000). The psychopathology inventory for mentally retarded adults: Factor structure and comparisons between subjects with or without dual diagnosis. *Research in Developmental Disabilities, 21*(4), 311-321.

Baldwin, M.L., and Johnson, W.G. (1998). Dispelling the myths about work disability. In T. Thomason, J.F. Burton, and D. Hyatt (eds.), *New approaches to disability in the workplace* (1st ed.). Madison, WI: Industrial Relations Research Association.

Balthazar, E.E., and English, G.E. (1969). A factorial study of unstructured ward behaviors. *American Journal of Mental Deficiency, 74*(3), 353-360.

Baroff, G.S. (1991). *Developmental disabilities: Psychological aspects*. Austin, TX: Pro-Ed.

Batshaw, M.L., and Perret, Y.M. (1992). Hearing. In M.L. Batshaw, and Y.M. Perret (eds.), *Children with disabilities: A medical primer* (3rd ed., pp. 321-349). Baltimore, MD: Paul H. Brookes.

Bayley, N. (1993). *Bayley scales of infant development: Manual* (2nd ed.). New York: Psychological Corporation.

Beck, F.W., Spurlock, D.W., and Lindsey, J.L. (1988). The WISC-R and the Woodcock-Johnson tests of achievement: Correlations for exceptional children. *Perceptual and Motor Skills, 67*(2), 587-590.

Bell-Dolan, D.J., and Allan, W.D. (1998). Assessing elementary school children's social skills: Evaluation of the parent version of the matson evaluation of social skills with youngsters. *Psychological Assessment, 10*, 140-149.

Biasini, F.J., Grupe, L., Huffman, L., and Bray, N.W. (1999). Mental retardation: A symptom and a syndrome. In S.D. Netherton, D.L. Holmes, and C.E. Walker (eds.), *Comprehensive textbook of child and adolescent disorders*. New York: Oxford University Press.

Binet, A., and Simon, T. (1916). New methods for the diagnosis of the intellectual level of subnormals. In A. Binet, T. Simon, and E.S. Kite, *The development of intelligence in children (the Binet-Simon scale)* (pp. 37-90). Baltimore, MD: Williams and Wilkins Company.

Blacher, J. (1982). Assessing social cognition in young mentally retarded and nonretarded children. *American Journal of Mental Deficiency, 86*, 473-484.

Blacher, J. (2001). Transition to adulthood: Mental retardation, families, and culture. *American Journal on Mental Retardation, 106*(2), 173-188.

Blacher, J., Lopez, S.R., Shapiro, J., and Fusco, J. (1997). Contributions to depression in Latina mothers with and without children with retardation: Implications for caregiving. *Family Relations, 46*(4), 325-334.

Blake, D.D., and Andrasik, F. (1986). Between-measure consistency in social-skills assessment. *Journal of Psychopathology and Behavioral Assessment, 8*, 227-240.

Bloom, B.S. (1964). *Stability and change in human characteristics*. New York: Wiley.

Bolla, K.I., Lindgren, K.N., Bonaccorsy, C., and Bleecker, M.L. (1990). Predictors of verbal fluency (FAS) in the healthy elderly. *Journal of Clinical Psychology, 46*(5), 623-628.

Borkowski, J.G. (1985). Signs of intelligence: Strategy generalization and metacognition. In S.R. Yussen (ed.), *The growth of reflection in children* (pp. 105-144). Orlando, FL: Academic Press.

Bornstein, R.F. (1996). Face validity in psychological assessment: Implications for a unified model of validity. *American Psychologist, 51*(9), 983-984.

Borthwick-Duffy, S.A. (1994). Prevalence of destructive behaviors: A study of aggression, self-injury, and property destruction. In T. Thompson, and D.B. Gray (eds.), *Destructive behavior in developmental disabilities: Diagnosis and treatment* (Vol. 170, pp. 3-23). Thousand Oaks, CA: Sage Publications.

Borthwick-Duffy, S.A. (2001). Assessment of adaptive behavior. Unpublished paper written for the National Research Council Committee on Disability Determination for Mental Retardation. Graduate School of Education, University of California, Riverside.

Borthwick-Duffy, S.A., and Eyman, R.K. (1990). Who are the dually diagnosed? *American Journal on Mental Retardation, 94*(6), 586-595.

Borthwick-Duffy, S.A., Lane, K.L., and Widaman, K.F. (1997). Measuring problem behaviors in children with mental retardation: Dimension and predictors. *Research in Developmental Disabilities, 18*(6), 415-433.

Boyd, R.D. (1989). What a difference a day makes: Age related discontinuities and the Battelle developmental inventory. *Journal of Early Intervention, 13*(2), 114-119.

Boyle, C.A., Yeargin-Allsop, M., Doernberg, N.S., Holmgreen, P., Murphy, C.C., and Schendel, D.E. (1996). Prevalence of selected developmental disabilities in children 3-10 years of age: The metropolitan Atlanta developmental disabilities surveillance program, 1991. *Morbidity and Mortality Weekly Report. CDC Surveillance Summaries, 45*(2), 1-14.

Bracken, B.A. (1984). *Bracken basic concept scale*. San Antonio, TX: Psychological Corporation.

Bracken, B.A. (1987). Limitations of preschool instruments and standards for minimal levels of technical adequacy. *Journal of Psychoeducational Assessment, 5*(4), 313-326.

Bracken, B.A. (1988). Ten psychometric reasons why similar tests produce dissimilar results. *Journal of School Psychology, 26*(2), 155-166.

Bracken, B.A. (1998). *Bracken basic concept scale - Revised*. San Antonio, TX: Psychological Corporation.

Bracken, B.A. (2000). Maximizing construct relevant assessment: The optimal preschool testing situation. In B.A. Bracken (ed.), *The psychoeducational assessment of preschool children* (3rd ed., pp. 33-44). Needham Heights, MA: Allyn and Bacon.

Bracken, B.A., and McCallum, R.S. (1998). *Universal Nonverbal Intelligence Test (UNIT)*. Itasca, IL: Riverside Publishing.

Bracken, B.A., Howell, K.K., and Crain, R.M. (1993). Prediction of Caucasian and African-American preschool children's fluid and crystallized intelligence: Contributions of maternal characteristics and home environment. *Journal of Clinical Child Psychology, 22*(4), 455-463.

Bracken, B.A., Prasse, D.P., and McCallum, R.S. (1984). Peabody Picture Vocabulary Test - Revised: An appraisal and review. *School Psychology Review, 13*(1), 49-60.

Braddock, D.L., Hemp, R., Parish, S.L., and Rizzolo, M.C. (2000). *The state of the states in developmental disabilities: 2000 study summary*. Chicago: University of Illinois at Chicago, Department of Disability and Human Development.

Brask, B.H. (1970). A prevalence investigation of childhood psychosis. In *Nordic symposium on the care of psychotic children* (pp. 45-153). Oslo, Norway: Barnepsychiatrist Forening, Universitetsforlagets Trykningssentral.

Brockley, J.A. (1999). History of mental retardation: An essay review. *History of Psychology, 2*(1), 25-36.

Brosgole, L., Gioia, J.V., and Zingmond, R. (1986). Facial-affect and postural-affect recognition in the mentally handicapped and normal young children. *International Journal of Neuroscience, 30*(1-2), 127-144.

Brown, L., and Leigh, J.E. (1986). *Adaptive behavior inventory*. Austin, TX: Pro-Ed.

Brown, L., Sherbenou, R.J., and Johnsen, S.K. (1997). *Test of Nonverbal Intelligence: A language-free measure of cognitive ability* (3rd ed.). Austin, TX: Pro-Ed.

Bruininks, R.H., Woodcock, R.W., Weatherman, R.F., and Hill, B.K. (1984). *Scales of independent behavior*. Allen, TX: DLM Teaching Resources.

Bruininks, R.H., Woodcock, R.W., Weatherman, R.F., and Hill, B.K. (1996). *SIB-R: Scales of independent behavior-Revised*. Chicago: Riverside Publishing.

Bruininks, R.H., Woodcock, R.W., Weatherman, R.F., and Hill, B.K. (2000). *Scales of independent behavior - revised*. Itasca, IL: Riverside Publishing.

Bryant, B.R., Bryant, D.P., and Chamberlain, S. (1999). Examination of gender and race factors in the assessment of adaptive behavior. In R.L. Schalock, and D.L. Braddock (eds.), *Adaptive behavior and its measurements: Implications for the field of mental retardation* (pp. 141-160). Washington, DC: American Association on Mental Retardation.

Bullis, M., and Foss, G. (1986). Assessing the employment-related interpersonal competence of mildly mentally retarded workers. *American Journal of Mental Deficiency, 91*(1), 43-50.

Bunn, D. (1992). Synthesis of expert judgment and statistical forecasting models for decision support. In G. Wright, and F. Bolger (eds.), *Expertise and decision support* (pp. 251-268). New York: Plenum Press.

Burd, L., Fisher, W., and Kerbeshian, J. (1987). A prevalence study of pervasive developmental disorders in North Dakota. *Journal of the American Academy of Child and Adolescent Psychiatry, 26*(5), 700-703.

Burke, B.G., Sauser, W.I., Kemery, E.R., and Dyer, F.N. (1989). Intelligence and physical fitness as predictors of success in early infantry training. *Perceptual and Motor Skills, 69*(1), 263-271.

Burkhauser, R.V. (1998). Policies to make work pay for people with disabilities. In K. Rupp, and D.C. Stapleton (eds.), *Growth in disability benefits: Explanations and policy implications*. Kalamazoo, MI: Upjohn Institute for Employment Research.

Burkhauser, R.V., Daly, M.C., and Houtenville, A.J. (2001). How working-age people with disabilities fared over the 1990s business cycle. In P.P. Budetti, R.V. Burkhauser, J.M. Gregory, and H.A. Hunt (eds.), *Ensuring health and income security for an aging workforce*. Kalamazoo, MI: Upjohn Institute for Employment Research.

Burt, C. (1909). Experimental tests of general intelligence. *British Journal of Psychology, 3*, 94-177.

Burt, C. (1949). The structure of the mind: A review of the results of factor analysis. *British Journal of Educational Psychology, 19*, 176-199.

Butterworth, J., and Strauch, J.D. (1994). The relationship between social competence and success in the competitive workplace for persons with mental retardation. *Education and Training in Mental Retardation and Developmental Disabilities, 29*(2), 118-133.

Campbell, D.T., and Fiske, D.W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*, 81-105.

Campbell, D.T., Stanley, J.C., and Gage, N.L. (1963). *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally.

Campbell, M., and Malone, R.P. (1991). Mental retardation and psychiatric disorders. *Hospital and Community Psychiatry, 42*(4), 374-379.

Campbell, S.K., Siegel, E., Parr, C.A., and Ramey, C.T. (1986). Evidence for the need to renorm the Bayley Scales of Infant Development based on the performance of a population-based sample of 12-month-old infants. *Topics in Early Childhood Special Education, 6*(2), 83-96.

Campione, J.C., and Brown, A.L. (1978). Toward a theory of intelligence: Contributions from research with retarded children. *Intelligence, 2*(3), 279-304.

Caplan, R., Guthrie, D., and Foy, J.G. (1992). Communication deficits and formal thought disorder in schizophrenic children. *Journal of the American Academy of Child and Adolescent Psychiatry, 31*(1), 151-159.

Carrey, N.J. (1995). Itard's 1828 memoire on "Mutism caused by a lesion of the intellectual functions": A historical analysis. *Journal of the American Academy of Child and Adolescent Psychiatry, 34*(12), 655-1661.

Carroll, J.B. (1993). *Human cognitive ablilities: A survey of factor-analytic studies*. New York: Cambridge University Press.

Carroll, J.B. (1997). Psychometrics, intelligence, and public perception. *Intelligence, 24*(1), 25-52.

Castles, E.E., and Glass, C.R. (1986). Empirical generation of measures of social competence for mentally retarded adults. *Behavioral Assessment, 8*(4), 319-330.

Cattell, P. (1940). *The measurement of intelligence of infants and young children*. New York: Psychological Corporation.

Cattell, R.B. (1941). Some theoretical issues in adult intelligence testing. *Psychological Bulletin, 38*, 592.

Cattell, R.B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology, 54*(1), 1-22.

Cattell, R.B. (1971). *Abilities: Their structure, growth, and action*. Boston: Houghton Mifflin.

Cattell, R.B. (1986). Scales and the meaning of standardized scores. In R.B. Cattell, and R.C. Johnson (eds.), *Functional psychological testing: Principles and instruments* (pp. 79-104). New York: Brunner/Mazel.

Centers for Disease Control and Prevention. (1997). Surveillance for fetal alcohol syndrome using multiple sources: Atlanta, Georgia, 1981-1989. *Morbidity and Mortality Weekly Report, 46*(47), 1118-1120.

Chadsey-Rusch, J. (1992). Toward defining and measuring social skills in employment settings. *American Journal on Mental Retardation, 96*(4), 405-418.

Chan, K.Y., Drasgow, F., and Sawin, L.L. (1999). What is the shelf life of a test? The effect of time on the psychometrics of a cognitive ability test battery. *Journal of Applied Psychology, 84*(4), 610-619.

Chattin, S.H., and Bracken, B.A. (1989). School psychologists' evaluation of the K-ABC, McCarthy scales, Stanford-Binet IV, and WISC-R. *Journal of Psychoeducational Assessment, 7*(2), 112-130.

Chavira, V., Lopez, S.R., Blacher, J., and Shapiro, J. (2000). Latina mothers' attributions, emotions, and reactions to the problem behaviors of their children with developmental disabilities. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 41*(2), 245-252.

Chirikos, T.N. (1995). Medicare and the Social Security Disability Insurance program. *Health Affairs, 14*(4), 244-252.

Cialdella, P., and Mamelle, N. (1989). An epidemiological study of infantile autism in a

French department (Rhone): A research note. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 30*(1), 165-175.

Cicchetti, D.V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment, 6*(4), 284-290.

Clark, L.A., and Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment, 7*(3), 309-319.

Clausen, J. (1972). Quo vadis, AAMD? *Journal of Special Education, 6*(1), 51-60.

Clinger, M.D., Fine, M.A., Johnson, J., Schwartzman, J., and Drude, K.P. (1988). Assessment of the psychometric properties of the Adaptive Behavior Scale with psychiatric patients. *Journal of Consulting and Clinical Psychology, 56*(4), 542-548.

Committee on Classification of Feeble-Minded. (1910). *Journal of Psycho-Asthenics, 15*, 61-67.

Cook, T.D., and Campbell, D.T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Chicago: Rand McNally.

Coombs, C.H., Dawes, R.M., and Tversky, A. (1970). *Mathematical psychology: An elementary introduction*. Englewood Cliffs, NJ: Prentice-Hall.

Corbett, J., and Barton, L. (1992). *A struggle for choice: Students with special needs in transition to adulthood*. London: Routledge.

Cordero, J.F., Floyd, R.L., Martin, M.L., Davis, M., et al. (1994). Tracking the prevalence of FAS. *Alcohol Health and Research World, 18*(1), 82-85.

Craig, E.M., and Tasse, M.J. (1999). Cultural and demographic group comparisons in adaptive behavior. In R.L. Schalock, and D.L. Braddock (eds.), *Adaptive behavior and its measurements: Implications for the field of mental retardation* (pp. 119-140). Washington, DC: American Association on Mental Retardation.

Crick, N.R., and Dodge, K.A. (1994). A review and reformulation of social information-processing mechanisms in children's social adjustment. *Psychological Bulletin, 115*(1), 74-101.

Crocker, L.M., and Algina, J. (1986). *Introduction to classical and modern test theory*. New York: Holt, Rinehart, and Winston.

Cronbach, L.J. (1949). *Essentials of psychological testing*. New York: Harper and Brothers.

Cronbach, L.J. (1990). *Essentials of psychological testing* (5th ed.). New York: Harper and Row.

Cuellar, I., and Paniagua, F.A. (2000). *Handbook of multicultural mental health*. San Diego, CA: Academic Press.

Cunningham, J.A., Dollinger, S.J., Satz, M., and Rotter, N.S. (1991). Personality correlates of prejudice against AIDS victims. *Bulletin of the Psychonomic Society, 29*(2), 165-167.

Currie, J., and Gruber, J. (1996a). Health insurance eligibility, utilization of medical care, and child health. *Quarterly Journal of Economics, 111*(2), 431-466.

Currie, J., and Gruber, J. (1996b). Saving babies: The efficacy and cost of recent changes in the medicaid eligibility of pregnant women. *Journal of Political Economy, 104*(6), 1263-1296.

Daniels, S.M., and West, J. (1998). Return to work for SSI and DI beneficiaries: Employment policy challenges. In K. Rupp, and D.C. Stapleton (eds.), *Growth in* *disability benefits: Explanations and policy implications* (pp. 359-364). Kalamazoo, MI: Upjohn Institute for Employment Research.

Das, J.P., Naglieri, J.A., and Kirby, J.R. (1994). *Assessment of cognitive processes: The PASS theory of intelligence*. Needham Heights, MA: Allyn and Bacon.

de Jung, J.E., Holen, M.C., and Edmonson, B. (1973). Test of social inference for retarded adolescents: Measuring social-cue perception. *Psychological Reports, 32*(2), 603-618.

Dekker, R., and Koole, F.D. (1992). Visually impaired children's visual characteristics and intelligence. *Developmental Medicine and Child Neurology, 34*(2), 123-133.

Denning, C.B., Chamberlain, J.A., and Polloway, E.A. (2000). An evaluation of state guidelines for mental retardation: Focus on definition and classification practices. *Education and Training in Mental Retardation and Developmental Disabilities, 35*(2), 226-232.

Denno, D.W. (1986). Victim, offender, and situational characteristics of violent crime. *Journal of Criminal Law and Criminology, 77*(4), 1142-1158.

Dodge, K.A. (1986). A social information processing model of social competence in children. In M. Perlmutter (ed.), *Minnesota symposium in child psychology* (pp. 77-125). Hillsdale, NJ: Lawrence Erlbaum Associates.

Doll, E.A. (1927). Borderline diagnosis. *Proceedings of the American Association for the Study of the Feebleminded, 32*, 45-59.

Doll, E.A. (1936a). Current thoughts on mental deficiency. *Proceedings of the American Association on Mental Deficiency, 41*, 33-49.

Doll, E.A. (1936b). *The Vinelend social maturity scale* (No. 3). Vineland, NJ: Training School.

Doll, E.A. (1953). *The measurement of social competence. A manual for the Vineland Social Maturity Scale*. Washington, DC: Educational Test Bureau.

Doll, E.A. (1967). Historical review of mental retardation 1800-1965: A symposium. *American Journal of Mental Deficiency, 72*, 165-189.

Doman, A. (1967). Appraisal of emotionally incongrous situations by mental retardates of differing mental ages. *The Journal of General Psychology, 77*(1), 131-143.

Donkar, E.N. (2000). The Social Security Disability Insurance Program: Projections for the coming decade. Paper presented at the National Academy of Social Insurance Conference. L. Frieden, and J.L. Mashaw, (eds.), *Disability income policy: Opportunities and challenges in the next decade - A policy education seminar.* December 15, 2000.

Drotar, D., and Sturm, L. (1988). Prediction of intellectual development in young children with early histories of nonorganic failure-to-thrive. *Journal of Pediatric Psychology, 13*(2), 281-296.

DuBois, P.H. (1970). *A history of pschological testing*. Boston: Allyn and Bacon.

Dunn, L.M. (1959). *Peabody picture vocabulary test*. Circle Pines, MN: American Guidance Service.

Dunn, L.M., and Dunn, L.M. (1981). *Peabody picture vocabulary test - Revised* (Special ed.). Circle Pines, MN: American Guidance Service.

Duran, R.P. (1989). Testing of linguistic minorities. In R.L. Linn (ed.), *Educational measurement* (3rd ed., pp. 573-587). New York: Macmillan.

Dykens, E.M. (in press). Anxiety and fears in Williams syndrome. *Developmental Neuropsychology*.

Dykens, E.M. (2002). Are jigsaw puzzle skills "spared" in children with Prader-Willi syndrome? *Journal of Child Psychology and Psychiatry 43*(3), 343-352.

Dykens, E.M., and Cassidy, S.B. (1999). Prader-Willi syndrome. In S. Goldstein, and C.R. Reynolds (eds.), *Handbook of neurodevelopmental and genetic disorders in children* (pp. 525-554). New York: Guilford Press.

Dykens, E.M., and Kasari, C. (1997). Maladaptive behavior in children with Prader-Willi syndrome, Down syndrome, and nonspecific mental retardation. *American Journal on Mental Retardation, 102*(3), 228-237.

Dykens, E.M., and Smith, A.C.M. (1998). Distinctiveness and correlates of maladaptive behavior in children and adolescents with Smith-Magenis syndrome. *Journal of Intellectual Disability Research, 42*(6), 481-489.

Dykens, E.M., Hodapp, R.M., and Finucane, B.M. (2000). *Genetics and mental retardation syndromes: A new look at behavior and interventions*. Baltimore, MD: Paul H. Brookes.

Dykens, E.M., Hodapp, R.M., Walsh, K., and Nash, L.J. (1992). Adaptive and maladaptive behavior in Prader-Willi syndrome. *Journal of the American Academy of Child and Adolescent Psychiatry, 31*(6), 1131-1136.

Dykens, E.M., Leckman, J.F., and Cassidy, S.B. (1996). Obsessions and compulsions in Prader-Willi syndrome. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 37*(8), 995-1002.

Eaton, L.F., and Menolascino, F.J. (1982). Psychiatric disorders in the mentally retarded: Types, problems, and challenges. *American Journal of Psychiatry, 139*(10), 1297-1303.

Edgerton, R.B. (1993). *The cloak of competence* (revised ed.). Berkeley, CA: University of California Press.

Editorial Board. (1996). Definition of mental retardation. In J.W. Jacobson, and J.A. Mulick (eds.), *Manual of diagnosis and professional practice in mental retardation* (pp. 13-53). Washington, DC: American Psychological Association.

Einfeld, S.L. (1992). Clinical assessment of psychiatric symptoms in mentally retarded children. *Australian New Zealand Journal of Psychiatry, 26*(1), 48-63.

Einfeld, S.L., and Tonge, B.J. (1996). Population prevalence of psychopathology in children and adolescents with intellectual disability: II. Epidemiological findings. *Journal of Intellectual Disability Research, 40*(2), 99-109.

Elias, M.J., Zins, J.E., Weissberg, R.P., Frey, K.S., Greenberg, M.T., Haynes, N.M., Kessler, R., Schwab-Stone, M.E., and Shriver, T.P. (1997). *Promoting social and emotional learning: Guidelines for educators*. Alexandria, VA: Association for Supervision and Curriculum Development.

Elliott, C.D. (1990). *Differential Abillity Scales: Introductory and technical handbook*. San Antonio, TX: Psychological Corporation.

Embretson, S.E., and Hershberger, S.L. (1999). *The new rules of measurement: What every psychologist and educator should know*. Mahwah, NJ: Lawrence Erlbaum Associates.

Evans, J. (1989). *Bias in human reasoning: Causes and consequences*. London: Erlbaum.

Evans, L.D., and Bradley-Johnson, S. (1988). A review of recently developed measures of adaptive behavior. *Psychology in the Schools, 25*(3), 276-287.

Faas, L.A., and D'Alonzo, B.J. (1990). WAIS-R scores as predictors of employment success and failure among adults with learning disabilities. *Jourrnal of Learning Disabilities, 23*(5), 311-316.

Fidler, D.J., Hodapp, R.M., and Dykens, E.M. (2000). Stress in families of young children with Down syndrome, Williams syndrome, and Smith-Magenis syndrome. *Early Education and Development, 11*(4), 395-406.

Floyd, F.J., and Widaman, K.F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment, 7*(3), 286-299.

Flynn, J.R. (1984). The mean IQ of Americans: Massive gains 1932 to 1978. *Psychological Bulletin, 95*(1), 29-51.

Flynn, J.R. (1987). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological Bulletin, 101*(2), 171-191.

Flynn, J.R. (1994). IQ gains over time. In R.J. Sternberg (ed.), *The encyclopedia of human intelligence* (pp. 617-623). New York: Macmillan.

Flynn, J.R. (1999). Searching for justice: The discovery of IQ gains over time. *American Psychologist, 54*(1), 5-20.

Fombonne, E. (1997). The prevalence of autism and other pervasive developmental disorders in the UK. *Autism, 1*, 227-229.

Fombonne, E. (1999). The epidemiology of autism: A review. *Psychological Medicine, 29*(4), 769-786.

Fombonne, E., and du Mazaubrun, C. (1992). Prevalence of infantile autism in four French regions. *Social Psychiatry and Psychiatric Epidemiology, 27*(4), 203-210.

Forness, S.R., Keogh, B.K., MacMillan, D.L., Kavale, K.A., and Gresham, F.M. (1998). What is so special about IQ? The limited explanatory power of cognitive abilities in the real world of special education. *Remedial and Special Education, 19*(6), 315-322.

Foss, G., and Bostwick, D. (1981). Problems of mentally retarded adults: A study of rehabilitation service consumers and providers. *Rehabilitation Counseling Bulletin, 25*(2), 66-73.

Foss, G., Cheney, D., and Bullis, M. (1986). *TICE: Test of Interpersonal Competence for Employment Manual*. Santa Barbara, CA: James Stanfield and Company.

Foster, A. (1988). Childhood blindness. *Eye, 2*, S27-S36.

Fox, D.R. (1994). Observations on disability evaluation in the Social Security Administration. *Journal of Social Behavior and Personality, 9*(5), 237-246.

Fox, H.B., McManus, M.A., Almeida, R.A., and Lesser, C. (1997). Medicaid managed care policies affecting children with disabilities: 1995 and 1996. *Health Care Financing Review, 18*(4), 23-36.

Frailberg, S. (1977). *Insights from the blind*. New York: Basic Books.

Frankenberger, W., and Fronzaglio, K. (1991). States' definitions and procedures for identifying children with mental retardation: Comparison over nine years. *Mental Retardation, 29*(6), 315-321.

Fujiura, G.T., and Yamaki, K. (2000). Trends in demography of childhood poverty and disability. *Exceptional Children, 66*(2), 187-199.

Galligan, B. (1990). Serving people who are dually diagnosed: A program evaluation. *Mental Retardation, 28*(6), 353-358.

Gallimore, R., and Goldenberg, C. (2001). Analyzing cultural models and settings to connect minority achievement and school improvement research. *Educational Psychologist, 36*(1), 45-56.

Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. New York, NY: Basic Books.

Garrett, B., and Glied, S. (2000). Does state AFDC generosity affect child SSI participation? *Journal of Policy Analysis and Management, 19*(2), 275-295.

Geisinger, K.F. (1992). *Psychological testing of Hispanics*. Washington, DC: American Psychological Association.

Geisinger, K.F. (1998). Psychometric issues in test interpretation. In J.H. Sandoval, C.L. Frisby, K.F. Geisinger, J.D. Scheuneman, and J.R. Grenier (eds.), *Test interpretation and diversity: Achieving equity in assessment* (pp. 17-30). Washington, DC: American Psychological Association.

Goddard, H.H. (1908). The Binet and Simon tests of intellectual capacity. *Training School, 5*, 3-9.

Goestason, R. (1985). Psychiatric illness among the mentally retarded: A Swedish population study. *Acta Psychiatrica Scandinavica, 71*(Suppl 318), 3-117.

Golden, C.J. (1990). *Clinical interpretation of objective psychological tests* (2nd ed.). Boston: Allyn and Bacon.

Gomez, R., and Hazeldine, P. (1996). Social information processing in mild mentally retarded children. *Research in Developmental Disabilities, 17*(3), 217-227.

Good, W.V., and Hoyt, C.S. (1989). Behavioral correlates of poor vision in children. *International Opthalmology Clinics, 29* (1), 57-60.

Gorsuch, R.L. (1983). *Factor analysis* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Goulden, K.J., Shinnar, S., Koller, H., Katz, M., and Richardson, S.A. (1991). Epilepsy in children with mental retardation: A cohort study. *Epilepsia, 32*(5), 690-697.

Grad, S. (2000). The role of disability benefits in the economic well-being of beneficiaries. Paper presented at the National Academy of Social Insurance Conference. L. Frieden, and J.L. Mashaw, (eds.), *Disability income policy: Opportunities and challenges in the next decade - A policy education seminar.* December 15, 2000.

Greenspan, S. (1999). A contextualist perspective on adaptive behavior. In R.L. Schalock, and D.L. Braddock (eds.), *Adaptive behavior and its measurements: Implications for the field of mental retardation* (pp. 61-77). Washington, DC: American Association on Mental Retardation.

Greenspan, S., and Driscoll, J. (1997). The role of intelligence in a broad model of personal competence. In D.P. Flanagan, J.L. Genshaft, et al. (eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 131-150). New York: Guilford Press.

Greenspan, S., and Granfield, J.M. (1992). Reconsidering the construct of mental retardation: Implications of a model of social competence. *American Journal on Mental Retardation, 96*(4), 442-453.

Greenspan, S., and Shoultz, B. (1981). Why mentally retarded adults lose their jobs:

Social competence as a factor in work adjustment. *Applied Research in Mental Retardation, 2*(1), 23-38.

Greenwald, D.F., Harder, D.W., Gift, T.E., Strauss, J.S., Ritzler, B., and Kokes, R. (1989). IQ as a prognostic indicator in adult psychiatric first-admissions. *Journal of Clinical Psychology, 45*(1), 37-50.

Gresham, F.M. (1991). Conceptualizing behavior disorders in terms of resistance to intervention. *School Psychology Review, 20*(1), 23-36.

Gresham, F.M., and Elliott, S.N. (1987). The relationship between adaptive behavior and social skills: Issues in definition and assessment. *Journal of Special Education, 21*(1), 167-181.

Gresham, F.M., and MacMillan, D.L. (1997). Social competence and affective characteristics of students with mild disabilities. *Review of Educational Research, 67*(4), 377-415.

Gresham, F.M., MacMillan, D.L., and Siperstein, G.N. (1995). Critical analysis of the 1992 AAMR definition: Implications for school psychology. *School Psychology Quarterly, 10*(1), 1-19.

Grigal, M., Test, D.W., Beattie, J., and Wood, W.M. (1997). An evaluation of transition components of individualized education programs. *Exceptional Children, 63*(3), 357-372.

Grossman, H.J. (1973). *Manual on terminology and classification in mental retardation*. Washington, DC: American Association on Mental Deficiency.

Grossman, H.J. (1977). *Manual on terminology and classification in mental retardation*. Washington, DC: American Association on Mental Deficiency.

Grossman, H.J. (1983). *Classification in mental retardation*. Washington DC: American Association on Mental Deficiency.

Guilford, J.P. (1967). *The nature of human intelligence*. New York: McGraw-Hill.

Guilford, J.P. (1974). Rotation problems in factor analysis. *Psychological Bulletin, 81*(8), 498-501.

Guilford, J.P., and Hoepfner, R. (1971). *The analysis of intelligence*. New York: McGraw-Hill.

Gumpel, T. (1994). Social competence and social skills training for persons with mental retardation: An expansion of a behavioral paradigm. *Education and Training in Mental Retardation, 29*(3), 194-201.

Gumpel, T., and Wilson, M.S. (1996). Application of a Rasch analysis to the examination of the perception of facial affect among persons with mental retardation. *Research in Developmental Disabilities, 17*(2), 161-171.

Halfon, N., and Newacheck, P.W. (1999). Prevalence and impact of parent-reported disabling mental health conditions among U.S. children. *Journal of the American Academy of Child and Adolescent Psychiatry, 38*(5), 600-609.

Hambleton, R.K., and Rodgers, J.H. (1995). *Item bias review*. Washington, DC: ERIC Clearinghouse on Assessment and Evaluation.

Haring, T.G. (1992). Review of the adaptive behavior: Street survival skills questionnaire. In J.C. Conoley, and J.J. Kramer (eds.), *The eleventh mental measurements yearbook, 11*. Lincoln, NE: Buros Institute of Mental Measurements.

Harrington, R.G. (1985). Review of the batelle developmental inventory. In J. Keyser, and R.C. Sweetland (eds.), *Test critiques* (Vol. 2, pp. 72-82). Austin, TX: Pro-Ed.

Harris, J.A. (1977). Emotion recognition in average and below average intelligence adults. Nashville, TN: George Peabody College.

Harrison, P.L. (1984). The application of the Vineland Adaptive Behavior Scales in educational settings. *Techniques, 1*(2), 101-112.

Harrison, P.L. (1987). Research with adaptive behavior scales. *Journal of Special Education, 21*(1), 37-68.

Harrison, P.L. (1998). Review of AAMR adaptive behavior school-residential and community - 2nd edition. In J.C. Conoley, and J.C. Impara (eds.), *The thirteenth mental measurements yearbook* (pp. 3-5). Lincoln, NE: Buros Institute of Mental Measurements.

Harrison, P.L., and Oakland, T. (2000a). *Adaptive behavior assessment system*. San Antonio, TX: Psychological Corporation.

Harrison, P.L., and Oakland, T. (2000b). *Adaptive behavior assessment system (manual)*. San Antonio, TX: Psychological Corporation.

Harrison, P.L., and Robinson, B. (1995). Best practices in the assessment of adaptive behavior. In A. Thomas, and J. Grimes (eds.), *Best practices in school psychology* (3rd ed., pp. 753-762). Washington, DC: The National Association of School Psychologists.

Hart, B. (2000). A natural history of early language experience. *Topics in Early Childhood Special Education, 20*(1), 28-32.

Hart, B., and Risley, T.R. (1992). American parenting of language-learning children: Persisting differences in family-child interactions observed in natural home environments. *Developmental Psychology, 28* (6), 1096-1105.

Hauser, W.A., and Hesdorffer, D.C. (1990). *Epilepsy: Frequency, causes, and consequences*. Landover, MD: Epilepsy Foundation of America.

Havercamp, S.M., and Reiss, S. (1997). The Reiss screen for maladaptive behavior: Confirmatory factor analysis. *Behaviour Research and Therapy, 35*(10), 967-971.

Haynes, S.N., Richard, D.C.S., and Kubany, E.S. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological Assessment, 7*(3), 238-247.

Heber, R. (1959). A manual on terminology and classification in mental retardation: A monograph supplement. *American Journal of Mental Deficiency, 64*(2), 1-111.

Heber, R. (1961). Modifications in the manual on terminology and classification in mental retardation. *American Journal of Mental Deficiency, 65*, 499-500.

Heflinger, C.A., Cook, V.J., and Thackrey, M. (1987). Identification of mental retardation by the system of multicultural pluralistic assessment: Nondiscriminatory or nonexistent? *Journal of School Psychology, 25*(2), 177-183.

Hemingson, C. (1998). The view from SSA's New Hampshire district office. In K. Rupp, and D.C. Stapleton (eds.), *Growth in disability benefits: Explanations and policy implications* (p. 307). Kalamazoo, MI: Upjohn Institute for Employment Research.

Herman, M.S., and Shantz, C.U. (1983). Social problem solving and mother-child interactions of educable mentally retarded children. *Journal of Applied Developmental Psychology, 4*(3), 217-226.

Herz, E.J., Chawla, A.J., and Gavin, N.I. (1998). Preventive services for children under Medicaid, 1989 and 1992. *Health Care Financing Review, 19*(4), 25-44.

Hickson, L., and Khemka, I. (1999). Decision making and mental retardation. In L.M. Glidden (vol. ed.), *International review of research in mental retardation* (Vol. 22, pp. 227-265). San Diego, CA: Academic Press.

Hickson, L., Golden, H., Khemka, I., Urv, T., and Yamusah, S. (1998). A closer look at interpersonal decision-making in adults with and without mental retardation. *American Journal on Mental Retardation, 103*(3), 209-224.

Hill, B.K. (1999) *Adaptive and maladaptive behavior scales.* [Online]. Available: http://www.isd.net/bhill/icap.htm [2002, March 22].

Hobson, R.P., Ouston, J., and Lee, A. (1989). Recognition of emotion by mentally retarded adolescents and young adults. *American Journal on Mental Retardation, 93*(4), 434-443.

Hodapp, R.M., and Dykens, E.M. (1996). Mental retardation. In E.J. Mash, and R.A. Barkley (eds.), *Child psychopathology* (pp. 362-389). New York: Guilford Press.

Hodapp, R.M., Dykens, E.M., and Masino, L.L. (1997). Families of children with Prader-Willi syndrome: Stress-support and relations to child characteristics. *Journal of Autism and Developmental Disorders, 27*(1), 11-24.

Hollon, S.D., and Flick, S.N. (1988). On the meaning and methods of clinical significance. *Behavioral Assessment, 10*(2), 197-206.

Holman, J.G., and Bruininks, R.H. (1985). Assessing and training adaptive behaviors. In K.L. Lakin, and R.H. Bruininks (eds.), *Strategies for achieving community integration of developmentally disabled citizens* (pp. 73-104). Baltimore, MD: Paul H. Brooks.

Horn, J.L. (1985). Remodeling old models of intelligence. In B.B. Wolman (ed.), *Handbook of intelligence: Theories, measurements, and applications* (pp. 267-300). New York: Wiley.

Horn, J.L. (1988). Thinking about human abilities. In J.R. Nesselroade, and R.B. Cattell (eds.), *Handbook of multvariate experimental psychology (2nd ed.) Perspectives on individual differences* (pp. 645-685). New York: Plenum Press.

Horn, J.L. (1998). A basis for research on age differences in cognitive capabilities. In J.J. McArdle, and R.W. Woodcock (eds), *Human cognitive abilities in theory and practice* (pp. 57-91). Mahwah, NJ: Lawrence Erlbaum Associates.

Horn, J.L., and Knapp, J.R. (1973). On the subjective character of the empirical base of Guilford's structure-of-intellect model. *Psychological Bulletin, 80*(1), 33-43.

Horn, J.L., and Knapp, J.R. (1974). Thirty wrongs do not make a right: Reply to Guilford. *Psychological Bulletin, 81*(8), 502-504.

Hoshino, Y., Kumashiro, H., Yashima, Y., Tachibana R., and Watanabe M. (1982). The epidemiological study of autism in Fukushima-ken. *Folia Psychiatrica Et Neurologica Japonica, 36*(2), 115-124.

Hoynes, H.W., and Moffitt, R. (1999). Tax rates and work incentives in the Social Security Disability Insurance program: Current law and alternative reforms. *National Tax Journal, 52*(4), 623-654.

Hundert, J., Morrison, L., Mahoney, W., Mundy, F., and Vernon, M.L. (1997). Parent and teacher assessments of the developmental status of children with severe, mild/

moderate, or no developmental disabilities. *Topics in Early Childhood Special Education, 17*(4), 419-434.

Hunter, J.E. (1986). Cognitive ability, cognitive aptitude, job knowledge, and job performance. *Journal of Vocational Behavior, 29*(3), 340-362.

Hunter, J.E., and Hunter, R.F. (1984). Validity and utility of alternative predictors of job performance. *Psychological Bulletin, 96*(1), 72-98.

Hunter, J.E., and Schmidt, F.L. (1990). *Methods of meta-analysis: Correcting error and bias in research findings*. Newbury Park, CA: Sage Publications.

Hutton, J.B., Dubes, R., and Muir, S. (1992). Assessment practices of school psychologists: Ten years later. *School Psychology Review, 21*(2), 271-284.

Iglehart, J.K. (1999). The American health care system - Medicaid. *New England Journal of Medicine, 340*(5), 403-408.

Illinois Planning Council on Developmental Disabilities. (1991). *Policy statements of the Illinois Planning Council on Developmental Disabilities*. Springfield, IL: Illinois Planning Council on Developmental Disabilities.

Individuals with Disabilities Education Act (IDEA) amendments of 1997. (1997). PL 105-17, Title 20 U.S.C. 1400 et seq.

Individuals with Disabilities Education Act (IDEA) of 1990. (1990). PL 101-476, Title 20 U.S.C. 1400 et seq.

Institute of Medicine. (1974). *Assessment of medical care in children* (Contrasts in health status No. 3). Washington, DC: National Academy Press.

Itard, J.M.G. (1932). *The wild boy of Aveyron*. New York, NY: The Century Co.

Jackson, H.J., King, N.J., and Heller, V.R. (1981). Social skills assessment and training for mentally retarded persons: A review of research. *Australian Journal of Developmental Disablilities, 7*(3), 113-123.

Jacobson, J.W. (1982). Problem behavior and psychiatric impairment within developmentally disabled population: I. Behavior frequency. *Applied Research in Mental Retardation, 3*(2), 121-139.

Jacobson, J.W. (1990). Assessing the prevalence of psychiatric disorders in a developmentally disabled population. In E. Dibble, and D.B. Gray (eds.), *Assessment of behavior problems in persons with mental retardation living in the community*. DHHS publication no. ADM 90-1642. Rockville, MD: National Institute of Mental Health.

Jacobson, J.W. (1992). The new definition of mental retardation: Reclassification or misclassification? Paper presented at the Gatlinberg Conference on Theory and Research in Mental Retardation. April, 1992.

Jacobson, J.W. (1997). Environmental postmodernism and rehabilitation of the borderline of mental retardation. Paper presented at the Annual Convention of the American Psychological Association. August, 1997.

Jacobson, J.W. (1999). Dual diagnosis services: History, progress and perspectives. In N. Bouras (ed.), *Psychiatric and behavioural disorders in developmental disabilities and mental retardation* (pp. 329-357). New York, NY: Cambridge University Press.

Jacobson, J.W., and Janicki, M.P. (1983). Observed prevalence of multiple developmental disabilities. *Mental Retardation, 21*(3), 87-94.

Jacobson, J.W., and Mulick, J.A. (1996). *Manual of diagnosis and professional practice in mental retardation*. Washington, DC: American Psychological Association.

Janicki, M.P., and Jacobson, J.W. (1982). The character of developmental disabilities in New York state: Preliminary observations. *International Journal of Rehabilitation Research, 5*(2), 191-202.

Jencks, C. (1998). Racial bias in testing. In C. Jencks, and M. Phillips (eds.), *The black-white test score gap* (pp. 55-85). Washington, DC: Brookings Institution.

Jenkinson, J., and Nelms, R. (1994). Patterns of decision-making behaviour by people with intellectual disability: An exploratory study. *Australia and New Zealand Journal of Developmental Disabilities, 19*(2), 99-109.

Jensen, A.R. (1980). *Bias in mental testing*. New York: Free Press.

Jensen, A.R. (1981). *Straight talk about mental tests*. New York: Free Press.

Jensen, A.R. (1984). Test validity: g versus the specificity doctrine. *Journal of Social and Biological Structures, 7*(2), 93-118.

Jensen, A.R. (1998). *The g Factor: The science of mental ability*. Westport, CT: Praeger.

Joint Committee on Standards for Educational and Psychological Testing of the American Educational Research Association, the American Psychological Association, and the National Council on Measurement in Education. (1985). *Standards for educational and psychological testing*. Washington, DC: American Psychological Association.

Joint Committee on Standards for Educational and Psychological Testing of the American Educational Research Association, the American Psychological Association, and the National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

Kamphaus, R.W. (1987a). Conceptual and psychometric issues in the assessment of adaptive behavior. *Journal of Special Education, 21*(1), 27-35.

Kamphaus, R.W. (1987b). Critiques of school psychological materials. *Journal of School Psychology, 25*, 97-100.

Kamphaus, R.W., Petoskey, M.D., and Rowe, E.W. (2000). Current trends in psychological testing of children. *Professional Psychology: Research and Practice, 31*(2), 155-164.

Kapell, D., Nightingale, B., Rodriguez, A., Lee, J.H., Zigman, W.B., and Schupf, N. (1998). Prevalence of chronic medical conditions in adults with mental retardation: Comparison with the general population. *Mental Retardation, 36*(4), 269-279.

Karno, M., Hough, R.L., Burnam, M.A., Escobar, J.I., et al. (1987). Lifetime prevalence of specific psychiatric disorders among Mexican Americans and non-Hispanic whites in Los Angeles. *Archives of General Psychiatry, 44*(8), 695-701.

Kaufman, A.S. (1975). Factor analysis of the WISC-R at eleven age levels between 6 1/2 and 16 1/2 years. *Journal of Consulting and Clinical Psychology, 43*(2), 135-147.

Kaufman, A.S. (1979). *Intelligent testing with the WISC-R*. New York: Wiley.

Kaufman, A.S. (1994). *Intelligent testing with the WISC-III*. New York: Wiley.

Kaufman, A.S., and Kaufman, N.L. (1983). *Kaufman assessment battery for children*. Circle Pines, MN: American Guidance Service.

Kaufman, A.S., and Kaufman, N.L. (1990). *Kaufman brief intelligence test*. Circle Pines, MN: American Guidance Service.

Kazimour, K.K., and Reschly, D.J. (1981). Investigation of norms and concurrent validity for the Adaptive Behavior Inventory for Children. *American Journal of Mental Deficiency, 85*(5), 512-520.

Keltner, B.R. (1992). Caregiving by mothers with mental retardation. *Family and Community Health, 15*(2), 10-18.

Kerlin, I. (1887). Moral imbecility. In *Proceedings of the Association of Medical Officers of American Institutions for Idiotic and Feebleminded Persons* (pp. 32-37).

Kiely, M. (1987). The prevalence of mental retardation. *Epidemiologic Reviews, 9*, 194-218.

Koegel, P., and Edgerton, R.B. (1984). Black "six hour retarded children" as young adults. In R.B. Edgerton (ed.), *Lives in process: Mildly retarded adults in a large city* (pp. 145-171). Washington, DC: American Association on Mental Deficiency.

Kohlberg, L. (1969). *Stages in the development of moral thought and action*. New York: Rinehart and Winston.

Koller, H., Richardson, S.A., Katz, M., and McLaren, J. (1982). Behavior disturbance in childhood and the early adult years in populations who were and were not mentally retarded. *Journal of Preventive Psychiatry, 1*, 453-468.

Konigsmark, B.W., and Gorlin, R.J. (1976). *Genetic and metabolic deafness*. Philadelphia, PA: Saunders.

Krueger, A.B. (1997). How will labor market trends affect people with disabilities? In V.P. Reno, J.L. Mashaw, and W.D. Gradison (eds.), *Disability: Challenges for social insurance, health care financing, and labor market policy*. Washington, DC: National Academy of Social Insurance.

Kubik, J. (2000). Fiscal federalism and welfare policy: The role of states in the growth of child SSI. Syracuse, NY: Syracuse University, unpublished manuscript.

Kuhlman, F. (1920). The results of mental re-examinations of the feeble-minded. *Journal of Psycho-Asthenics, 25*(147-160).

Lambert, N.M., Nihira, K., and Leland, H. (1993a). *AAMR Adaptive Behavior Scale (ABS-S:2): School* (2nd ed.). Austin, TX: Pro-Ed.

Lambert, N.M., Nihira, K., and Leland, H. (1993b). *AAMR Adaptive Behavior Scales-School (ABS-S:2) Examiner's Manual*. Austin, TX: Pro-Ed.

Larson, S.A., Lakin, K.C., Anderson, L., Kwak, N., and Lee, J.H. (2000). *Prevalence of mental retardation or developmental disabilities: Analysis of the 1994/1995 NHIS-D. MR/DD Data Brief 1(2)*. Minneapolis, MN: Research and Training Center.

Larson, S.A., Lakin, K.C., Anderson, L., Kwak, N., Lee, J.H., and Anderson, D. (2001). Prevalence of mental retardation and developmental disabilities: Estimates from the 1994/1995 National Health Interview Survey Disability Supplements. *American Journal on Mental Retardation, 106*(3), 231-252.

Leffert, J.S., and Siperstein, G.N. (1996). Assessment of social-cognitive processes in children with mental retardation. *American Journal on Mental Retardation, 100*(5), 441-455.

Leffert, J.S., and Siperstein, G.N. (in press). Social cognition: The key to understanding adaptive behavior in individuals with mild mental retardation. In L.M. Glidden (ed.), *International review of research in mental retardation*. San Diego, CA: Academic Press.

Leffert, J.S., Siperstein, G.N., and Millikan, E. (2000). Understanding social adaptation in children with mental retardation: A social-cognitive perspective. *Exceptional Children, 66*(4), 530-545.

Leiter, R.G. (1948). *Leiter International Performance Scale*. Chicago, IL: Stoelting.

Leland, H., Shellhaas, M., Nihira, K., and Foster, R. (1967). Adaptive behavior: A new dimension in the classification of the mentally retarded. *Mental Retardation Abstracts, 4*(3), 359-387.

Linaker, O. (1991). DSM-III diagnoses compared with factor structure of the Psychopathology Instrument for Mentally Retarded Adults (PIMRA), in an institutionalized, mostly severely retarded population. *Research in Developmental Disabilities, 12*(2), 143-153.

Lindsay, P.H., Shapiro, A., Musselman, C.R., and Wilson, A. (1988). Predicting language development in deaf children using subscales of the Leiter international performance scale. *Canadian Journal of Psychology, 42*(2), 144-162.

Linkenhoker, D.D., and McCarron, L.T. (1983). *Street Survival Skills Questionnaire*. Dallas, TX: Common Market Press.

Loeb, P.A. (1996). *Independent living scales (ILS)*. Sidcup, Kent, UK: Psychological Corporation.

Loevinger, J. (1957). Objective tests as instruments of psychological theory. *Psychological Reports, 3*, 635-694.

Lotter, V. (1966). Epidemiology of autistic conditions in young children: I. Prevalence. *Social Psychiatry, 1*, 124-137.

Lovaas, O.I. (1987). Behavioral treatment and normal educational and intellectual functioning in young autistic children. *Journal of Consulting and Clinical Psychology, 55*(1), 3-9.

Lubin, R.A., Jacobson, J.W., and Kiely, M. (1982). Projected impact of the functional definition of developmental disabilities: The categorically disabled population and service eligibility. *American Journal of Mental Deficiency, 87*(1), 73-79.

Luria, A.R. (1966a). *Higher cortical functions in man.* New York: Basic Books.

Luria, A.R. (1966b). *Human brain and psychological processes.* New York: Harper and Row.

Luster, T., and Dubow, E. (1992). Home environment and maternal intelligence as predictors of verbal intelligence: A comparison of preschool and school-age children. *Merrill-Palmer Quarterly, 38*(2), 151-175.

Lyon, G.R. (1996). Learning disabilities. *The Future of Children, 6*(1), 54-76.

MacMillan, D.L., and Reschly, D.J. (1996). Issues of definition and classification. In W.E.J. MacLean (ed.), *Handbook of mental deficiency: Psychological theory and research* (3rd ed., pp. 47-74). Hillsdale, NJ: Lawrence Erlbaum Associates.

MacMillan, D.L., Gresham, F.M., and Bocian, K.M. (1998). Curing mental retardation and causing learning disabilities: Consequences of using various WISC-III IQs to estimate aptitude of Hispanic students. *Journal of Psychoeducational Assessment, 16*(1), 36-54.

MacMillan, D.L., Gresham, F.M., and Siperstein, G.N. (1993). Conceptual and psychometric concerns about the 1992 AAMR definition of mental retardation. *American Journal on Mental Retardation, 98*(3), 325-335.

MacMillan, D.L., Gresham, F.M., and Siperstein, G.N. (1995). Heightened concerns over the 1992 AAMR definition: Advocacy versus precision. *American Journal on Mental Retardation, 100*(1), 87-95.

MacMillan, D.L., Gresham, F.M., Siperstein, G.N., and Bocian, K.M. (1996). The labyrinth of I.D.E.A.: School decisions on referred students with subaverage general intelligence. *American Journal on Mental Retardation, 101*(2), 161-174.

MacMillan, D.L., Siperstein, G.N., and Gresham, F.M. (1996). A challenge to the viability of mild mental retardation as a diagnostic category. *Exceptional Children, 62*(4), 356-371.

MacMillan, D.L., Siperstein, G.N., and Leffert, J.S. (in press). Children with mild mental retardation: A challenge for classification practices. In H. Switaky, and S. Greenspan (eds.), *What is mental retardation? (Ideas for the new century)*. Washington, DC: American Association on Mental Retardation.

Maheady, L., Maitland, G., and Sainato, D.M. (1984). The interpretation of social interactions by mildly handicapped and nondisabled children. *The Journal of Special Education, 18*(2), 151-159.

Mank, D.M. (2001). Employability of individuals with mental retardation. Unpublished paper written for the National Research Council, the Committee on Disability Determination for Mental Retardation. Department of Curriculum and Instruction, Indiana University.

Marshak, L.E., Seligman, M., and Prezant, F. (1999). *Disability and the family life cycle: Recognizing and treating developmental challenges*. New York: Basic Books.

Martel, J., McKelvie, S.J., and Standing, L. (1987). Validity of an intuitive personality scale: Personal responsibility as a predictor of academic achievement. *Educational and Psychological Measurement, 47*(4), 1153-1163.

Mashaw, J.L., and Reno, V.P. (1996). *Balancing security and opportunity: The challenge of disability income policy*. Washington, DC: National Academy of Social Insurance.

Matazow, G.S., Kamphaus, R.W., Stanton, H.C., and Reynolds, C.R. (1991). Reliability of the Kaufman Assessment Battery for Children for black and white students. *Journal of School Psychology, 29*(1), 37-41.

Matilainen, R., Airaksinen, E., Mononen, T., Launiala, K., and Kaariainen, R. (1995). A population-based study on the causes of mild and severe mental retardation. *Acta Pediatrica, 84*(3), 261-266.

Matson, J.L., and Bamburg, J.W. (1998). Reliability of the assessment of dual diagnosis (ADD). *Research in Developmental Disabilities, 19*(1), 89-95.

Matson, J.L., Barrett, R.P., and Helsel, W.J. (1988). Depression in mentally retarded children. *Research in Developmental Disabilities, 9*(1), 39-46.

Matson, J.L., Helsel, W.J., Bellack, A.S., and Senatore, V. (1983). Development of a rating scale to assess social skill deficits in mentally retarded adults. *Applied Research in Mental Retardation, 4*, 399-407.

Matthias, J.L., and Nettelbeck, T. (1992). Validity of Greenspan's models of adaptive and social intelligence. *Research in Developmental Disabilities, 13*(2), 113-129.

McAdoo, H.P. (1995). Stress levels, family help patterns, and religiosity in middle- and working-class African American single mothers. *Journal of Black Psychology, 21*(4), 424-449.

McArdle, J.J., and Woodcock, R.W. (1998). *Human cognitive abilities in theory and practice*. Mahwah, NJ: Erlbaum.

McArthur, G.M., Hogben, J.H., Edwards, V.T., Heath, S.M., and Mengler, E.D. (2000). On the "specifics" of specific reading disability and specific language impairment. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 41*(7), 869-874.

McCarney, S.B. (1983). *Adaptive behavior evaluation scale*. Columbia, MO: Hawthorne Educational Services.

McCarney, S.B. (1988). *Parent rating of student behavior*. Columbia, MO: Hawthorne Educational Services.

McCarthy, D. (1972). *McCarthy scales of children's abilities*. Cleveland, OH: Psychological Corporation.

McClaren, J., and Bryson, S.E. (1987). Review of recent epidemiological studies in mental retardation. *American Journal on Mental Retardation, 92*, 243-254.

McCullough, P.G., and Rutenberg, S.K. (1988). Launching children and moving on. In B. Carter, and M. McGoldrick (eds.), *The changing family life cycle: A framework for family therapy* (2nd ed., pp. 285-309). New York: Gardner Press.

McFadden, T.U. (1996). Creating language impairments in typically achieving children: The pitfalls of "normal" normative sampling. *Language, Speech, and Hearing Services in Schools, 27*(1), 3-9.

McFall, R.M. (1982). A review and reformulation of the concept of social skills. *Behavioral Assessment, 4*(1), 1-33.

McGrew, K.S., and Bruininks, R.H. (1989). The factor structure of adaptive behavior. *School Psychology Review, 18*(1), 64-81.

McGrew, K.S., and Bruininks, R.H. (1990). Defining adaptive and maladaptive behavior within a model of personal competence. *School Psychology Review, 19*(1), 53-73.

McIntyre, L.L., Blacher, J., and Baker, B.L. (2002). Behaviour problems/mental health problems in young adults with intellectual disability: The impact on families. *Journal of Intellectual Disability Research 46*(3), 239-249.

McNamara, J.R., Hollmann, C., and Riegel, T. (1994). A preliminary study of the usefulness of the Behavior Assessment Systems for Children in the evaluation of mental health needs in a Head Start population. *Psychological Reports, 75*(3, Pt 1), 1195-1201.

Menolascino, F.J., and Fleisher, M.H. (1991). Developmental concepts in mental retardation and mental illness. *Comprehensive Mental Health Care, 1*(1), 45-56.

Mercer, J., and Lewis, J. (1978). *Adaptive behavior inventory for children*. San Antonio, TX: Psychological Corporation.

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika, 58*(4), 525-543.

Messick, S. (1989). Meaning and values in test validation: The science and ethics of assessment. *Educational Researcher, 18*(2), 5-11.

Messick, S. (1993). Validity. In R.L. Linn (ed.), *Educational measurement* (3rd ed., pp. 13-104). Phoenix, AZ: Oryx Press.

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist, 50*(9), 741-749.

Meyer, L.H., Cole, D.A., McQuarter, R., and Reichle, J. (1990). Validation of the Assessment of Social Competence (ASC) for children and young adults with developmental disabilities. *Journal of the Association for Persons with Severe Handicaps, 15*(2), 57-68.

Meyers, C.E., Nihira, K., and Zetlin, A. (1979). The measurement of adaptive behavior. In N.R. Ellis (ed.), *Handbook of mental deficiency: Psychological theory and research*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Miller, L.T., and Vernon, P.A. (1992). The general factor in short-term memory, intelligence, and reaction time. *Intelligence, 16*(1), 5-29.

Millham, J., Chilcutt, J., and Atkinson, B.L. (1978). Comparability of naturalistic and controlled observation assessment of adaptive behavior. *American Journal of Mental Deficiency,* (83), 52-59.

Miner, C.A., and Bates, P.E. (1997). The effect of person centered planning activities on the IEP/transition planning process. *Education and Training in Mental Retardation and Developmental Disabilities, 32*(2), 105-112.

Mitchell, L., and Lambourne, R.D. (1979). An association between high intellectual ability and an imaginative and analytic approach to the discussion of open questions. *British Journal of Educational Psychology, 49*(1), 60-72.

Monti, P.M. (1983). The social skills intake interview: Reliability and convergent validity assessment. *Journal of Behavior Therapy and Experimental Psychiatry, 14*, 305-310.

Mor, V. (1998). A modern lexicon of disability. In S. Allen, and V. Mor (eds.), *Living in the community with disability: Service needs, use, and systems* (pp. 353-371). New York: Springer.

Moran, M.R. (2001). Review of the adaptive behavior evaluation scale. In J.C. Conoley, and J.J. Kramer (eds.), *The fourteenth mental measurements yearbook, 14*. Lincoln, NE: Buros Institute of Mental Measurements.

Morton, K.L., and Green, V. (1991). Comprehension of terminology related to treatment and patients' rights by inpatient children and adolescents. *Journal of Clinical Child Psychology, 20* (4), 392-399.

Moss, S., Emerson, E., Kiernan, C., Turner, S., Hatton, C., and Alborz, A. (2000). Psychiatric symptoms in adults with learning disability and challenging behaviour. *British Journal of Psychiatry, 177*, 452-456.

Moss, S., Prosser, H., and Goldberg, D. (1996). Validity of the schizophrenia diagnosis of the Psychiatric Assessment Schedule for Adults with Developmental Disability (PAS—ADD). *British Journal of Psychiatry, 168*(3), 359-367.

Moss, S., Prosser, H., Costello, H., Simpson, N., Patel, P., Rowe, S., Turner, S., and Hatton, C. (1998). Reliability and validity of the PAS-ADD checklist for detecting psychiatric disorders in adults with intellectual disability. *Journal of Intellectual Disability Research, 42*(2), 173-183.

Naglieri, J.A., and Das, J.P. (1997). *Cognitive assessment system*. Itasca, IL: Riverside Publishing.

National Commission on Childhood Disability. (1995). *Supplemental security income for children with disabilities: Report to Congress*. Washington, DC: National Commission on Childhood Disability.

National Council on Disability. (1996). *Improving the implementation of the Individuals* *with Disabilities Education Act: Making schools work for all of America's children, Supplement.* Washington, DC: National Council on Disability.

National Research Council. (2001). *Educating children with autism*. Washington, DC: National Academy Press.

Neisser, U. (1998). *The rising curve: Long-term gains in IQ and related measures*. Washington, DC: American Psychological Association.

Neisser, U., Boodoo, G., Bouchard, T.J., Boykin, A.W., Brody, N., Ceci, S.J., Halpern, D.F., Loehlin, J.C., Perloff, R., Sternberg, R.J., and Urbina, S. (1996). Intelligence: Knowns and unknowns. *American Psychologist, 51*(2), 77-101.

Newborg, J., Stock, J.R., Wnek, L., Guidubaldi, J., and Svinicki, J.G. (1984). *Battelle Developmental Inventory*. Allen, TX: DLM Teaching Resources.

Nezu, C.M., Nezu, A.M., and Gill-Weiss, M.J. (1992). *Psychopathology in persons with mental retardation: Clinical guidelines for assessment and treatment*. Champaign, IL: Research Press.

Nihira, K., Leland, H., and Lambert, N.M. (1993). *AAMR Adaptive Behavior Scales-Residential and Community (ABS-RC:2)*. Austin, TX: Pro-Ed.

Northern, J.L., and Downs, M.P. (1991). *Hearing in children* (4th ed.). Baltimore, MD: Williams and Wilkins.

Nunnally, J.C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.

Nunnally, J.C., and Bernstein, I.H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.

Oakland, T., and Parmelee, R. (1985). Mental measurement of minority-group children. In B.B. Wolman (ed.), *Handbook of intelligence: Theories, measurements, and applications* (pp. 699-736). New York: Wiley.

Ochoa, S.H., Powell, M.P., and Robles-Pina, R. (1996). School psychologists' assessment practices with bilingual and limited-English-proficient students. *Journal of Psychoeducational Assessment, 14*(3), 250-275.

Oehler-Stinnett, J. (1989). Review of the battelle developmental inventory. In J.C. Conoley, and J.J. Kramer (eds.), *The tenth mental measurements yearbook* (pp. 10-75). Lincoln, NE: Buros Institute of Mental Measurements.

Opitz, J.M. (1996). Historiography of the causal analysis of mental retardation. Speech to the Annual Gatlinburg Conference on Research and Theory in Mental Retardation. March, 1996.

Otis, A.S., and Lennon, R.T. (1979). *Otis-Lennon School Ability Test*. New York: Psychological Corporation.

O'Toole, B.I., and Stankov, L. (1992). Ultimate validity of psychological tests. *Personality and Individual Differences, 13*(6), 699-716.

Paal, N., Skinner, S., and Reddig, C. (1988). The relationship of non-verbal intelligence measures to academic achievement among deaf adolescents. *Journal of Rehabilitation of the Deaf, 21*(3), 8-11.

Painter, M.J., and Gaus, L.M. (1991). Neonatal seizures: Diagnosis and treatment. *Journal of Child Neurology, 6*(2), 101-108.

Palmer, F.B., and Hoon, A.H. (1995). Cerebral palsy. In S. Parker, and B. Zukerman (eds.), *Behavioral and developmental pediatrics: A handbook for primary care* (pp. 88-94). Boston: Little, Brown.

Paniagua, F.A. (1998). *Assessing and treating culturally diverse clients: A practical guide* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Parish, S.L., and Braddock, D.L. (2001). The context of Social Security Administration's disability determination process for mental retardation. Unpublished paper written for the Committee on Disability Determination for Mental Retardation. Department of Disability and Human Development, University of Illinois at Chicago.

Patrick, J.L., and Reschly, D.J. (1982). Relationship of state educational criteria and demographic variables to school-system prevalence of mental retardation. *American Journal of Mental Deficiency, 86*(4), 351-360.

Pearson, D.A., and Lachar, D. (1994). Using behavioral questionnaires to identify adaptive deficits in elementary school children. *Journal of School Psychology, 32*(1), 33-52.

Petersen, C.D., and Bascetta, C.A. (1999). Rethinking federal disability policy: Return to work strategies from the U.S. private sector. In S.S. Nagel (ed.), *The substance of public policy* (pp. 265-274). Commack, NY: Nova Science.

Pfeiffer, S.I., and Baker, B.L. (1994). Residential treatment for children with dual diagnoses of mental retardation and mental disorder. In J. Blacher (ed.), *When there's no place like home: Options for children living apart from their natural families*. Baltimore, MD: Paul H. Brookes.

Philips, I., and Williams, N. (1975). Psychopathology and mental retardation: A study of 100 mentally retarded children: I. Psychopathology. *American Journal of Psychiatry, 132*(12), 1265-1271.

Poon, L.W., Messner, S., Martin, P., Noble, C.A., Clayton, G.M., and Johnson, M.A. (1992). The influences of cognitive resources on adaptation and old age. *International Journal of Aging and Human Development, 34*(1), 31-46.

Porteus, S. (1921). A social rating scale for defectives. *Journal of Psycho-Asthenics, 26*, 117-126.

Post, M.W.M., de Witte, L.P., and Schrijvers, A.J.P. (1999). Quality of life and the ICIDH: Towards an integrated conceptual model for rehabilitation outcomes research. *Clinical Rehabilitation, 13*(1), 5-15.

Poteat, G.M., Wuensch, K.L., and Gregg, N.B. (1988). An investigation of differential prediction with the WISC-R. *Journal of School Psychology, 26*(1), 59-68.

President's Committee on Employment of People with Disabilities. (1994). *Operation people first: Toward a national disability policy*. Washington, DC: President's Committee on Employment of People with Disabilities.

President's Committee on Mental Retardation. (1970). *The six hour retarded child*. Washington, DC: United States Government Printing Office.

Pring, L. (1984). A comparison of the word recognition processes of blind and sighted children. *Child Development, 55*(5), 1865-1877.

Prosser, H., Moss, S., Costello, H., Simpson, N., Patel, P., and Rowe, S. (1998). Reliability and validity of the Mini PAS-ADD for assessing psychiatric disorders in adults with intellectual disability. *Journal of Intellectual Disability Research, 42*(4), 264-272.

Prout, H.T. (1993). Assessing psychopathology in persons with mental retardation: A review of the Reiss Scales. *Journal of School Psychology, 31*(4), 535-540.

Ramey, C.T., Mulvihill, B.A., and Ramey, S.L. (1996). Prevention: Social and educational factors and early intervention. In J.W. Jacobson, and J.A. Mulick (eds.), *Manual of diagnosis and professional practice in mental retardation* (pp. 215-227). Washington, DC: American Psychological Association.

Ramey, S., and Ramey, C.T. (1999). Early experience and early intervention for children "at risk" for developmental delay and mental retardation. *Mental Retardation and Development Disabilities Research Reviews, 5*(1), 1-10.

Raven, J.C., Court, J.H., and Raven, J.E. (1986). Section 1: General overview. In *Manual for Raven's Progressive Matrices and Vocabulary Scales*. London: H. K. Lewis.

Reiss, S. (1990). Prevalence of dual diagnosis in community-based day programs in the Chicago metropolitan area. *American Journal on Mental Retardation, 94*(6), 578-585.

Reiss, S. (1994). Issues in defining mental retardation. *American Journal on Mental Retardation, 99*(1), 1-7.

Reiss, S., and Szyszko, J. (1983). Diagnostic overshadowing and professional experience with mentally retarded persons. *American Journal of Mental Deficiency, 87*(4), 396-402.

Reiss, S., and Valenti-Hein, D. (1994). Development of a psychopathology rating scale for children with mental retardation. *Journal of Consulting and Clinical Psychology, 62*(1), 28-33.

Reiss, S., Levitan, G.W., and Szyszko, J. (1982). Emotional disturbance and mental retardation: Diagnostic overshadowing. *American Journal of Mental Deficiency, 86*(6), 567-574.

Reschly, D.J. (1981a). Evaluation of the effects of SOMPA measures on classification of students as mildly mentally retarded. *American Journal of Mental Deficiency, 86*(1), 16-20.

Reschly, D.J. (1981b). Psychological testing in educational classification and placement. *American Psychologist, 36*(10), 1094-1102.

Reschly, D.J. (1990). Best practices in adaptive behavior. In A. Thomas, and J. Grimes (eds.), *Best practices in school psychology* (Vol. 29-42). Washington, DC: National Association of School Psychology.

Reschly, D.J. (1992). Special education decision making and functional behavioral assessment. In W. Stainback, and S. Stainback (eds.), *Controversial issues confronting special education: Divergent perspectives* (pp. 127-138). Boston: Allyn and Bacon.

Reschly, D.J., and Ward, S.M. (1991). Use of adaptive measures and overrepresentation of black students in programs for students with mild mental retardation. *American Journal on Mental Retardation, 96*(3), 257-268.

Reynolds, C.R., and Kaiser, S.M. (1990). Test bias in psychological assessment. In T.B. Gutkin, and C.R. Reynolds (eds.), *The handbook of school psychology* (2nd ed., pp. 487-525). New York: Wiley.

Roberts, C., McCoy, M., Reidy, D., and Crucitti, F. (1993). A comparison of methods of assessing adaptive behaviour in pre-school children with developmental disabilities. *Australia and New Zealand Journal of Developmental Disabilities, 18*(4), 261-272.

Roberts, J., and Baird, J.T. (1972). *Behavior patterns in children in school.* (Report No. HSM 72-1042). Washington, DC: United States Government Printing Office.

Rogers, S.J. (1998). Neurospsychology of autism in young children and its implications for intervention. *Mental Retardation and Development Disabilities Research Reviews, 4*(2), 104-112.

Rogers-Dulan, J., and Blacher, J. (1995). African American families, religion, and disability: A conceptual framework. *Mental Retardation, 33*(4), 226-238.

Rutter, M., Simonoff, E., and Plomin, R. (1996). Genetic influences on mild mental retardation: Concepts, findings, and research implications. *Journal of Biosocial Science, 28*(4), 509-526.

Rutter, M., Tizard, J., Yule, W., Graham, P., and Whitmore, K. (1976). Research report: Isle of Wight studies, 1964-1974. *Psychological Medicine, 6*(2), 313-332.

Salvia, J., and Ysseldyke, J.E. (2001). *Assessment*. Boston: Houghton Mifflin.

Salzberg, C.L., Lignugaris-Kraft, B., and McCuller, G.L. (1988). Reasons for job loss: A review of employment termination studies of mentally retarded workers. *Research in Developmental Disabilities, 9*(2), 153-170.

Salzberg, C.L., Likins, M., McConaughy, E.K., and Lignugaris-Kraft, B. (1986). Social competence and employment of retarded persons. In N. Ellis, and N.W. Bray (eds.), *International review of research in mental retardation* (Vol. 14, pp. 225-257). New York: Academic Press.

Sameroff, A.J. (1986). Environmental context of child development. *The Journal of Pediatrics, 109*(1), 192-200.

Sattler, J.M. (1988). *Assessment of children* (3rd ed.). San Diego, CA: Author.

Sattler, J.M. (1990). *Assessment of children* (revised reprint, 3rd ed.). San Diego, CA: Author.

Scanlon, W.J. (2000). *Children with disabilities: Medicaid can offer important benefits and services.* (Report No. GAO/T-HEHS-00-152). Washington, DC: United States General Accounting Office.

Scheerenberger, R.C. (1983). *A history of mental retardation*. Baltimore, MD: Paul H. Brookes.

Scheuneman, J.D., and Oakland, T. (1998). High-stakes testing in education. In J.H. Sandoval, C.L. Frisby, K.F. Geisinger, J.D. Scheuneman, and J.R. Grenier (eds.), *Test interpretation and diversity: Achieving equity in assessment* (pp. 77-103). Washington, DC: American Psychological Association.

Schmidt, F.L., and Hunter, J.E. (1977). Development of a general solution to the problem of validity generalization. *Journal of Applied Psychology, 62*(5), 529-540.

Schmidt, F.L., and Hunter, J.E. (1981). Employment testing: Old theories and new research findings. *American Psychologist, 36*(10), 1128-1137.

Seguin, E. (1856). Origin of the treatment and training of idiots. *American Journal of Education, 2*, 145-152.

Seguin, E. (1866). *Idiocy and its treatment by the physiological method*. New York: William Wood.

Shealy, R., and Stout, W.F. (1993). A model-based standardization approach that separates true bias/DIF from group ability differences and detects test bias/DTF as well as item bias/DIF. *Psychometrika, 58*(2), 159-194.

Shepard, L., Camilli, G., and Averill, M. (1981). Comparison of procedures for detecting test-item bias with both internal and external ability criteria. *Journal of Educational Statistics, 6*(4), 317-375.

Sherrod, L.R., Haggerty, R.J., and Featherman, D.L. (1993). Introduction: Late adolescence and the transition to adulthood. *Journal of Research on Adolescence, 3*(3), 217-226.

Shoham-Vardi, I., Davidson, P.W., Cain, N.N., Sloane-Reeves, J.E., Giesow, V.E., Quijano, L.E., and Houser, K.D. (1996). Factors predicting re-referral following crisis intervention for community-based persons with developmental disabilities and behavioral and psychiatric disorders. *American Journal on Mental Retardation, 101*(2), 109-117.

Siebert, L.A. (1962). Otis IQ scores of delinquents. *Journal of Clinical Psychology, 18*(4), 517.

Siperstein, G.N., and Leffert, J.S. (1997). Comparison of socially accepted and rejected children with mental retardation. *American Journal on Mental Retardation, 101*(4), 339-351.

Sireci, S.G., and Geisinger, K.F. (1998). Equity issues in employment testing. In J.H. Sandoval, C.L. Frisby, K.F. Geisinger, J.D. Scheuneman, and J.R. Grenier (eds.), *Test interpretation and diversity: Achieving equity in assessment* (pp. 105-140). Washington, DC: American Psychological Association.

Sitlington, P.L. (1996). Transition to living: The neglected component of transition programming for individuals with learning disabilities. *Journal of Learning Disabilities, 29*(1), 31-39.

Smith, C.R. (1989). Review of the adaptive behavior inventory. In J.C. Conoley, and J.J. Kramer (eds.), *The tenth mental measurements yearbook, 10*. Lincoln, NE: Buros Institute of Mental Measurements.

Smith, D.C. (1986). Interpersonal problem-solving skills of retarded and nonretarded children. *Applied Research in Mental Retardation, 7*(4), 431-442.

Smith, I.L., and Greenberg, S. (1979). Hierarchical assessment of social competence. *American Journal of Mental Deficiency, 83*, 551-555.

Smith, T., Groen, A.D., and Wynn, J.W. (2000). Randomized trial of intensive early intervention for children with pervasive developmental disorder. *American Journal on Mental Retardation, 105*(4), 269-285.

Social Security Administration. (1999). *Red book on work incentives: A summary guide to social security and supplemental security income work incentives for people with disabilities*. Baltimore, MD: Department of Health and Human Services.

Social Security Administration. (2000). 20 CFR: Parts 404 and 416 [Regulations 4 and 16] RIN 0960-AF40. Supplemental Security Income: Determining disability for a child under age 18. *Federal Register, 65 (No. 176)*. 54747-54790.

Social Security Administration. (2001a). *Annual statistical supplement*. Washington, DC: U.S. Government Printing Office.

Social Security Administration. (2001b). *Children Receiving SSI, December 2000*. Washington, DC: U.S. Government Printing Office.

Social Security Administration. (2002). *Disability evaluation under Social Security*. Washington, DC: U.S. Government Printing Office.

Sparrow, S.S., and Cicchetti, D.V. (1985). Diagnostic uses of the Vineland Adaptive Behavior Scales. *Journal of Pediatric Psychology, 10*(2), 215-225.

Sparrow, S.S., Balla, D.A., and Cicchetti, D.V. (1984a). *Vineland adaptive behavior scales*. Circle Pines, MN: American Guidance Service.

Sparrow, S.S., Balla, D.A., and Cicchetti, D.V. (1984b). *Vineland adaptive behavior scales (Interview ed.)*. Circle Pines, MN: American Guidance Service.

Sparrow, S.S., Balla, D.A., and Cicchetti, D.V. (1984c). *Vineland Adaptive Behavior Scales: Interview edition, expanded form manual*. Circle Pines, MN: American Guidance Service.

Spearman, C.E. (1904a). 'General intelligence,' objectively determined and measured. *American Journal of Psychology, 15*(2), 201-293.

Spearman, C.E. (1904b). The proof and measurement of association between two things. *American Journal of Psychology, 15*(1), 72-101.

Spearman, C.E. (1914). The theory of two factors. *Psychological Review, 21*(2), 101-115.

Spearman, C.E. (1927). *The abilities of man: Their nature and measurement*. New York: Macmillan.

Spearman, C.E. (1939). Thurstone's work re-worked. *Journal of Educational Psychology, 30*, 1-16.

Spector, J.E. (1999). Precision of age norms in tests used to assess preschool children. *Psychology in the Schools, 36*(6), 459-471.

Spivack, G., and Shure, M.B. (1974). *Social adjustment of young children: A cognitive approach to solving real-life problems*. San Francisco: Jossey-Boss.

Spreat, S. (1999). Psychometric standards for adaptive behavior assessment. In R.L. Schalock (ed.), *Adaptive behavior and its measurement: Implications for the field of mental retardation* (pp. 103-108). Washington, DC: American Association on Mental Retardation.

Stapleton, D.C., Coleman, K.A., Dietrich, K.A., and Livermore, G.A. (1998). Econometric analysis of DI and SSI application and award growth. In K. Rupp, and D.C. Stapleton (eds.), *Growth in disability benefits: Explanations and policy implications* (pp. 34-80). Kalamazoo, MI: Upjohn Institute for Employment Research.

State, M.W., Dykens, E.M., Rosner, B., Martin, A., and King, B.H. (1999). Obsessive-compulsive symptoms in Prader-Willi and "Prader-Willi-like" patients. *Journal of the American Academy of Child and Adolescent Psychiatry, 38*(3), 329-334.

Sternberg, R.J. (1985). Human intelligence: The model is the message. *Science, 230,* 1111-1118.

Sternberg, R.J. (1986). *Intelligence applied: Understanding and increasing your intellectual skills.* New York: Harcourt Brace Jovanovich.

Sternberg, R.J. (1996). *Successful intelligence.* New York: Simon and Schuster.

Stewart, R.H. (1965). Effect of continuous responding on the order effect in personality impression formation. *Journal of Personality and Social Psychology, 1*(2), 161-165.

Stinnett, T.A. (1997). AAMR Adaptive Behavior Scale-School: Test review. *Journal of Psychoeducational Assessment, 15*(4), 361-372.

Stinnett, T.A., Havey, J.M., and Oehler-Stinnett, J. (1994). Current test usage by practicing school psychologists: A national survey. *Journal of Psychoeducational Assessment, 12*(4), 331-350.

Stogdill, R.M. (1948). Personal factors associated with leadership: A survey of the literature. *Journal of Psychology, 25*, 35-71.

Stone, W.L. (1997). Autism in infancy and early childhood. In D.J. Cohen, and F.R. Volkmar (eds.), *Handbook of autism and pervasive developmental disorders* (2nd ed., pp. 266-282). New York: Wiley.

Strauss, D.J., Shavelle, R.M., and Anderson, T.W. (1998). Life expectancy of children with cerebral palsy. *Pediatric Neurology, 18*(2), 143-149.

Streissguth, A.P., Barr, H.M., Kogan, J., and Bookstein, F.L. (1996). *Understanding the occurrence of secondary disabilities in clients with fetal alcohol syndrome (FAS) and fetal alcohol effects (FAE)*. Seattle, WA: University of Washington School of Medicine.

Sturmey, P., and Bertman, L.J. (1994). Validity of the Reiss Screen for Maladaptive Behavior. *American Journal on Mental Retardation, 99*(2), 201-206.

Sturmey, P., and Ley, T. (1990). The Psychopathology Instrument for Mentally Retarded Adults: Internal consistencies and relationship to behaviour problems. *British Journal of Psychiatry, 156*, 428-430.

Swets, J.A., Dawes, R.M., and Monahan, J. (2000). Psychological science can improve diagnostic decisions. *Psychological Science in the Public Interest, 1*(1), 1-26.

Switzky, H.N., Greenspan, S., and Granfield, J.M. (1996). Adaptive behavior, everyday intelligence and the constitutive definition of mental retardation. *Advances in Special Education, 10*, 1-24.

Szymanski, L.S. (1987). Prevention of psychosocial dysfunction in persons with mental retardation. *Mental Retardation, 25*(4), 215-218.

Teplin, S.W. (1983). Development of blind infants and children with retrolental fibroplasia: Implications for physicians. *Pediatrics, 71*(1), 6-12.

Terman, L.M. (1916). *The measurement of intelligence; An explanation of and a complete guide for the use of the Stanford revision and extension of the Binet-Simon Intelligence Scale*. New York: Houghton Mifflin.

Terman, L.M., and Merrill, M.A. (1937). *Measuring intelligence: A guide to the administration of the new revised Stanford-Binet tests of intelligence*. New York: Houghton Mifflin.

Terman, L.M., and Merrill, M.A. (1960). *Stanford-Binet Intelligence Scale; Manual for the third revision form L-M*. Boston: Houghton Mifflin.

Terman, L.M., and Merrill, M.A. (1973). *The Stanford-Binet Intelligence Scale: Manual for the third revision form L-M*. Chicago: Riverside.

Ticket to Work and Work Incentives Improvement Act of 1999. (1999). (PL 106-170).

Thompson, F.R., McGrew, K.S., and Bruininks, R.H. (1999). Adaptive and maladaptive behavior: Functional and structural characteristics. In R.L. Schalock (ed.), *Adaptive behavior and its measurement: Implications for the field of mental retardation* (pp. 15-42). Washington, DC: American Association on Mental Retardation.

Thompson, J.S. (2001). Identifying persons as having mental retardation: A simulation study. Unpublished paper written for the National Research Council, Committee on Disability Determination for Mental Retardation. Department of Psychology, University of California, Davis.

Thomson, G.H. (1951). *The factorial analysis of human ability* (5th ed.). London: University of London Press.

Thorndike, R.L. (1971). Concepts of culture-fairness. *Journal of Educational Measurement, 8*(2), 63-70.

Thorndike, R.L. (1986). The role of general ability in prediction. *Journal of Vocational Behavior, 29*(3), 332-339.

Thorndike, R.L., Hagen, E.P., and Sattler, J.M. (1986). *The Stanford-Binet Intelligence Scale* (4th ed.). Chicago: Riverside Publishing.

Thorndike, R.M., and Lohman, D.F. (1990). *A century of ability testing*. Chicago: Riverside Publishing.

Thurstone, L.L. (1938a). The perceptual factor. *Psychometrika, 3*, 1-17.

Thurstone, L.L. (1938b). Primary mental abilities. *Psychometric Monographs, 1*, ix-121.

Thurstone, L.L. (1947). *Multiple-factor analysis; A development and expansion of the vectors of the mind*. Chicago: University of Chicago Press.

Thurstone, L.L., and Thurstone, T.G. (1941). Factorial studies of intelligence. *Psychometric Monographs, 2*, 94.

Tiedeman, D.V. (1978). The eighth mental measurements yearbook. In O.K. Buros (ed.). Highland Park, NJ: Gryphon Press.

Towbin, K.E. (1997). Pervasive developmental disorder not otherwise specified. In D.J. Cohen, and F. R. Volkmar (eds.), *Handbook of autism and pervasive developmental disorders* (2nd ed., pp. 123-147). New York: Wiley.

Travarozzi, M. (2000). Private employers, disability management and return-to-work. Paper presented at the National Academy of Social Insurance Conference. L. Frieden, and J.L. Mashaw, (eds.), *Disability income policy: Opportunities and challenges in the next decade - A policy education seminar*. December 15, 2000.

Trower, P. (1982). Toward a generative model of social skills: A critique and synthesis. In J.P. Curran, and P.M. Monti (eds.), *Social skills training: A practical handbook for assessment and treatment* (pp. 399-427). New York: Guilford Press.

Tucker, L.R. (1940). The role of correlated factors in factor analysis. *Psychometrika, 5*, 141-152.

U.S. Department of Education. (1994). *The sixteenth annual report to Congress on the implementation of the Individuals with Disabilities Education Act*. Washington, DC: U.S. Government Printing Office.

U.S. Department of Education. (1999). *Twenty-first annual report to Congress on the implementation of the Individuals with Disabilities Education Act*. Washington, DC: U.S. Department of Education.

U.S. Department of Education. (2000). *To assure the free appropriate public education of all children with disabilities: Twenty-second annual report to Congress on the Implementation of the Education of the Individuals with Disabilities Education Act*. Washington, DC: Office of Special Education Programs.

U.S. Department of Health and Human Services. (2001). *Mental health: Culture, race, and ethnicity—A supplement to mental health: R report of the surgeon general*. Rockville, MD: Substance Abuse and Mental Health Services Administration, Center for Mental Health Services.

U.S. General Accounting Office. (1995). *Social Security: New functional assessments for* *children raise eligibility questions.* (Report No. GAO/HEHS-95-66). Washington, DC: U.S. General Accounting Office.

U.S. General Accounting Office. (1996). *SSA Disability: Program redesign necessary to encourage return to work.* (Report No. GAO/HEHS-96-66). Washington, DC: U.S. General Accounting Office.

U.S. General Accounting Office. (1997). *Social Security: Disability programs lag in promoting return to work.* (Report No. GAO/HEHS-97-46). Washington, DC: U.S. General Accounting Office.

U.S. General Accounting Office. (1999). *SSI children: Multiple factors affect families' costs for disability-related services.* (Report No. GAO/HEHS-99-99). Washington, DC: U.S. General Accounting Office.

U.S. General Accounting Office. (2000). *Medicaid managed care: Challenges in implementing safeguards for children with special needs.* (Report No. GAO/HEHS-00-37). Washington, DC: U.S. General Accounting Office.

U.S. Government Printing Office. (1918). *Examiners guide for psychological examining in the Army*. Washington, DC: U.S. Government Printing Office.

U.S. House of Representatives, Committee on Ways and Means, Subcommittee on Social Security. (1998). *Social Security Disability Insurance: Multiple factors affect beneficiaries' ability to return to work.* (Report No. GAO/HEHS 98-39). Washington, DC: U.S. General Accounting Office.

Utley, C.A., Lowitzer, A.C., and Baumeister, A.A. (1987). A comparison of AAMD's definition, eligibility criteria, and classification schemes with state department of education guidelines. *Education and Training in Mental Retardation, 22*(1), 35-43.

Valdivia, R. (1999) *The implications of culture on developmental delay* [Web Page]. URL http://ericec.org/digests/e589.html [2002, March].

Van Hasselt, V.B., Hersen, M., and Bellack, A.S. (1981). The validity of role play tests for assessing social skills in children. *Behavior Therapy, 12*, 202-216.

Vance, H.B., and Gaynor, P.E. (1976). A note on cultural difference as reflected in the Wechsler Intelligence Scale for Children. *Journal of Genetic Psychology, 129*(1), 171-172.

Venter, A., Lord, C., and Schopler, E. (1992). A follow-up study of high-functioning autistic children. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 33*(3), 489-507.

Vernon, P.E. (1950). *The structure of human abilities* (Methuen's manuals of modern psychology No. 1). London: Methuen.

Vernon, P.E. (1961). *The structure of human abilities* (2nd ed.). (Methuen's manuals of modern psychology). London: Methuen.

Vernon, S.W., and Roberts, R.E. (1982). Prevalence of treated and untreated psychiatric disorders in three ethnic groups. *Social Science and Medicine, 16*(17), 1575-1582.

Vining, E.P.G., and Freeman, J.M. (1996). Epilepsy and developmental disabilities. In A.J. Capute, and P.J. Accardo (eds.), *Developmental disabilities in infancy and childhood* (2nd ed., pp. 511-520). Baltimore, MD: Paul H. Brookes.

Voelker, S.L., Shore, D.L., Brown-More, C., Hill, L.C., Miller, L.T., and Perry, J. (1990). Validity of self-report of adaptive behavior skills by adults with mental retardation. *Mental Retardation, 28*(5), 305-309.

Volkmar, F.R., Cook, E.H., Pomeroy, J., Realmuto, G., and Tanguay, P. (1999). Practice parameters for the assessment and treatment of children, adolescents, and adults with autism and other pervasive developmental disorders. *Journal of the American Academy of Child and Adolescent Psychiatry, 38*(12, Suppl), 32S-54S.

Volkmar, F.R., Klin, A., and Cohen, D.J. (1997). Diagnosis and classification of autism and related conditions: Consensus and issues. In D.J. Cohen, and F.R. Volkmar (eds.), *Handbook of autism and pervasive developmental disorders* (2nd ed., pp. 5-40). New York, NY: Wiley.

Volkmar, F.R., et al. (1994). Field trial for autistic disorder in DSM-IV. *American Journal of Psychiatry, 151*(9), 1361-1367.

Walker, D., Greenwood, C.R., Hart, B., and Carta, J. (1994). Prediction of school outcomes based on early language production and socioeconomic factors. *Child Development, 65*, 606-621.

Waller, N.G., Thompson, J.S., and Wenk, E. (2000). Using IRT to separate measurement bias from true group differences on homogeneous and heterogeneous scales: An illustration with the MMPI. *Psychological Methods, 5*(1), 125-146.

Walls, R.T., and Werner, T.J. (1977). Vocational behavior checklists. *Mental Retardation, 15*(4), 30-35.

Warburg, M., Federiksen, P., and Rattleff, J. (1979). Blindness among 7,720 mentally retarded children in Denmark. *Clinical Developmental Medicine, 73*, 56-67.

Wasseman, J., and Bracken, B.A. (2002). Selecting appropriate tests: Psychometric and pragmatic considerations. In J. Carlson, and B. Waterman (eds.), *Social and personal assessment of school-aged children: Developing interventions for educational clinical settings*. Boston: Allyn and Bacon.

Watkins, C.E., Campbell, V.L., Nieberding, R., and Hallmark, R. (1995). Contemporary practice of psychological assessment by clinical psychologists. *Professional Psychology: Research and Practice, 26*(1), 54-60.

Watson, J.E., Aman, M.G., and Singh, N.N. (1988). The psychopathology instrument for mentally retarded adults: Psychometric characteristics, factor structure, and relationship to subject characteristics. *Research in Developmental Disabilities, 9*(3), 277-290.

Wechsler, D. (1939). *The measurement of adult intelligence*. Baltimore, MD: Williams and Wilkins.

Wechsler, D. (1949). *Wechsler Intelligence Scale for Children*. San Antonio, TX: Psychological Corporation.

Wechsler, D. (1955). *Wechsler Adult Intelligence Scale*. New York: Psychological Corporation.

Wechsler, D. (1967). *Wechsler Preschool and Primary Scale of Intelligence*. New York: Psychological Corporation.

Wechsler, D. (1974). *WISC-R, Wechsler Intelligence Scale for Children—Revised*. New York: Psychological Corporation.

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale-Revised (WAIS-R)*. New York: Psychological Corporation.

Wechsler, D. (1989). *WPPSI-R, Wechsler Preschool and Primary Scale of Intelligence—Revised*. San Antonio, TX: Psychological Corporation.

Wechsler, D. (1991). *WISC-III, Wechsler Intelligence Scale for Children* (3rd ed.). San Antonio, TX: Psychological Corporation.

Wechsler, D. (1999). *WechsleR Abbreviated Scale of Intelligence*. San Antonio, TX: Psychological Corporation.

Wehman, P. (1996). *Life beyond the classroom: Transition strategies for young people with disabilities* (2nd ed.). Baltimore, MD: Paul H. Brookes.

White, M.J., Nichols, C.N., Cook, R.S., Spengler, P.M., et al. (1995). Diagnostic overshadowing and mental retardation. *American Journal on Mental Retardation, 100*(3), 293-298.

Widaman, K.F., and McGrew, K.S. (1996). The structure of adaptive behavior. In J.W. Jacobson, and J.A. Mulick (eds.), *Manual of diagnosis and professional practice in mental retardation* (pp. 97-110). Washington, DC: American Psychological Association.

Widaman, K.F., Borthwick-Duffy, S.A., and Little, T.D. (1991). The structure and development of adaptive behaviors. In N.W. Bray (ed.), *International review of research in mental retardation* (Vol. 17, pp. 1-54). San Diego, CA: Academic Press.

Widaman, K.F., Gibbs, K.W., and Geary, D.C. (1987). Structure of adaptive behavior: I. Replication across fourteen samples of nonprofoundly mentally retarded people. *American Journal of Mental Deficiency, 91*(4), 348-360.

Widaman, K.F., Stacy, A.W., and Borthwick-Duffy, S.A. (1993). Construct validity of dimensions of adaptive behavior: A multitrait-multimethod evaluation. *American Journal on Mental Retardation, 98*(2), 219-234.

Wilbur, H. (1882). Some of the abnormal characteristics of idiocy and the methods adopted in obviating them. *Proceedings of the Association of Medical Officers of American Institutions for Idiotic and Feebleminded Persons,* (190-201).

Wilson, R.S. (1978). Synchronies in mental development: An epigenetic perspective. *Science, 202*(4371), 939-948.

Windle, M., and Blane, H.T. (1989). Cognitive ability and drinking behavior in a national sample of young adults. *Alcoholism: Clinical and Experimental Research, 13*(1), 43-48.

Witt, J.C., Elliott, S.N., Daly, E.J., Gresham, F.M., and Kramer, J.J. (1998). *Assessment of at-risk and special needs children* (2nd ed.). Boston: McGraw-Hill.

Wolber, G.J., Reynolds, B., Ehrmantraut, J.E., and Nelson, A.J. (1997). In search of a measure of intellectual functioning for an inpatient psychiatric population with low cognitive ability. *Psychiatric Rehabilitation Journal, 21*(1), 59-63.

Wolraich, M.L., Felice, M.E., and Drotar, D. (1996). *The classification of child and adolescent mental diagnoses in primary care: Diagnostic and Statistical Manual for Primary Care (DSM-PC) Child and Adolescent Version*. Elk Grove Village, IL: American Academy of Pediatrics.

Woodcock, R.W., Mather, N., McGrew, K.S., and Schrank, F.A. (2001). *Woodcock-Johnson III tests of cognitive abilities*. Itasca, IL: Riverside Publishing.

World Health Organization. (1988). *International Classification of Disease, 9th edition, Cinical Modification*. Ann Arbor, MI: Commission on Professional and Hospital Activities.

World Health Organization. (1993). *ICD-10 Classification of Mental and Behavioural Disorders: Diagnostic Criteria for Research*. Geneva, Switzerland: World Health Organization.

World Health Organization. (1996). *ICD-10 Guide for Mental Retardation*. Geneva, Switzerland: World Health Organization.

World Health Organization. (2000). *ICIDH-2: International Classification of Functioning, Disability, and Health—Pre-final version*. Geneva, Switzerland: World Health Organization.

Yamaki, K. *Employment and income status of adults with developmental disabilities living in the community*. Unpublished doctoral dissertation, University of Illinois at Chicago.

Yelowitz, A. (1998). The impact of health care costs and Medicaid on SSI participation. In K. Rupp, and D.C. Stapleton (eds.), *Growth in disability benefits: Explanations and policy implications*. Kalamazoo, MI: Upjohn Institute for Employment Research.

Yepsen, L.N. (1941). Defining mental deficiency. *American Journal of Mental Deficiency, 46*, 200-205.

Zhu, J., Chen, H., and Tulsky, D. (1999). WISC-III reliability data for special clinical groups. Paper presented at the Annual Convention of the National Association of School Psychologists. April, 1999.

Zigler, E., Balla, D.A., and Hodapp, R.M. (1984). On the definition and classification of mental retardation. *American Journal of Mental Deficiency, 89*(3), 215-230.

Zigman, W.B., Schupf, N., Sersen, E., and Silverman, W. (1996). Prevalence of dementia in adults with and without Down syndrome. *American Journal on Mental Retardation, 100*(4), 403-412.

Biographical Sketches

Daniel J. Reschly (*Chair*) is professor of education and psychology and chair of the Department of Special Education at Peabody College, Vanderbilt University. He has published more than 100 articles, chapters, and books on the topics of school psychology professional practices, mild mental retardation, the assessment of disabilities in minority children and youth, and legal issues in special education. He has served as an expert witness in various school-related court cases for the state of Iowa for over a decade. He was a member of the National Research Council's Committee on Goals 2000 and the Inclusion of Students with Disabilities and currently serves on the Committee on Representation of Minority Children in Special Education and Gifted Programs. He has a Ph.D. in school psychology from the University of Oregon.

Harolyn M. E. Belcher is currently a neurodevelopmental pediatrician at the Kennedy Krieger Institute and assistant professor in the

Department of Pediatrics at The Johns Hopkins University School of Medicine. She has served as chair on both the nominations subcommittee and the minority access to health care committee for the section on children with disabilities at the American Academy of Pediatrics. She has published extensively and continues to receive federal and state funding for work in the area of developmental disabilities in infancy and childhood and the effects of prenatal drug exposure on development. She has an M.D. from the College of Medicine at Howard University.

Jan Blacher is professor of education at the University of California, Riverside. Her research and writing has addressed family adjustment, the transition to adulthood, dual diagnosis, and out-of-home placement in white and Hispanic families. Her longitudinal research on families of children and young adults with severe disabilities spans 19 years and is funded by the National Institute on Child Health and Human Development. She has served as consulting editor for the *American Journal on Mental Retardation* and *Mental Retardation* and is frequently asked to appear as an expert in cases involving right to education or placement for children with autism, mental retardation, and developmental disabilities. She is a fellow of the American Association on Mental Retardation and Division 33 of the American Psychological Association. She has an A.B. in psychology from Brown University and a Ph.D. in special education/developmental psychology from the University of North Carolina, Chapel Hill.

Bruce A. Bracken is a professor in the School of Education at the College of William and Mary. He has written more than 100 articles and chapters on school assessments and has developed numerous behavioral, cognitive, and nonverbal intelligence assessment tools. He is the editor and cofounder of the *Journal of Psychoeducational Assessment* and has been a member of the American Psychological Association's Committee on Psychological Testing and Assessment. He is a member of the National Association of School Psychologists. He has a Ph.D. in educational psychology from the University of Georgia.

Elizabeth M. Dykens is associate professor in the Department of Psychiatry and Biobehavioral Sciences and associate director for training at the Mental Retardation Research Center and research and training at the Tarjan Center for Developmental Disabilities at the University of California, Los Angeles. Her research identifies the neurobehavioral and developmental phenotypes of individuals with genetic mental retardation syndromes. She examines genetic, psychosocial, and developmental sources of individual differences within syndromes. She has a Ph.D. in clinical psychology from the University of Kansas.

Christine R. Hartel (*Study Director*) is director of the Board on Behavioral, Cognitive, and Sensory Sciences at the National Research Council. Previously, she was associate executive director for science at the American Psychological Association. At the National Institute on Drug Abuse, she awarded grants and contracts for basic and clinical research on drug abuse in the fields of neuroscience, behavior, pharmacology, and chemistry. She has written many articles and edited two books on drug abuse and illicit drug demand reduction. As a research psychologist at the U.S. Army Research Institute for the Behavioral and Social Sciences, she received the army's highest civilian award for technical excellence. She is a fellow of the American Psychological Association and a member of the Society for Neuroscience. She has a Ph.D. in biopsychology from the University of Chicago.

Thomas Hehir is a lecturer in the Graduate School of Education at Harvard University. He served as director of the U.S. Department of Education's Office of Special Education Programs from 1993 to 1999. As director, he was responsible for federal leadership in implementing the Individuals with Disabilities Education Act (IDEA). Previously he was associate superintendent for the Chicago Public Schools, where he was responsible for special education services and student support services; he implemented major changes in the special education service delivery system. He has published articles on special education, special education in the reform movement, due process, and least restrictive environment issues. He has a Ph.D. in education from Harvard University.

John W. Jacobson is a planner and applied researcher in the New York State Office of Mental Retardation and Developmental Disabilities. He has published extensively on a diverse range of topics related to developmental disabilities, including population disability demographics, program evaluation, clinical services management and development, psychiatric disability, definition of developmental disabilities and mental retardation, adaptive development, and treatment efficacy. His current work focuses on the clinical and managerial issues in the delivery of individual support and housing services for people with developmental disabilities. He is also involved in training professionals and direct personnel on the impact of comorbid psychiatric diagnoses in individuals with mental retardation. He has a Ph.D. in general applied psychology and developmental disabilities across the life span from the University of Vermont, Burlington.

Bette R. Keltner is the dean of the School of Nursing at Georgetown University and codirector of the National Center for Family Diversity and Developmental Disabilities. Her research interests include families of children with mental retardation, the special needs of mothers with developmental disabilities, and rural family support for people with disabilities. She has published extensively in the area of culturally diverse families and developmental disabilities and has been the principal investigator on a number of federally funded studies on mental retardation and minority families facing challenging life circumstances. She has a Ph.D. in nursing from the University of Texas, Austin.

Marty W. Krauss is professor of social welfare and director of the Starr Center for Mental Retardation in the Heller Graduate School at Brandeis University. She has published widely in the area of life-span development of individuals with mental retardation and is particularly interested in the social support and coping styles of aging mothers of adult children with mental retardation and mental illness. She received the Future Leaders in Mental Retardation Award from the Joseph P. Kennedy, Jr., Foundation and is an elected fellow of the American Association on Mental Retardation. She has a Ph.D. in social welfare from Brandeis University.

Tracy G. Myers (*Senior Research Associate*) joined the National

Research Council to serve as senior research associate for the study. Prior to joining the National Research Council, he worked at Westat, Inc., on experimental and quasi-experimental evaluations of human services intervention programs targeted toward children and their families. His areas of interest include risk and protective factors that affect child and adolescent health and psychopathology, psychological assessment, minority mental health, and program evaluation. He has a Ph.D. in clinical psychology from the University of Maryland, College Park.

Hans Palmer is the W. M. Keck distinguished service professor and professor of economics at Pomona College where he served as Vice President for Academic Affairs and the Dean of the college from 1998-2001. He was the recipient of a National Science Foundation Faculty Fellowship to Oxford, England, where he worked with a health services evaluation group. He has also served as a visiting scholar at a number of American universities studying the economics of health care and aging. He has published on long-term care and on the effects of health care on the economy. His current research is examining home health care and service use by the elderly in community settings, the effects of medical research on the U.S. economy and on health insurance, and a grade of membership analysis of veterans with spinal cord impairments. He has a Ph.D. in economics from the University of California, Berkeley.

Sharon L. Ramey is the founding director of the Civitan International Research Center at the University of Alabama, Birmingham, and professor of psychiatry, psychology, pediatrics, neurobiology, sociology, maternal and child health, and nursing. She is a developmental psychologist whose professional interests include the study of intellectual and developmental disabilities, early experience and early intervention, the changing American family, and transition to school. She is actively engaged in research on American Indian families, strategies to enhance prenatal care for high-risk women, the prevention of developmental disabilities, and the effects of welfare reform on families and children. She is the author of numerous articles and three books, has made presentations at national and international conferences, and continues to receive substantial funding on developmental disabilities from federal and private agencies. She has a Ph.D. in developmental psychology from the University of Washington.

James Shanteau is professor of psychology and has served as director of the Institute for Social and Behavioral Research at Kansas State University. He is former program director of the decision, risk, and management science program at the National Science Foundation. He has published numerous articles and books in the areas of judgment and decision making behavior, expert decision making, and the evaluation of psychological measurement and model testing techniques. He has received funding from the National Science Foundation, the Department of Transportation, and the National Institutes of Health for his research related to judgment theory and decision analysis. He has a Ph.D. in experimental psychology from the University of California, San Diego.

Wayne P. Silverman is the chief research scientist and chairman of the Department of Psychology of the New York State Institute for Basic Research in Developmental Disabilities. He has received multiple grants from the National Institutes of Health and has contributed to numerous publications on topics ranging from basic studies of cognition to applied studies of the characteristics and service needs of people with profound mental retardation and multiple disabilities. He has collaborated with colleagues working in biomedical research for many years and now directs a large multidisciplinary program project that is studying the effects of aging and Alzheimer's disease on the cognitive, functional, and health status of adults with mental retardation. He has a Ph.D. in cognitive psychology from the State University of New York at Buffalo.

Gary N. Siperstein is a professor in the Human Services Center and the director of the Center for Social Development and Education at the University of Massachusetts, Boston. He has carried out federally funded research for the past 25 years in the area of mental retardation, with a specific focus on the social aspects of the disability. His research on the social acceptance of children with disabilities in the classroom has been carried out both nationally and internationally. He has published numerous articles and chapters on mental retardation and related disabilities. More recently, he and his colleagues have written on the issues of diagnosis and classification of individuals with mental retardation. He has a Ph.D. in developmental psychology and special education from the Ferkauf Graduate School of Psychology at Yeshiva University.

Sara S. Sparrow is professor of psychology and the chief psychologist at the Yale University Child Study Center. Since receiving a Ph.D. in clinical psychology and neuropsychology in 1968, she has published numerous articles and book chapters and presented many papers on cognitive and behavioral assessments in children with developmental learning disabilities and mental retardation. She has been an editor and a consulting reviewer for numerous journals, including the *Journal of Clinical Neuropsychology*, the *Journal of Abnormal Psychology*, *Developmental Psychology,* and the *Journal of the American Academy of Child Psychiatry*. She also has conducted numerous workshops on the administration and interpretation of the Vineland Adaptive Behavior Scales with community and state agencies that determine disabilities associated with mental retardation.

Keith F. Widaman is professor of psychology at the University of California, Davis, and the director of its Center for Quantitative Social Science Research. He has received numerous awards and honors for his research and teaching activities in the area of developmental psychology and has published numerous refereed articles and book chapters. He has served on the editorial board of numerous journals on mental retardation, psychological measurement, and statistical methods, including *Psychological Methods* and the *American Journal on Mental Retardation*. He has been actively involved in research on lifespan development of individuals with mental retardation for over 20 years and has been awarded various federal grants from the National Institute for Child Health and Human Development and the U.S. Department of Education. He has a Ph.D. in developmental psychology from The Ohio State University.

Index